Walter Frenz *Editor*

Handbook Industry 4.0

Law, Technology, Society

Springer

Handbook Industry 4.0

Walter Frenz
Editor

Handbook Industry 4.0

Law, Technology, Society

Springer

Editor
Walter Frenz
Unit for Mining, Environmental and
European Law
RWTH Aachen University
Aachen, Germany

ISBN 978-3-662-64447-8 ISBN 978-3-662-64448-5 (eBook)
https://doi.org/10.1007/978-3-662-64448-5

This Springer imprint is published by the registered company Springer-Verlag GmbH, DE part of Springer Nature.
The registered company address is: Heidelberger Platz 3, 14197 Berlin, Germany

Foreword by the former Minister President of the State of North Rhine-Westphalia, Armin Laschet

The term “Industry 4.0” first caught on in Germany at the Hannover Messe show in 2011 and has since become the international catchphrase for the ongoing transformation of industry and manufacturing by smart technologies. This may be because the term serves two purposes: it refers to the fourth industrial revolution currently underway—following on from mechanisation, electrification and automatisation—and is also a reminder, through its adoption of the software naming format, of what is driving this latest chapter, namely the digital transformation of industrial production.

Nowadays we associate industrial production with high-tech companies, modern, well-paid jobs, excellent training and great opportunities for young people. Industry in the Germany of today—as in many other countries around the world—is about climate protection, environmental standards and, not least, widespread social engagement.

That said, digitalisation is proving a great challenge and seriously questioning the viability of some existing business models. This applies all the more in times of a global pandemic. Production must be flexibly converted, value chains changed, processes adapted. Companies must demonstrate their capacity to innovate and transform their operations.

However, digitalisation and increased flexibility also provide great opportunities. We will be well placed to combine our strong manufacturing tradition with the possibilities afforded by Industry 4.0. The prospects will be excellent

- for the construction of machines that detect when their own parts need to be replaced
- for the intelligent networking of machines and processes using information and communication technology. This opens up new horizons to companies in the form of flexible production, modifiable factories, optimised logistics, closed-loop processes that minimise waste, etc.
- and for companies that can embrace highly flexible processes and expand their manufacture of individualised products. Customers and business partners are

becoming part of commercial and creative processes that can be controlled and optimised almost in real time

When people think of the Internet, they tend to conjure up images of Google, Facebook or Amazon because these major US-based platforms dominate the "consumer Internet." But another Internet, the "Internet of Things," is emerging. The industrial DNA of the German economy puts us in a strong position to lead the way in this emerging "industrial Internet."

While the systemic shifts inherent to Industry 4.0 have long since made themselves felt in a large number of companies, many small and medium-sized firms are still tentative in their approach. This is the perfect time, then, for the release of this Guide, which not only provides an overview of Industry 4.0 but also sets about applying practical solutions to real-life challenges.

For all the evocativeness of the term "Industry 4.0," it is in danger of being narrowly associated with manufacturing. In reality, the challenges posed by digitalisation are wider and require answers to a whole raft of questions: How will digital technologies change our lives and work practices? How are we adapting our education, training and qualification systems? How can we protect our data and safeguard our infrastructure? What new legislation is needed? Which norms and standards should be adopted?

This Handbook avoids the pitfall of taking too narrow a view, illuminating instead the Industry 4.0 theme in all its facets. For this reason, I commend the book and hope it attracts a large readership. The Guide sets out the newest scientific advances that can be implemented by firms today, findings that can be harnessed to bring about the digital transformation of established business models and develop new goods and services. Coming from North Rhine-Westphalia, the industrial heartland of Germany, I hope that it can serve as reference book for the challenges that lie ahead when we set about to shape the fourth industrial revolution globally.

North Rhine-Westphalia, Germany Armin Laschet

Foreword by the former Minister of State, Dorothee Bär

Industrie 4.0: A New Understanding of the Role Played by Humans and Technology

When we talk about processes of digitalisation, about the transition to the high-tech age, we use a wide range of concepts that seek to capture the full extent of these changes. We talk about revolution, about disruption and challenges, the likes of which we have not had to grapple with since the eighteenth century.

In addition to this, there are concepts whose potential significance we do not see at first sight, but which stand for probably the most far-reaching developments that we have observed in the context of technological progress as they not only require greater knowledge and skills, but also a reappraisal of our own basic attitude to specific circumstances.

One of these concepts is "Industrie 4.0." What at first appears to be a large but ultimately well-defined area turns out on closer inspection to be a sector that directly or indirectly affects almost all people in our society, thus reflecting the ubiquity of the digital transformation.

One reason for this is that a large part of the developments we are talking about in connection with Industrie 4.0 not only affect major companies, but above all small and medium-sized enterprises, which today more than ever are the backbone of our entrepreneurial world, thus making themselves felt in all conceivable areas of people's lives.

On the other hand, however, this is also due to the fact that all the products made by these companies, both large and small, are central to the way in which all of us lead our lives, opening doors and creating opportunities whose diversity and long-term significance have been much less pronounced to date. This applies both to our daily working environment as well as to our private lives. In the course of the increasing scope for individual development, the latter is at least as important as the former.

Making this clear to members of the public, and above all demonstrating this to the many entrepreneurs in a very tangible way, without arousing scepticism or even fear, is a very important factor on the path to the near and distant future, along with a comprehensive and constantly expanding and improving infrastructure and the necessary legal framework conditions.

It is increasingly important in this regard not only to talk in terms of theoretical visions of the future, but also to show in a most concrete way what changes industry and business are experiencing in a digitalised world and what technological progress in individual areas, industries and sectors will entail. Now that the digital transformation has become established as an unstoppable and promising development in the social and political discourse in general, it is now time at long last to shape the developments it involves in a tangible manner and to translate plans into action.

To this end, we must first remind ourselves that a vision is always preceded by a will to set out on a journey, which in turn is followed by acquiring suitable equipment. If you want to climb Everest, you must first really want to do so and decide to set off having made suitable plans and having acquired the necessary knowledge about the risks and the best routes. Once the plan has been put together, the next step is to get hold of the correct equipment and to learn the skills to use it in the right way.

For Industrie 4.0, this involves understanding what the term actually means. We must be aware that is not just about the technological further development of specific machine processes, but that this also, primarily, necessitates a fundamental recasting of the relationship between humans and machines. If, in the field of the Internet of Things, technology interacts with technology and this happens more and more independently, this requires, to some extent, a new understanding of the role of humans within our traditional communication model.

This example of machine-to-machine communication in particular thus shows very clearly that we must not only be mindful not to lose control over technology, but that we can only prevent this loss of control if we continue to fully understand this exchange of information and data.

In the future, refusing to get involved in certain developments will therefore no longer mean simply staying on the sidelines or being left behind, but can very quickly lead to a loss of human freedom in connection with making decisions about technical processes and actions.

Applied to small and medium-sized enterprises in particular, this means creating the necessary awareness of new developments and, to this end, joining forces with the research sector, for example.

Centres of excellence and test environments must therefore be created to simulate and advance processes and developments under realistic conditions. Once the latter has been accomplished, dialogue must take place to share results and experiences and make them available to as many people as possible.

In addition, there must be a regular and intensive dialogue between the worlds of politics, business and trade unions with regard to norms, standards and legal frameworks, and all those involved must be consulted when it comes to shaping life and work in a modern country in a globalised world in such a way that what is good is preserved and what is bad is improved. Nobody should then have to fear the future, and all of us must work to relegate social decline and the real or perceived loss of security to the realm of literary dystopia. These things should no longer be a feature of the evening news.

So, the great field of Industrie 4.0 is supported by several major pillars at the end of the day. This is about IT architectures and the standards they entail.

This is a question of reliable IT security, such as the security of data and business secrets, the protection of which, in turn, is a foundation of trust and a basic prerequisite for people to engage in technological and process innovations. Moreover, this is also about the absolutely vital qualification of people, about knowledge transfer, skills development and prevention concepts, which must be developed and implemented together with specialists, companies and social partners.

Industrie 4.0—this at first glance somewhat vague term stands, on closer inspection, for the gigantic spectrum of modern-day progress. After all, it not only comprises modern manufacturing processes and new supply chains. This term not only stands for innovative production methods and a hitherto unprecedented interconnection between man and machine or the all but autonomous communication between one machine and another.

Rather, it stands for a new understanding of the role played by humans and technology. It stands for a society that has to rethink familiar things in a challenging way owing to technological progress, but which is given the opportunity thanks to the developments of the digital transformation to achieve and expand general prosperity, as well as a responsible approach to our environment and social security. The latter three aspects form the three essential foundations of democracy and tolerant coexistence.

German Bundestag, Berlin, Germany Dorothee Bär

Foreword by the rector of the RWTH Aachen University, Ulrich Rüdiger

When this book was published, the COVID-19 pandemic was acting as a catalyst for digitisation in a wide variety. At the RWTH Aachen University, it also meant that we had to go through a restructuring process of our fields of work in the areas of research, teaching and administration in a very short time. In this process, the university pursued the goal of keeping various projects running without any losses. We were only able to do this by resorting to digital solutions—much faster than originally expected. Of course, even before the pandemic, the networking and digitisation of the world was developing rapidly and in all areas of life to make processes faster, smarter and easier.

Within the last decade, the term "Industry 4.0" was coined and the sphere of influence of digitisation has unfolded such potentials that it would be negligent to leave them unobserved and, above all, unused. It is of utmost importance for the economic and scientific development of our country to investigate the opportunities and obligations associated with the networking of people, machines and products, as well as the associated digitisation processes. As one of the leading technical universities in Europe, the RWTH Aachen University is at the forefront of researching and developing this topic in an application-oriented way. It is very important and helpful that this book looks at the topic of Industry 4.0 from the perspectives of a wide variety of disciplines. After all, interdisciplinary research is the future.

During the last years, the RWTH Aachen University has developed into an "Integrated Interdisciplinary University for Science and Technology." In our concept for the German Excellence Strategy, we demonstrate through our motto "Knowledge. Impact. Networks." that the university strives for the convergence of knowledge, methods and findings from the humanities to business and engineering, to medicine, and to the natural and life sciences to penetrate and further develop complex systems. In the interest of exploiting the potential of Industry 4.0, eradicate the dangers and provide tailored assistance to as many companies as possible, there are already many research projects on these aspects at the RWTH Aachen University—for example on digital twins, connectivity, 5G networks, networks, adaptive production, predictive quality and many more topics. These research areas are of course not detached from each other, but they can rather learn from each other and build on each other to achieve the best possible research results. We conduct this research to be able to design value chains more efficiently and to help industrial companies to act in a more resource-conserving way. In this regard, the RWTH Campus GmbH plays an important role as a unique infrastructure for cooperation between the university and industry by tackling the major challenge in creating an innovation-driven deep tech ecosystem.

One of the RWTH Aachen University's Clusters of Excellence is called "Internet of Production." The Internet of Production is often described as the centerpiece of Industry 4.0. It offers real-time information availability of relevant data. In this cluster, more than 25 institutes and research facilities have joined forces to develop the future of digitalisation in production in Aachen. Scientists from the fields of production technology, materials science and economics, as well as industrial engineering and ergonomics, psychology and, in particular, computer science, are tackling these challenges in an interdisciplinary manner to pave the way for the digital age of production.

Many of the scientists involved have contributed articles to this book, and so is the interdisciplinary, integrated research approach of RWTH Aachen University also reflected in this compendium. Experts from different disciplines examine the topic of Industry 4.0 to obtain an overall view, and to illuminate the future-oriented fields of action in a holistic manner.

I am very grateful for the great contributions to this book, and especially for the effort that the editor has undertaken. The book examines the complex topic from various angles, and I hope that it will provide in-depth support for further research on the topic of Industry 4.0.

RWTH Aachen University, Aachen, Germany

Ulrich Rüdiger

Preface

Digitisation is advancing with mighty steps and not only brings about numerous technical innovations, but also presents law and society with immense challenges. This is why this handbook Industry 4.0 covers all three areas together. The legal developments are currently of particular importance. The commission presented principles for business-to-business data sharing and provides a transparent model with mutual respect and shared value creation. The EU level therefore also plays a decisive role for digitisation—as does the federal and state levels—with corresponding contributions in this volume from the former Minister of State for Digital Bär as well as from the former Prime Minister of North Rhine-Westphalia Laschet and State Minister of Justice of North Rhine-Westphalia Biesenbach.

The Commission paper does not provide an answer to the question of data ownership—or should there not be such a thing at all? These and other questions (contract and liability law, intellectual property and copyright law, competition law, administrative and procurement law, data protection, etc.) are dealt with by numerous well-known lawyers from various fields of practice. Data security is of decisive importance.

The technical developments and fields of application of Industry 4.0 are diverse. They range from the Internet of Production, robotics and artificial intelligence, each with characteristics in manufacturing technology, mechanics and mechanical engineering (digital shadows) through electromobility and autonomous driving (including legal issues) to climate protection and energy technology/economy, infrastructure and medicine. Above all, colleagues from the RWTH Aachen University, but also from various other domestic and foreign universities, documented the state of research and development and pointed out perspectives.

Ethical questions arise, especially for autonomous driving and healthcare (big data-based), but also in finance, for example. Furthermore, the whole of society is being gripped by digitisation. This applies above all to the digital transformation of companies as well as universities and the world of work, as well as gender. There are drastic consequences, which are presented by well-known experts from a wide variety of perspectives. The humanities are also affected by digitisation (digital humanities, computational social science).

In conclusion, the handbook Industry 4.0 comprises over 66 articles that cover the whole range of relevant fields of digitisation and together illuminate all three dimensions: law, technology and society. I would like to thank all the authors very much for their content-rich, exciting contributions, which enable an up-to-date, interdisciplinary overview of the entire spectrum of Industry 4.0. My special thanks go to Dr. Kristina Fischer M.A., Ms. Désirée Dietrich B.A. and Ms. Elena Schiwinger, who supported me in my editorial work, as well as Mr. Michael Quandt, who produced various translations. Dr. Brigitte Reschke from Springer Nature was happy to take up my idea for this handbook and accompanied the development with great interest.

Please submit suggestions to Univ. Prof. Dr. jur. Walter Frenz, RWTH Aachen University, Wüllnerstraße 2, 52062 Aachen, frenz@bur.rwth-aachen.de.

Aachen, Germany Walter Frenz
1 April 2022

Contents

Part I
Legal Aspects

Aspects of the Digital Transformation of the Judiciary

Peter Biesenbach

Contents

P. Biesenbach (✉)
Ministry of Justice of the State of North Rhine-Westphalia, Minister of Justice, Düsseldorf, Germany
e-mail: peter.biesenbach@landtag.nrw.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_1

Digital change is the topic of our time. Digitization is changing our lives—comparable to the industrial revolution from the middle of the nineteenth century—at breathtaking speed. It covers all areas of life, including the judiciary.

1 Current State of Digitization and Perspectives on Technical IT Equipment

1.1 Legal Framework

The legal framework for the digitization of the judiciary can essentially be found in three **federal laws**. Due to the Law to promote electronic legal transactions with the courts of October 10, 2013 (the "**eJustice Law**"), the opening of **electronic legal transactions** in proceedings under the German ZPO (Code of Civil Procedure), FamFG (Act on Court Procedure in Family Matters and Nonlitigious Matters), ArbGG (Labor Court Act), SGG (Social Court Act), VwGO (Administrative Court Code), and FGO (Tax Court Rules) was regulated. The Law on the introduction of electronic files in the judiciary and for the further promotion of electronic legal transactions of July 5, 2017, regulates the opening of electronic legal transactions in criminal matters and the obligation to keep electronic files from January 1, 2026. Finally, electronic legal transactions for communication with bailiffs were opened by the Law implementing Regulation (EU) No. 655/2014 and amending other civil procedural, land register and property law provisions and amending the judicial collection regulations of November 21, 2016 (EuKoPfVODG).

The following time frames result from these legal framework conditions.

On January 1, 2018, facultative electronic legal transactions were set to be opened nationwide. Since then, documents could in principle also be submitted in **electronic form** to all courts, public prosecutors, and bailiffs. At the same time, a passive obligation to use the special **attorney mailbox** (beA) for the legal profession was standardized. North Rhine-Westphalia has not made use of the legally provided opt-out option (the federal states were able to postpone the legal opening of electronic legal transactions by an ordinance in whole or in part until December 31, 2019).

By January 1, 2020, at the latest, electronic legal communication also had to be opened by those federal states that have made use of the opt-out option. At the same time, there is an opt-in option for all countries from this point in time to oblige lawyers, authorities, and legal entities under public law to actively use electronic legal transactions by means of a legal ordinance.

A nationwide obligation of "professional submitters" (lawyers, legal entities under public law, etc.) to participate in electronic legal transactions exists from January 1, 2022. Eventually, the judiciary is obliged to manage files electronically from January 1, 2026.

1.2 State of Digitization in the Judiciary in North Rhine-Westphalia

1.2.1 Opening of the Electronic Legal Communication

As stipulated by law, the optional **electronic legal communication** in the judiciary of the state of North Rhine-Westphalia shall be opened nationwide for proceedings under the ZPO, FamFG, ArbGG, SGG VwGO, FGO, and StPO (Criminal Procedure Code) and directly with the bailiffs on January 1, 2018. Starting in the first quarter of 2019, the sending of electronic documents by the judiciary should be enabled and tested in practice. This possibility is to be extended to all higher regional courts, regional courts, the public prosecutor's offices, as well as some larger local courts in the further course of the year.

In view of the fact that the productive operation of the special **electronic attorney's mailbox** (beA) could not commence on January 1, 2018, as provided by law and the Federal Bar Association, but with a considerable delay on September 3, 2018, due to security problems and the lack of adoption by registered lawyers, the courts and public prosecutors would initially have to be cautious about sending documents electronically across the board. Hence, the first consequence for the courts and public prosecutor's offices is the necessary, labor-intensive digitization of paper receipts at a high level. Here, it is important for everyone involved to advertise intensively and continuously for the use of the beA by the legal profession.

1.2.2 Implementation of the Electronic File

According to legal requirements, all courts and public prosecutor's offices must introduce **electronic files** across the board by 2026, which would enable continuous electronic processing from the receipt of a document through to the handling of the case to the comprehensive delivery of documents.

1.2.2.1 IT Centralization as a Prerequisite for Electronic Legal Communication and Electronic Files

The judiciary of the state of North Rhine-Westphalia intends—for technical and organizational reasons and also because of corresponding suggestions from the State Audit Office for economic reasons—to implement the nationwide introduction of electronic files on the basis of **centralized information technology (IT) operations**.

The central IT operating office of the judiciary (ZBS), which was set up for this purpose in Münster, was put into operation in August 2016 with the start of the productive operation for the so-called EHUG (German Law on Electronic Commercial and Company Registers) proceedings at the Bonn Regional Court. In January 2017, the first pilot court, the Krefeld Regional Court, was successfully transferred to the ZBS with all of its data. The first specialized court followed in April 2017, the Düsseldorf Social Court, which also introduced a new specialized procedure (EUREKA-Fach). After the social justice system, the centralization of the courts was also started in March 2018. In addition, the jurisdiction areas of a first higher regional court (Hamm), the first local courts, and the first public prosecutor's offices have been transferred to the ZBS on a pilot basis.

As of January 15, 2019, the IT operations of all regional courts, six local courts, the Hamm Higher Regional Court, three public prosecutor's offices, 21 labor courts, and six social courts were centralized. According to the current plans, IT centralization should be completed by the end of 2021.

1.2.2.2 Current State of Electronic Filing

Electronic filing is currently being tested in civil matters at the Hamm Higher Regional Court, six regional courts (Bochum, Bielefeld, Detmold, Hagen, Krefeld, and Bonn), and two local courts. For several chambers of all regional courts involved, electronic filing has now even been mandated by statutory order. Therefore, there is no longer any parallel paper-based file management.

This leading electronic file is also being tested in practice in numerous **arbitration bodies** in all of the state's financial courts, the Münster Higher Administrative Court, and the Minden Administrative Court.

Further plans envisage the **successive introduction** of electronic files in other departments of ordinary jurisdiction, the public prosecutor's offices, and the specialized courts, which, according to current expectations, should be fully completed by the beginning of 2025.

1.2.3 Outlook on Factual IT Equipment

After the experience gained so far from the piloting with the electronic file and the centralized IT operation in the ZBS Münster, there is a need for additional equipping of the workplaces for the employees compared to the previous IT equipment.

In order to enable **ergonomic work**, all workplaces of the nonjudicial services should be equipped with additional monitors (i.e., a total of two screens), as well as swivel arms, signature cards, and card readers. In order to reconcile family and work, it might also be desirable to equip a larger number of nonjudicial employees with a mobile device in order to accommodate the additional possibility of **working from home**. The prerequisite for this—as in principle for all material equipment—is a sufficient financial leeway, which can only be specified by the budget legislator.

Since judges are constitutionally guaranteed the right to work regardless of location but the use of private devices is not an option for reasons of IT security, they must also be equipped with **mobile devices** without any preconditions. In order to meet the individual requirements of judges for ergonomic work, in particular for reading and processing the most extensive electronic files, in some cases, this group of people should be offered several equipment variants.

The introduction of electronic files into the judiciary also makes it necessary to equip the courtrooms accordingly. The idea is to provide a touch monitor and a standard personal computer (PC) for each judge on the bench. In addition, depending on the size of the room, large wall monitors or projectors should be installed to keep the meeting open to the audience.

1.3 Means and Channels of Communication

Internal communication in the judiciary will probably continue to be dominated by telephones and e-mails in the next few years. Nevertheless, new and more modern communication options could significantly improve cooperation in the judiciary, both ergonomically and functionally.

The current telephone system of the judiciary, which is based on ALLIP, offers improved possibilities for "hands-free" telephoning and for conducting **telephone conferences**. In addition, video conference systems will increasingly be used to support business meetings and avoid business trips.

A future linking of all **exchange systems** of the judiciary for the purpose of nationwide coordination of appointments based on all the calendars kept in the judiciary will facilitate the organization of the work. **Remote maintenance systems** (Remote Assistance) will be used to connect experts for troubleshooting or technical support on the use of IT systems.

2 Organizational Questions

The digitization of the judiciary will change existing jobs in the long term because in the future digital world, files can be accessed from any workstation and simultaneously. The organizational processes existing up to now, among other things, involve the need to transport files, which will no longer be necessary in the future, wherein faster and parallel work will be possible. Nevertheless, the established and proven organizational processes of the judiciary can also be transferred to the digital world to a large extent. The e^2A application, which is used as an electronic file in the judiciary and can therefore be mentioned as *pars pro toto*, largely reproduces the work steps that have been tried and tested in the paper file and are the result of decades of experience. Currently, the focus is on how to cope with the considerable challenges involved in this conversion phase (keyword: "media disruption").

In addition to the effects on judicial equipment and spatial arrangements, it must be taken into account the extent to which the **work processes** will change after the introduction of the electronic file and whether the digitization will result in new tasks for the employees or, if necessary, the elimination of obsolete tasks. This can only be conclusively assessed in full when the electronic file has been introduced across the board in the work area. At the current early stage, the experience obtained from the **piloting courts** does not yet allow a reliable assessment of whether fundamental organizational adjustments are useful or necessary. It can be assumed that further knowledge can be gathered as the piloting stage progresses. To this end, there is an intensive exchange of communication with the pilot authorities. The design of the work processes will also be considered beyond the already implemented electronic accessibility of the courts, in particular from the perspective of the **citizen seeking legal protection**, although the necessary formal rigor and **legal certainty** of judicial action must not set **unrealistic expectations** of the judiciary.

Various sets of rules are already being continuously checked and adapted to the digital world and the new challenges it poses. Under local leadership, a **nationwide working group** is focusing on regulations, some of which date back to the 1930s, with the aim of making them applicable to both paper and electronic files. In addition, the first changes to the rules of procedure and the detailed provisions for handling incoming and outgoing mails in the case of electronic legal transactions have already been made.

The content-related effects of digitization on work results are of fundamental importance but may still be difficult to predict. The judicial activity is not characterized by the completion of uniform business processes from the point of **view of efficiency**, but characterized by the careful and in-depth examination of often atypical facts. It is important for the citizens and the economy that the high quality of case law be maintained and to furthermore strengthen it. Electronic files offer a wide range of options for accessing documents and **structuring information**. This provides opportunities for the decision-makers to organize the respective lecture as effectively and correctly as possible, especially with regard to particularly complex procedures. The shift of work areas to digital processes and the increasing flexibility of work also require a high degree of adaptability. The wealth of information has to be processed intellectually. Therefore, the acceleration and automation of work processes must not **increase stress factor levels**. Rather, they must be perceived as a relief. Because of these factors, health management is facing new challenges through digitization.

In terms of content, there may also be opportunities in the future to readjust judicial tasks, legal clerk tasks, and the tasks of office clerks. A comprehensive analysis of the possible assignments of tasks from all branches of services is also carried out in consideration of this. As an example, the registry rules for the courts and public prosecutor's offices of the state of North Rhine-Westphalia were amended to enable civil servants of the police sergeant to convert a handwritten law enforcement or judicial document correspondingly signed by responsible persons to a replacement electronic document, including the qualified electronic signatures.

If, after the widespread introduction of electronic files, it becomes apparent that work processes and workflows are changing, it will also have to be examined to what extent these changes are so fundamental that they have **an impact on the job**

profiles in the judiciary and according adjustments to training and further education are necessary.

With the advancement of digitization and the widespread introduction of electronic files, new possibilities emerge with regard to the temporal and local conditions of work. Digitization makes it possible to fundamentally rethink whether there's a need to be present in the workplace (keyword: **telecommuting**). Once the technical prerequisites for working at home have been met, in all branches of services, the questions will arise as to the extent of the flexibility and in which cases presence at the workplace would be necessary. Here, especially the aspect of the **compatibility of work and family or work and care** and thus the question of the attractiveness of the judiciary as a modern employer speak in favor of increasing flexibility in terms of time and location. While changes in attendance in the judicial service are likely to show up automatically and immediately with the introduction of the electronic file, there will be a need for action in terms of **enabling telecommuting** among all other branches of services. Among others, for reasons of employee motivation and satisfaction, it will have to be checked whether and to what extent the flexibilization of the workplace is possible for everyone and which factually compelling areas require boundaries.

It is therefore important to consider the effects on the presence culture with regard to the perception of the judiciary for the citizens of North Rhine-Westphalia, but also to the professional exchange and social interaction between employees in all branches of service. The exchange of information on a collegial level is of fundamental importance for effective cooperation and, thus, for sufficient task completion. At the same time, it helps prevent misunderstandings or conflicts and creates a good **working atmosphere**. A good culture of discussion can also ensure that the so-called institutional knowledge is retained over the long term. The natural transfer of knowledge between the employees of a court or an authority must also be ensured during the digital transformation. Even though the fundamental importance of physical presence and personal exchange for effective work in the judiciary cannot be neglected even in these times of digital working life, it will have to be examined the extent to which offers of social and professional exchange can also be made digitally available (keyword: new **forms of communication**).

For these reasons, too, communication (which has always been important) with regular institutionalized discussions is becoming even more crucial in these times of digital change.

Overall, the judiciary in North Rhine-Westphalia is currently undergoing a change of dogma, which affects many business processes that have been handed down for decades—the "judicial culture," so to speak—and makes the introduction of complex, especially very technical, processes necessary. All **career groups** are affected by these changes. This fundamental change in almost all business processes should be accompanied not only by acceptance management during the term of the project as well as employee training and further education measures but also by comprehensive further training of managers since the success of this process crucially depends not only on the competence of employees but also on how they are being motivated. The executives and **junior executives** are of particular importance on this matter.

3 Legal Policy Engagement of North Rhine-Westphalia's "Working Group Digital New Start"

3.1 Reliable Legal Framework?

The importance of digitization for the state, society, and economy will grow faster over time. Nevertheless, the associated opportunities for companies and the society are not just faced with technical and economic challenges. A digital society also needs a reliable legal framework so that **freedom, equality, democracy,** and **justice** are preserved.

3.1.1 Does Our Civil Code Need an "Update"?

In view of the rapid digitization of our everyday life, the question arises whether the applicable law meets the new requirements. In a nutshell, does our civil code need an "update?"

Despite multiple reforms and changes, one looks in vain for the regulation of many digital processes in the code of law that came into force on January 1, 1900. However, digital processes have become a matter of course for all of us and happen thousands of times every day. In fact, lawyers, judges, corporate lawyers, and all other practitioners have to resolve issues from the digital world with regulations from the analog age on a daily basis.

This mostly works surprisingly well. Nevertheless, a thorough examination of this topic had been called for, and so the Conference of the Justice Ministers of the states, upon the initiative of North Rhine-Westphalia, set up the working group "Digital New Start" under our leadership in June 2015. A large number of federal states and also the Federal Ministry of Justice and Consumer Protection participated. After an extensive examination of the question of whether there is a need for legislative action in civil law in the course of judicial digitization, the working group issued two reports: the first one with more than 400 pages on the **Justice Ministers Conference** in June 2017 and a second one on the Autumn Conference 2018. Both reports, as well as a third report on the Spring Conference of the Justice Ministers in June 2019, are available on the Internet (at https://www.justiz.nrw) under the title of the working group (Digital New Start).

The deliberations of the working group are based on the principle that there is no need for legislative action as long as the applicable law provides viable norms. If this applies, it shall be left to the courts to find appropriate solutions subsumed under existing norms. Furthermore, any need for **legislative action** should primarily be taken into account by supplementing existing regulations with special publications, if necessary. The stated objective was to avoid further fragmentation of civil law as much as possible.

The report of the working group, published in 2017, deals with two subject areas that are less relevant to the keyword "Industry 4.0"—namely, "**digital personal rights**" and "**digital inheritance**"—among others with the following topics and questions.

3.1.2 Does the Legal Quality of Digital Data Require a Legal Provision?

The working group investigated the question of whether the legal quality of digital data should be determined by law, for example, by creating an exclusive right. Data can have great emotional, for companies also existential, importance as well as a high economic value. Trading data is commonplace. Legal laypersons have the tendency to think that data "belong" to them. However, the applicable law does not recognize data ownership or any other absolute right to digital data. Rather, data enjoy legal protection through a variety of different approaches. For example, the ownership of the storage medium and, under certain conditions, the information content of the data is protected, e.g., by copyright. Overall, this protection of data in civil law can be described as some kind of "patchwork rug," which is made up of many different parts that together form—from today's perspective—a sufficiently closed protection system. With regard to the trade of data or the **foreclosure** in storage media or databases, there are no gaps in the applicable law that are required to be closed by the legislature.

The creation of a right of ownership to data would also be associated with difficulties. One example is the delimitation of data to which an absolute right should exist. Particularly under the aspects of legal clarity and legal certainty, it is likely to be problematic to want to record every digital item of data regardless of the data content or whether a "pertinence threshold" is exceeded. It would also be questionable under which criteria the data should be assigned. It is conceivable, for example, to group data according to personal impact, creative process (**scripturact**), or the way people see things. Finally, it should be clarified how the right to data should relate to other rights. This pertains, above all, to possible collisions with the right to the storage medium as well as (other) rights to data content.

Another topic that the **working group** is currently investigating, and on which it will present a result in spring, concerns the question of whether and, if so, to whom machine-generated data are to be assigned or who has or should have access to it. Such data are not included in the copyright protection of data collections, and they cannot be protected as trade secrets. Therefore, the suggestion of ancillary copyright on such data seems worth checking. The fact that machine data are of immense economic importance is shown in the dispute between Lufthansa, Airbus, and

Boeing over the access to and the right to use data collected from aircraft.[1] The approximately 1.5 terabytes of data that are generated every day is, for both the airline and the aircraft manufacturer, equally valuable. The sovereignty over the data gives access to the knowledge gained and to be gained. The dispute between Lufthansa, Airbus, and Boeing is **paradigmatic** of the question of whether the European and German legal systems provide the rules necessary for the functioning of the data economy.

3.1.3 "Big Data"

Given the enormous extent which data collection and processing have gained, as well as the constant refinement and expansion of the possibilities for evaluating data volumes, the working group has also turned to the topic of "Big Data." It has therefore presented a report at the Autumn Conference of the Justice Ministers and the Minister of Justice in November 2018 (also available at https://www.justiz.nrw).

There is hardly a section of social life in which application areas for big data are inconceivable. There are some areas that appear positive at first sight. It is anticipated that there will be potential in the areas of early detection and treatment of diseases, development of new therapies using personalized medicine and precision medicine, prevention and management of major natural disasters, shortening of **design and production cycles** in industries, and acceleration of the conception of new materials. Application in areas of national security and defense are also seen, for example, in the development of complex **encryption techniques** as well as the tracing of cyberattacks and corresponding countermeasures.

At the same time, however, areas of application are emerging that harbor risks or are simply incompatible with our legal system. In almost every US state, software is used to calculate the likelihood of recidivism in courts of law.[2] In Great Britain and the United States, job applicants are sorted out in online selection processes without human involvement. In doing so, knowledge gained from credit scoring is also processed.

The question to be answered is whether and, if so, in which areas legal adjustments are necessary. The working group—based on the task it received from the Justice Ministers Conference—dealt with the application areas of big data and algorithms that play a role in our everyday life: personalized advertising, personalized hit lists and continuously updated information offers (so-called **newsfeed**), as well as personalized prices. By working out the opportunities and risks in these areas

[1] Die Welt, published on July 21st, 2018, available at https://www.welt.de/wirtschaft/article179728238/Luftfahrt-Der-erbitterte-Streit-um-das-Gedaechtnis-von-Flugzeugen.html.

[2] Stern, published January 30th, 2018, available at https://www.stern.de/panorama/stern-crime/compas-bei-gericht%2D%2Dwo-ein-algorithmus-bestimmt%2D%2Dob-jemand-rueckfaellig-wird-7843206.html.

mentioned, it came to the conclusion that there is a need for legislative action as follows.

On the one hand, offers on the Internet must be transparent if the pricing has been personalized for the individual consumer through the use of algorithms—some kind of "transparent price tag" is needed. On the other hand, it must be ensured that in the course of the creation of personalized hit lists on the Internet, especially in the case of "newsfeeds," the essential criteria for the sorting algorithm are disclosed. Only in this way will consumers be given the opportunity to make informed and self-determined decisions.

3.2 Only Occasional Need for Legislative Action in the Law of Obligations

The working group also dealt with numerous issues related to digital phenomena, characterized either by the autonomy of machines or products or by a "digital" subject matter of the contract. This was done primarily from the perspective of whether an independent legal regime for "**digital contracts**" is required or whether it is sufficient to supplement the current law of obligations, which has proven itself in the "analog" world, in individual aspects. The audit has shown that there is only a selective need for legislative action in the area of the **law of obligations**.

This applies, for example, to what is known as "paying with data," which the working group tested using social networks. Social networks are particularly suitable for generating income through personalized advertising. Accordingly, they are designed—which the majority of users are also well aware of—to "pay with data." The working group, therefore, dealt with the questions of whether personal data are to be recognized as contractual consideration, what civil law relevance the right to revoke **consent under data protection law** has, and whether a "button solution" is also required for "paying with data," such as is already in place when paying a sum of money for orders on the Internet. It is definitely questionable that the consent to be given by the user for the collection and use of his data is agreed in the form of "Terms of Use" or "Data Guidelines." However, this does not show the user in the same way as with the obligation to pay a sum of money what he is actually "paying" with. The working group, therefore, sees it as appropriate and in line with interests to legally establish a "button" solution for "paying with data" as well.

The working group also dealt with legal issues relating to the conclusion of contracts involving communicating "intelligent" objects, such as the refrigerator that automatically reorders milk. Usually, these do not pose any problems. The legal business doctrine of the German Civil Code leads to solutions that are appropriate to several interests, be it with regard to the submission and acceptance of declarations of intent or to possible contestations due to error.

3.3 Substantial Questions in Liability Law

Substantial questions, however, may arise in liability law, in particular with regard to noncontractual liability. In this respect, an intensive examination was required to determine whether a liability gap emerged in the operation of autonomous systems. The starting point for the deliberations was the assumption that such systems could lead to actions that could cause damage, which the manufacturer could not foresee despite all due care and which therefore cannot be blamed on the manufacturer according to the applicable principles of damage law. On the contrary, the operator of the autonomous system is only at fault if he has violated **duties of care** during its use. If there is no strict liability in such cases—such as based on the example of owner liability under road traffic law and animal owner liability—the question arises whether there is a risk of a liability gap that the legislature should close. It would also be possible, for example, to tighten product liability or to introduce a new, independent statutory liability regime.

Given the complexity of these issues, the working group pursued this topic under the heading "**Robotic Law**" in the past year. In this process, the team in particular examined the areas of autonomous driving as well as the use of autonomous systems in medical technology to determine whether our liability law in a worst-case scenario provides sufficient solutions or requires changes or additions. In view of the possible damage scenarios, the working group deliberately turned to these two product areas by way of example and initially left out the other areas in which the use of autonomous systems is conceivable or foreseeable. Experts from the automotive industry and the **medical technology** sector were interviewed in order to obtain firsthand information about which types of application of autonomous systems in the mentioned fields already exist today and which are to be expected in the next 10–15 years.

These surveys have shown that in relation to the examined areas on the European market, the use of autonomous systems can be expected in the foreseeable future, but not the introduction of self-learning systems. Autonomous systems are used when the system is able to carry out tasks without human control or supervision. They can manifest themselves in the form of **high-tech robotic systems** or intelligent software. Self-learning systems, on the other hand, are able to use examples in which they recognize patterns and regularities that they can generalize after the end of the learning phase and apply to unknown facts. They can therefore "teach" themselves new strategies and independently search for new, analyzable information. In this sense, their actions are no longer comprehensible and can no longer be checked by humans. On the one hand, it is not possible to determine how the machines achieve their results beyond the first algorithms. Moreover, their performance is based on data used in the learning process that may no longer be available or accessible.

Based on the rationale conveyed by the experts that no self-learning systems will be ready for the market in the foreseeable future but only so-called encapsulated autonomous systems will be used, the working group is currently examining the possible effects on our legal system to attend the **Spring Conference of the Justice Ministers** to be able to present a report on this in 2019.

In addition to relevant questions regarding the law of product approval, the working group examines which evaluation principles are based on the applicable fault and strict liability and to what extent they can be suitable instruments for dealing with the liability issues that arise when autonomous systems are being used. With respect to **strict liability**, it is necessary, for example, to answer the question whether the mere fact that an autonomous system is manufactured and placed on the market poses a special risk that could lead to strict liability on the part of the manufacturer. As a particularly relevant subcase of strict liability, **product liability** according to the Product Liability Act must also be presented separately. Among other things, the question to be investigated is whether the **difficulties in providing evidence** that may arise in connection with **error localization** in an autonomous system requires changes to the applicable legal situation. In addition, it must be checked carefully what obligations the manufacturer has with regard to **software updates**.

Since the work group's examinations have not yet been completed, no result can be reported at this point in time. Like the previous reports, this report is to be published after the Spring Conference of the Justice Ministers.

Finally, it should be emphasized how important is a reliable legal framework also for a digital society. Many actors, both citizens and companies, are unsettled by the legal classification of digital phenomena and the multitude of questions. They expect nothing less than legally secure, understandable, and at the same time balanced answers to their questions. We are working on this as legislators.

3.4 Legal Tech as an Opportunity and a Challenge for the Judiciary

3.4.1 What Is Legal Tech?

Legal technology, in short **legal tech**, in the broadest sense refers to information technology (IT) that is used in the legal field. There is no fixed, generally valid definition that is able to explain the meaning of the term more concretely. Rather, the term includes any software that can be used in the legal field. A widespread attempt at providing a meaning for legal tech resulted in the following definition: the use of modern computer-assisted, digital technologies to automate, simplify, and improve the discovery, application, access, and administration of law through innovation. In short, it is software technology that supports or replaces the work of lawyers.

The year 2016 is probably the most important year for legal tech in Germany. In 2016, the **Federal Bar Association** founded the working group "Digital Legal Advice," the Federal Association of Corporate Lawyers (BUJ) established the specialist group "Legal Tech," and the European Legal Tech Association founded its own legal tech association in Berlin. Above all, the German Lawyers Association decided to dedicate the 68th German Lawyers' Day (in 2017 in Essen) to the topic of "Innovation and Legal Tech." The accompanying media reports also made the topic accessible to a wider public. The term legal tech continued as a dominant topic in conferences and publications in 2017 and 2018. Therefore, many speak of the "hype" surrounding **legal tech**. The great attention paid to legal tech is currently disproportionate to the importance that it has in terms of the practical use of applications. However, if one looks at the potential for change that the technology brings with it for the future of legal activity, the high level of attention is justified. This also applies to the judiciary because the digital transformation of society is comprehensive, understood as a process of change based on digital technologies. Digitization allows completely new ways of cooperation, which has a significant impact on the administration of justice and its organization and external presentation. Technology is developing in leaps and bounds, but it is consistently misjudged: we overestimate the capabilities of software, and we underestimate its future effects. In addition, there is a clear crisis in access to justice and **consumer protection**.

3.4.2 General Theses

The discussion about legal tech as a threat or opportunity for consumers, the judiciary, and the rule of law is generally dominated by four theses:

1. What can be done by software will be done by software.
2. Digitization of society is also forcing the administration of justice to digitize—the rules of procedure are from the nineteenth century; federalism makes modernization difficult, but effective systems within the administration of justice combined with a reform of the rules of procedure are urgently needed.
3. Blockchain technology enables completely new transaction techniques and therefore has effects on the so-called intermediaries (banks, registrars, and notaries).
4. Technology makes nonstandard behavior more difficult.

The digitization of the economy is already fundamentally changing customer relationships. They intensify and thus go far beyond the specific delivery of goods or services. The mere purchase or service contract turns into an overall phenomenon. Almost all **e-commerce** companies are working on getting to know their customers even better, offering them security in processing, certainly also with the desire to receive good online reviews, which are extremely important in competition. This also includes the difficult and legally unresolved issue of "profiling" or "scoring." Anyone who was previously known to the owner of a business as a so-called regular customer is now only recognized through data, with all the advantages and disadvantages for pricing, delivery times, guarantees, and goodwill.

One might think that the increasing shift of trade into the digital sphere could not affect private law. However, this is a misconception. The more the contractual practice in the world of e-commerce evolves away from the standardized contractual content of the civil code, the more it is **up to the parties to determine which rules**

apply in their legal relationships. This ultimately leads to the question: what does digitization do with private law?

3.4.3 Privatization of Legal Protection?

There are two striking consequences if customer relationships would be massively and inexorably digitized.

Firstly, digital commerce "invents" its own rules, legal relationships are differentiated, and **e-commerce** then consists of a bundle of contracts. This phenomenon can already be found in the privatized trading infrastructure of e-commerce. In recent years, simple private legal rules have emerged, largely ignored by jurisprudence, which are followed by the balancing of party interests in the event of a conflict. Often these rules are predefined by a trading platform and are used in thousands of conflict cases every day. A prime example is the leading internal case decision of the payment service provider "PayPal": if there is a conflict between the merchant and the customer, the PayPal employees decide essentially according to the specification that money and goods must not be with the same person. So if a customer complains about the goods, he will get his money back if the shipment tracking does not record a delivery or as soon as he has returned the item. Whether he was entitled to a right of withdrawal, whether there was a material defect when the risk passed, and whether the seller is entitled to an objection are in principle completely irrelevant. State law, which recognizes these differentiations, is not formally waived but in fact no longer plays a role (cf. Fries 2016, p. 2860; Wernicke, lecture of January 23, 2018, given at the Bucerius Law School).

Some authors already speak of a "privatization of private law" in view of the factual rebalancing beyond the statutory standard models (Fries 2016, p. 2860; Wernicke, op. cit.).

Secondly, this finding also applies to law enforcement. The terms "**Ebay Law**" or "**Facebook Law**" are then used. A business practice that is primarily geared toward customer loyalty and therefore requires an extremely accommodating handling of customer complaints has developed throughout online trading, regardless of the legal requirements.

In this context, it is quite remarkable that the numbers recorded by the state have been declining for years (all numbers from Wagner 2017, pp. 243 ff.). The total number of civil court proceedings settled fell by almost 25% between 2005 and 2015. In view of the small number of cases, mediation or arbitration cannot be responsible for this. A closer look reveals that the decline is particularly noticeable in the case of purchased items: these fell from 164,000 to 140,000 submissions to local courts between 2005 and 2015; this is with certain probability a consequence of e-commerce and the changed forms of dispute resolution within customer relationships themselves. Traffic accident cases, for example, have not decreased, and medical liability cases have almost doubled. Of course, state civil law also applies in these legal relationships. A legal entity is still free to invoke state statutory law and to take legal action in ordinary courts. However, if PayPal law, for example, hits the legal perception of market participants so well that the assertion of legal claims remains the exception, the question of the importance of the rule of law may be asked in the future. Since the constitutionally guaranteed institution of justice also solves an individual case, state justice also contributes—with greater effort and resources, as well as procedural precautions—to an objectively correct and, in this sense, justified decision and makes contributions for the further development of the law and legal security.

Legal tech fundamentally changes these options for law enforcement: the aim is to reduce transaction costs at all levels and avoid conflicts and, if necessary, extrajudicial enforcement. As a result, state law loses its monopoly. Conversely, legal tech companies also facilitate the de facto enforcement of consumer law and thus make a contribution to easier access to law: companies such as Flightright (passenger rights), Geblitzt.de (fines), MyRight (diesel scandal), Mietright (rent brake), and CaseCheck (ALG-2 notifications) have created access to law in just a few years, which consumer protection associations were unable to provide before. In this way, consumers have the opportunity to enforce legal claims without taking an economic risk (no win, no fee) because those that conclude contracts on the Internet will be grateful for internet-based law enforcement mechanisms.

4 Judicial Enforcement (Where to Go to Online Courts?)

4.1 Formal Level

The judiciary is initially proving to be reluctant to adapt and shift to new technologies. The legislature has now opened the rules of procedure for electronic legal transactions, which are being implemented step by step. Digitization offers considerable potential for a modern judiciary. Electronic legal communication and electronic files initially (only) facilitate formal work processing but do not yet lead to a new work structure in terms of content. If one rightly grasps jurisprudence as a scarce and common good resource, it is difficult to understand why judges nowadays spend a considerable part of their working time on arranging the arguments of the parties and their lawyers and assigning them to one another so that the actual task, to weigh their arguments, is made possible in the first place. For this reason, discussions to introduce a mandatory document structure have been ongoing for years (since Calliess 2014, pp. 99 f.; approving Gaier 2013, pp. 2871 and 2874; Fries 2016, pp. 2860 and 2864). With the tools of legal tech, a digital document management system could be organized, which uses “Relationstechnik” to compare related statements and clearly separate the disputed from the undisputed. A piece of court software could then use the content-related and machine-readable form of prestructured writs, based on an analysis of the literature and case law, to carry out a preliminary examination of the content and present the legal situation in a draft judgment, adapted to a specific situation. However, this (well-known) proposal for the structure of the writ is currently not enforceable in the legal profession.

It would also be conceivable to digitize the judiciary beyond the formal level in terms of content. One could allow judges to access cases with identical or similar party submissions within the judiciary. They could then look through these and, if possible, make use of the reasons for the decision made there. This would not only make work easier for judges but also possibly contribute to the uniformity of court rulings. It would even be conceivable to let the software automatically generate legal information for the parties or even a preliminary judgment based on the analysis of similar cases, thus only entrusting the judge with the review of this proposal. This would free the judiciary from unnecessary preparatory work.

However, there is the agreement that digitization cannot and should not replace judges. The state in fact has to equip the judiciary with legal tech instruments, i.e., support software, so that it does not fall behind in the face of the increasingly complex procedures vis-à-vis the legal profession. In this respect, the judiciary will benefit from the possibilities of **legal tech** because support software can make the work of the courts more efficient. Nevertheless, legal tech instruments are always of a supportive nature. Judges must therefore always be able to see how results are achieved. The legal limits of legal tech are clearly where discretionary and weighing decisions are at stake because people trust people, not algorithms. There can be no jurisdiction without a human factor.

This is particularly evidenced by proceedings in family matters. However, family lawyers are already recognizing the advantages of legal tech. This is because family lawyers often use software to calculate maintenance, profit compensation, and pension compensation, and they do this even if they are not able to understand the individual calculation method. This reveals a substantial weakness of legal tech: as soon as judges read their judgment from a “subsumption automation” only, whose calculation methods they do not understand, the judge’s function shifts to the programmer. That cannot and must not be the case in the future.

4.2 The Online Court

The development from judicial online arbitration to real online courts, in which a court will be a service and not a place, is progressing. With Great Britain as a pioneer, and also the USA and the Netherlands (and China, in Hangzhou), the debate about online courts is picking up speed and has reached the so-called mainstream. In Great Britain, since the report by Lord Justice Briggs (Briggs Report) of July 2016, concrete efforts have been made to fully digitize the judiciary and to introduce an online court. This online court is said to have jurisdiction initially over cases involving claims of up to £25,000 but eventually over all other disputes, with the exception of highly complex and particularly important cases. This new court should be used by the parties themselves, with little or no support from lawyers, and should be designed with their own, user-friendly procedural rules (on this, Braegelmann 2018, pp. 215 et seq.).

However, concerns about these so-called online courts are expressed in broad forms. An oral hearing can ensure legal peace because the parties get closer. Not doing this would remove a core element of legal protection. Moreover, public scrutiny of the judicature would hardly be possible without an oral hearing. It is furthermore questionable whether evidence gathering, in particular witness interrogations, that are only conveyed electronically from a distance can have the same quality as evidence gathering during oral hearings. And to the extent that efforts are made to introduce **online courts** without the participation of lawyers, this can no longer be understood as real case law. In fact, this should require strict adherence to the rules of procedural law, in particular the principle of procedural equality of arms through lawyer representation.

It is therefore rightly said that online courts represent a so-called disruptive technology and will permanently change the structure of the legal process. Such a disruptive effect is fundamentally unacceptable because courts must remain recognizable in their activity and in their work. This presupposes places of jurisdiction in material form.

On the other hand, the judiciary should remain open to alternative dispute resolution procedures. On a voluntary basis and with the consent of the parties, there should be no objection to the operation of virtual courtrooms in the future. Especially if the personal appearance of the parties is not ordered and only applications are required and therefore the factual and legal situation is to be discussed with those who have legal knowledge, a video negotiation is a sensible middle ground between a written procedure and an oral hearing, for which party representatives would have to travel from distant places just to attend. In this respect, **video negotiations** can also contribute to a significant reduction in transaction costs.

5 Future Form of Justice

Civil litigation in the future will largely be prepared and conducted electronically. The cognitive process will generally be presented in electronic form. Paper-based processes will not disappear completely but will increasingly fade from the spotlight. Then there is also the question how the future working environment of the judiciary should look like. Specifically, this means whether and how court buildings are to be designed and organized in the future. If our parties involved in the process and a large number of our employees will work predominantly or at least to a considerable extent digitally, then the classic room concepts must also be rethought. Mobile working could also be implemented in open office environments based on the principle of desk sharing. Meeting rooms and consultation rooms remain indispensable, but the classic spatial structure of courthouses needs to be reconsidered.

6 Conclusion

1. The digitization of the law will result in a partial privatization of the law as well as the enforcement of the law in the sense of an outsourcing or forwarding of certain disputes in customer relationships and thus in business decisions.
2. The lack of publicity in online dispute resolution can weaken the validity of the law if it creates extensive legal reality alongside state jurisdiction.
3. Law enforcement, i.e., the use of coercion to implement the law, is left to the state and its judiciary alone. The state judiciary remains the constitutionally guaranteed institution of justice, so that online courts as autonomous forms of jurisdiction are not to be accepted. However, consideration can be given to providing **online courts** as alternative formats for special cases.
4. In front of state courts, a computer-aided structuring of processes and jurisprudence should be possible (IT-supported “Relationstechnik”).

References

Braegelmann T (2018) Online-Streitbeilegung. In: Hartung M, Bues M-M, Halbleib G (eds) Legal Tech, pp 215 et seq.
Calliess G-P (2014) Verhandlungen des 70. Deutschen Juristentages Hannover 2014 Bd. I: Gutachten Teil A: Der Richter im Zivilprozess – Sind ZPO und GVG noch zeitgemäß? Beck
Fries M (2016) PayPal Law und Legal Tech – was macht die Digitalisierung mit dem Privatrecht? NJW, 2860 et seq.
Gaier R (2013) Der moderne liberale Zivilprozess. NJW, 2871 et seq.
Wagner G (2017) Rechtsstandort Deutschland im Wettbewerb. Beck, München

Who Owns the Data?

Walter Frenz

Contents

W. Frenz (✉)
Unit for Mining, Environmental and European Law, RWTH Aachen University, Aachen, Germany
e-mail: frenz@bur.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_2

1 Digitisation and Law

Industry 4.0 includes all industries. The practical corporate application, which poses most legal questions, is about a procedure-oriented usage of resources to improve utilisation, to control stocks and to counteract if processes differ from current strategies. Business processes are supposed to be improved regarding connected logistics, management, system connections, buildings, facilities and other supplies as well as the connection of producers and users or rather customers. Additionally, it is intended to achieve process innovations. Moreover, the planning cycle has to be diminished and it must be adjusted to the market.

There is a constant need to process data; thus, it is necessary to secure and protect them. Therefore, **data security 4.0** has a significant part to play, specifically for ensuring data safety. If attacks on data by third parties cannot be prevented, questions of culpability and prosecution of infringements (cf. contributions by Momsen and Hartmann) should be brought to the forefront. At the turn of the year 2019, there was a sensational attack on the data of certain politicians and celebrities in Germany, which was leaked online shortly afterwards. This attack led to the arrest of the perpetrator and reignited the discussion about information technology (IT) security. Therefore, the Federal Ministry of the Interior announced an immediate reform of the German IT Security Act (Handelsblatt, 04 April 2019; draft bill from 27 March 2019). However, **cyberattacks** on accident control systems are especially challenging (cf. the paragraph by Müggenborg).

If several companies are involved or if one company depends on another, e.g. to use a cloud or receive a software, **numerous legal difficulties** can arise: Which rules apply to the realisation and handling of contracts, if data is involved? How to chide deficits? What are the consequences of one's legal liability?

Furthermore, there can be **severe problems regarding the application of competition law**. To what extent and for how long would cooperation be possible? Are there access claims, and how much do they cost? Does digitalisation need rules, and which parameters are essential for their development (extensive discussions can be found in the articles about competition law and standard setting by organisations)?

However, in the current market environment, ownership of data is crucial: **Who owns the data?** This question appears in supply chains as well as in corporate cooperation. The difficulty is shown in overview in the following brief summary. A more immersive view follows in the paragraphs regarding property law 4.0, intellectual property law 4.0 and copyright 4.0. Partly, a legal property assignment of data is declined because it is not definite and flexible enough. This certainly happens at the expense of deficient possibilities to counteract as well as a lack of sanctioning if a third party is unable to pay damages after an unauthorised usage (ct. Spindler 2018, pp. 151–170). If an entity is assigned to property rights regarding data, violations result in injunctive reliefs and claims for damages. This can also be objectionable under antitrust law but not consistently and only under special conditions (see the article standard setting by organisations).

2 Data Security and Protection

2.1 Source of EU Data Safety and Consequences

Digitisation evidently involves lots of data. Nevertheless, data security is not a topic of priority at the EU level but data protection. Actually, it is less about protection and secure shielding from unauthorised third party's use but rather more about protection of personal data. This particularly became evident in the *Schrem* decision (European Court of Justice (ECJ), Case C-362/14, ECLI:EU:C:2015:650, recital 38 et seq.—Schrem), which had an exceedingly great impact on the international scene. This decision concerned the onward transmission of data to the USA. Therefore, it is also conceivable that providers of platforms for Industry 4.0 track, save and redirect data to the USA, which is emerging through third party use. Based on the Safe-Harbor convention, the ECJ certainly has declined any redirection of personal data, which has not been checked in each case by case review.

The ECJ justified its decision with references to the basic rights provided under Articles 7 and 8 of the Charter of Fundamental Rights of the European Union (CFR), which concern personal integrity and the protection of personal data, respectively. "Limitations in relation to the protection of personal data must apply only in so far as it is strictly necessary" (ECJ, Cases C-203/15 and C-698/15, ECLI:EU:C:2016:970—Tele 2 Sverige; already Case C-293/12, ECLI:EU:C:2014:238, recitals 52 et seq.—Digital Rights Ireland). **However, corporate data** is (for the intervention of Article 8 EGRC in regard to corporate data, see Guckelberger 2011, p. 126; Heißl 2017, pp. 561 et seq. with further references; in contrast Frenz 2009, recital 1369) secured by fundamental economic rights, thus by corporate freedom and **intellectual property** (weakened in pursuance of ECJ, Case C-92 and 93/09, ECLI:EU:2010:662—Schecke und Eifert, except for the case that the name of the artificial person determines one or several natural persons). This has to be saved as well. Therefore, excelling common goods have to be named in order to justify the saving and onward transmission of corporate data.

2.2 EU General Data Protection Regulation: Regulation of Data Privacy

The data protection law was extensively codified in the General Data Protection Regulation (GDPR).[1] Its primary focus is the protection of personal data against unauthorised access and processing. Personal data, in respect to the *ratio legis*, may

[1] Regulation (EU) 2016/679 of the European Parliament and the council from 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), ABl. 2016, L 119 p. 1, in force since 25.5.2018.

be processed as little as possible (principle of data minimisation, Article 5 paragraph 1 lit. c GDPR) and may, in principle, not be maintained longer than it is necessary for processing purposes (**principle of storage limitation**, Article 5 paragraph 1 lit. e) GDPR). Technical and organisational measures to secure these principles must also be taken (Article 25 GDPR, more in Paal and Hennemann 2017, pp. 1697 and 1700). Thus, the integrity of personal data is ensured, but there is no protection against attacks by third parties.

Because of that, it is not the **external data security**, which is ensured, it is the data security itself regarding the integrity of personal data. At least the GDPR is in so far interconnected through Article 35 par. 7 lit. d) with the information security as its part of the data protection impact assessment (see the article by Wischmeyer). However, this also regards the indemnity of the protection of personal data. In order to do that, it is not possible to conclude property rights of data from the GDPR. This and the question of who is behind the data in terms of personal rights must be clearly separated (Riehm 2018, p. 87).

2.3 NIS Directive

Unrestricted saving and onward transmissions of data by providers of platforms for Industry 4.0 would be a recipe for industrial espionage. Such spying must be prevented in order to protect corporate property. This applies independently of the redirection of data. Infringement by third parties—especially cyberattacks—is enough (shortly in 3). Therefore, from a corporate point of view, it is less about data protection but more about data security.

On the EU level, the Commission has specified **unified minimum standards** in the Network and Information Security (NIS) Directive[2] to ensure a high common security level of network and information systems. The member states have to take care of network and information systems in convergence with this directive in their territories. Businesses also have to take suitable and appropriate actions to achieve this objective. This regards especially operators and providers (Articles 14 and 16 NIS-Directive), flanked by state units, which should be provided with sufficient capacities.

Therefore, the NIS Directive especially mandates corporations to adopt measures to ensure data security. The provisions of the NIS Directive, which came into force in August 2016, had to be integrated by the member states into their national legal framework until 9 March 2018. In Germany, the data security obligations of

[2]RL (EU) 2016/1148 of the European Parliament and oft the council from 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union, ABl. 2016 L 194, p. 1.

companies are regulated by the IT Security Legislation.[3] Since the introduction of the GDPR and the NIS Directive, companies have had to follow EU law guidelines regarding their dealings with data. The NIS Directive is complementary to the General Data Protection Regulation. Certainly, its focus is more about "business continuity" and "resilience of critical services" (see Piggin 2018, p. 44).

In particular, the NIS Directive is going to prepare the European Union against cyberattacks because the reliability and safety of network and information systems, as well as the services that are made possible through them, are exceptionally important for society. Therefore, these services must be protected. These include energy, water, medical care, digital service, and logistics. Piggin concludes: "Boards will need to review their governance processes, risk management and cyber security plans to ensure their security measures meet acceptable levels of cyber security deemed appropriate by the relevant competent authority. Evidence of cyber security capability and practices to protect essential services will be necessary to avoid potential non-compliance" (Piggin 2018, p. 44). Therefore, companies have to install mechanisms to prevent cyberattacks.

2.4 Cyberattacks and the Necessity of Strengthened Regulation and Prosecution

2.4.1 Intensified State Activities After the Last Cyberattack

As shown in the latest digital assault on ~1000 personalities in politics and the media at the beginning of 2019, when personal data such as phone numbers were leaked online, cyberattacks are still the main criminal issue in the realm of the Internet and digitalisation. Therefore, the Federal Minister of the Interior, Seehofer, demanded a **"cyber defence centre plus"** after consultation with the president of the Federal Office for Information Security (BSIG-E), Schönbohm, and the president of the Federal Criminal Police Office Münch. This centre is going to connect the different competences and responsibilities on federal as well as on state level to accomplish the defence against cyberattacks. The Federal Minister of Justice, Barley, furthermore ordered platform providers to instantly suspend accounts distributing the hacked data (Jahn, NJW-Newsletter from 9.1.2019).

While this second measure addresses private actors, the first proposal from the Federal Ministry of the Interior affects solely the state. This proposal moreover was made concrete through the draft bill on the IT Security Act 2.0 from 27 March 2019. It includes other criminal offences and puts no limit to investigatory powers.

[3] IT-Security Act (IT-Sicherheitsgesetz) from 17 July 2015 (BGBl. I p. 1324), amended by art. 5 paragraph 8 of the act from 18 July 2016 (BGBl. I p. 1666) see Roßnagel (2015), pp. 1206 et seq. Now draft bill from 27.03.2019 for an IT-Security Act 2.0.

The development of new technologies always contains new starting points for criminals, too. So far, conventional cyberattacks only affect the confidentiality, integrity and accessibility of data and services (Loukas 2015, p. 1). Due to the increase in cybercrimes, they constitute a major threat to society and the economy (Holzleitner and Reichl 2017, p. 14). A fortiori, the state has to react. The state protection obligations are providing the foundation. A manifest evidence for these hazards became evident because a single student was able to conduct the comprehensive hack above-mentioned. Therefore, the state has to apply countermeasures to protect the victims. Because of these kinds of crimes, measures also have to take into account the extended reach of criminals with the use of improved equipment.

The public expects appropriate reactions. Only 16% of those polled by ZDF Politbarometer on 1 January 2019 think that private data are (very) safe on the Internet. 79% see safety issues, and 75% hold the opinion that the state does not do enough to fight cybercrime. On the other hand, only 30% recognise their own fault. Moreover, 54% consider their cyber-security efforts to be sufficient (Europaticker from 11 January 2019 concerning the ZDF Politbarometer 1 January 2019).

2.4.2 Alignment with New Judicature

State protection obligations are not activated by virtue of public demand only; the German Constitution does not contain a referendum. However, the state can react to the sentiments of the population. Therefore, the Federal Constitutional Court (BVerfG) legitimised the (re-)preponement of the nuclear power phase-out based on state protection obligations regarding life and health, even taking into account a changing public risk awareness (BVerfG, verdict from 6.12.2016—1 BvR 2821/11 et al., recital 308). Nuclear power is an object lesson on how politics can react to a change of sentiment in the public.

The social relatedness generally constitutes a high design flexibility of the legislator (BVerfG, verdict from 6.12.2016—1 BvR 2821/11 et al., recital 268). If it is a matter of dangerous phenomena, the legislator will have to carefully observe the development based on his constitutional protection obligations and can therefore also decree appropriate tightening.[4] This especially applies to current incidents. Therefore, even a single incident is sufficient, which does not bring along any further local threats. Thus it is irrelevant that the cyberattack above-mentioned was carried out by a student and that the leaked data was not significant. The Federal Constitutional Court explicitly conceded the legislator to expedite the nuclear power phase-out to protect the health of the population and the environment, even

[4] Already BVerfG, decree from 08.08.1978—2 BvL 8/77—BVerfGE 49, 89 (130 et seq.)—Kalkar with relation to human dignity. See BVerfG, verdict from 6.12.2016—1 BvR 2821/11 et al. recital. 283, 303 with reference to art. 2 paragraph 2 and art. 20a constitution.

without new threat insights but against the backdrop of the reactor accident in Fukushima.[5]

2.4.3 Constitutional Protection Obligations Against Cyberattacks

Although cyberattacks do not contain the same risks as reactor accidents, they can still interfere with public and individual interests in a significant manner. The private use of the Internet brings about intense consequences. Many areas of life are already digitised. If access to private data is granted, it is possible to evolve a picture of an individual, which enables third parties to draw critical inferences and often to reveal major parts of the regarding identity. So the general right of personality is threatened.

In economics, professional activities are often based on the constant use and the accessibility of the Internet as well as the protection of sensitive data. It is therefore necessary to ensure an undisturbed use without the hazard of being spied on by competitors. As part of the continuous progression in digitalisation, business activities increasingly become inconceivable without data. Therefore, criminals should not be able to access data through data leaks. Thus, there are threats to the economic expansion and freedom of occupation, as well as existing data and therefore freedom of ownership.

Personal as well as business data can be affected by cyberattacks. They can compromise the user-provider relationship of trust considerably and fundamentally threaten functionality of the data exchange together with business activities in the field of digitisation as a whole. Consequently there is a **state protection obligation for data protection against private assaults** based on the **general right of personality, freedom of occupation and the freedom of ownership**.

All of these starting points require effective state measures to fight cybercrime. MacDonnell Ulsch even states that we are in a cyberwar, which evokes the necessity to raise consciousness about security breaches (MacDonnell Ulsch 2014, p. XV). State countermeasures have to be intensified even more. Public consciousness for security breaches must be raised or rather created by the state through **information and warnings**, too.

Beyond that, technical possibilities must be realised to prevent cyberattacks in advance. Apart from that, there must be criminal law provisions prosecuting and punishing cybercriminals. This is a task for the national **penology authorities**. Thus, its sufficient elaboration must be connected with an **effective prosecution** in order to create a deterrent impact. Such an impact can only be achieved by providing severe punishment for the act. Therefore, **threats to impose punishment, in line with §§ 202a et seq. and 303a et seq. Criminal Code,** must be made according to the IT Security Act 2.0.

[5] BVerfG, verdict from 6.12.2016—1 BvR 2821/11 and others, DVBl. 2017, 113 (headnote 6, recital 283, 304); see also VGH Kassel, verdict from 27.02.2013—6 C 825/11.T, ZNER 2013, 419 with critics from Becker, ZNER 2013, 339 because of existing real reference points.

Any strengthening of punishment needs sufficient technical mechanisms on different levels. These hardware- as well as software-oriented mechanisms are required to identify cyberattacks swiftly. These include personnel resources to secure effective prosecution—all of this as quickly as possible to prevent further damages. If legally high-protected interests are seriously threatened, the Federal Constitutional Court refers to the legislative observation and audit required by fundamental rights. Therefore, it is to be examined if the current standards and measures are sufficient. The legislator also can re-evaluate his causes anytime and e.g. advance the nuclear power phase-out; regarding state procedures. There is only an evidence-control of suitability and necessity in respect of the protected goods (BVerfG, verdict from 6.12.2016—1 BvR 2821/11 and others, recital 285 et seq.). Nevertheless, the state always has to assure that current measures are sufficient.

2.4.4 Inadequate IT Security Act

The German IT Security Act came into force on 25 July 2015. It guarantees improved network protection against cyberattacks. However, it regulates and demands, first and foremost, corporate efforts, which are intertwined with state control mechanisms. Considering the latest precarious situations, the Federal Ministry of the Interior now plans for a reform of the IT Security Act (ministerial draft bill from 27 March 2019) in order to accomplish its protection goals. Thus, corporate self-protection shall be accompanied better by state action—also through a strict restraint and prosecution of corporate espionage, which has to be deterred. This especially is a matter of penology (see the article by Hartmann in this book). Digital trespass should be punishable.

But the state cannot only confine itself to the creation of standards or the monitoring of corporate efforts; it also has to take its own precautionary measures to prevent security breaches from occurring in the first place (paragraph 7b sections 1 and 4 BSIG-E). Only by taking this measure can the government live up to the growing threats and protect small and medium-sized businesses, which cannot defend themselves properly against cyberattacks. They often lack financial and technical resources to install and operate ongoing defences, particularly in view of the fact that cyberattacks are becoming more advanced. The purpose of constitutional protection obligations is to protect those that cannot protect themselves sufficiently against threats by third parties or nature. Therefore, security breaches ought to be avoided.

Constitutional protection obligations allow an ample scope of measures for the kind of measures to be taken. The state only has to take measures, which are most certain and completely sufficient. Therefore, there follows from this that there is no precise predestined action (BVerfG, decree from 29.10.1987—2 BvR 624/83 and others, BVerfGE 77, 214 f.).

Consequently, the protection obligations by no means constitute a precept for an optimised effectivity, but a general efficacy. If standards for private individuals are not sufficient, government agencies have to get proactive. One way to do that is to consider increasing penalties. However, there is no **absolute security against cyberattacks**. Neither private nor state measures can grant full protection.

Nevertheless, the state must not undercut the **prohibition of excess**: It is necessary to provide an adequate and effective protection, which contains precautionary sufficient measures which are based on careful fact-finding and reasonable evaluation, whereby crosscurrent legally protected interests are to be taken into account (BVerfGE 88, 254). Depending on the extent of the threat, the state has to take countermeasures that presumably are going to grant sufficient protection without excessively limiting legally protected interests.

At the same time, the legislator has a far reaching **evaluation-prerogative**. In order to fight forest damages, several (existing) regulations were sufficient. The Federal Constitutional Court did not oblige the legislature in spite of an affirmed protection obligation in line with Article 14 paragraph 1 of the Constitution to a further reduction of private and corporate pollutant emissions or a compensation scheme for affected forest owners. Both are not mandatory under the Constitution (BVerfG, NJW 1998, 3264). The state has not yet been convicted to certain measures, except for the area of the protection of the *en ventre sa mére* as legally highest protected good (BVerfGE 39, 42 f.; 88, 254 f.). It has also not been convicted with regard to ozone limits (BVerfG, NJW 1996, 651; see prior judicature to nuclear energy BVerfGE 56, 73 et seq.). The decision concerning the nuclear power phase-out was legitimised (BVerfG, DVBl. 2017, 113 with remarks from Frenz) without it being explicitly mandatory. Therefore, this is also not to be expected for decisions on climate protection, although a constitutional challenge is pending (more about this Frenz 2019).

The state has to take immediate action against cyberattacks. However, it can hardly be expected to implement substantially more precisely determined actions. The currently revealed reforms in the form of a “cyber-centre of defence plus” and a realignment of the IT Security Act should be sufficient. The threats to impose punishment, in line with paragraphs 202a et seq. and 303a et seq. Criminal Code, are to be intensified. In conclusion, it is crucial that the state increases its efforts to secure data and does not shift the responsibility to private individuals because they apparently cannot handle it independently.

2.4.5 Constitutional Protection Obligations on EU Level

The constitutional protection obligations are to be endorsed on the EU level, too. However, they are also undetermined and apply mutatis mutandis with a wide margin of appreciation (more about this Frenz 2009, recital 360 et seq., 365 et seq.). Hence, specific requirements in terms of content do not arise at the EU level. This especially applies to regulations, which are more superior than the NIS Directive. On the EU level, care has to be taken that not only private obligations are regulated but also state actions are initiated.

Penology remains to be a classic legal subject among the EU member states. However, cybercrime is a distinctive transboundary issue. Therefore, measures on an

EU level seem to be particularly promising. This applies to the **cooperation between justice and criminal prosecution authorities** as well as bringing in **Europol**. Additionally, the threats to impose punishment have to be coherent throughout the Union, for it is only through sufficient guidelines regarding serious criminal offences and punishments, in line with Article 83 of the Treaty on the Functioning of the European Union (TFEU), can **Europe-wide minimum standards** be achieved. These include computer- as well as the organised crime as rule examples, which can be extended by unanimous council decision (so-called *dynamische Blankettermächtigung*—dynamic blanket authorisation in line with Article 83 paragraph 1 subparagraph 3 AEUV).[6] Further areas needing normative guidelines are the allocation of data and data exchange between private individuals. However, there are no specific regulations yet. Meanwhile, there is a practical interest to determine EU-wide coherent contractual principles and always being able to ascertain the ownership of data as part of digitalisation; the competition law by itself is not sufficient enough (a.m. paragraph 1). This results in a state obligation to regulate also those private relationships if their lack leads to serious difficulties in business traffic and inhibits further digitalisation processes. In any case, such regulation is justifiable in the light of state protection obligations, especially involving the corporate sector. Insofar, the European Union is moving forward.

3 EU Guideline for the Technical Implementation of Data Exchange

3.1 Open-Data Approach

The EU has developed **guidance on sharing private sector data**: the provision of data and the further use thereof by and between corporations can happen via an Open-Data approach through profitably exploiting data on a data marketplace or exchanging data on a closed platform.[7] By means of the Open-Data approach, data are made available to an "in principle open range of (re-)users with as few restrictions as possible and provided either without or for an extremely low fee." The profitable exploitation of data on a marketplace can take place "on the basis of bilateral contracts against remuneration." Meanwhile, data exchange can take place through a closed platform. "The data in this case may be supplied against monetary remuneration or against added-value services, provided e.g. inside the platform. This

[6]Bundestag and Bundesrat have to agree beforehand in Germany after Art. 23 par. 1 p. 2 constitution, BVerfG, verdict from 30.6.2009, 2 BvE 2/08 and others, recital 363—Lissabon; more in Frenz (2011), recital 3024 ff.

[7]European Commission. Commission staff working document, guidance on sharing private sector data in the European data economy, accompanying document to the Communication, COM (2018) 232 final, pp. 3 et seq. also for the following.

solution allows offering added-value services and thus provides for a more comprehensive solution for more stable data partnerships and allows for more mechanisms of control on the usage made of the data; model contract terms can lower the costs of drawing up data usage agreements." This compliance with applicable competition regulations is mandatory.

Data exchange between two or more companies must be **based on a contract**. The agreement regarding data usage or licensing must be in accordance with applicable legislation as there are laws that may be opposed to possible data exchange. In addition to and alongside the legal provisions, "the **strategic interests of each party** and competition are being preserved." The following considerations may help companies in the preparation and/or negotiation of agreements on data usage:

- Description of the data (should be made as concretely and precisely as possible)
- Quality level of the data
- Whether data sharing involves a data set or a data stream
- Compliance with legal obligations
- Respect for data protection legislation
- Transparent, clear understandable way of defining who has a right to access, (re)use and distribute data and under which conditions
- Limitations on the right to access and (re)use data
- Technical means of data access and/or exchange
- Means to protect data[8]

3.2 Legal Principles

Furthermore, the following legal principles have to be kept in mind[9]—first, **transparency**: "The relevant contractual agreements should identify in a transparent and understandable manner (i) the persons or entities that will have access to the data that the product or service generates, the type of such data, and at which level of detail; and (ii) the purposes for using such data."

Second, the **principle of shared value creation** applies, which recognises that several parties have contributed to the creation of data if data are generated as a by-product of using a product or service.

Third, the **commercial interests of all participants** must also be mutually respected: the relevant contractual agreements should also therefore accommodate

[8] European Commission. Commission staff working document, guidance on sharing private sector data in the European data economy, accompanying document to the Communication, COM (2018) 232 final, pp. 3 et seq.

[9] All of them laid down by the European Commission, Commission staff working document, guidance on sharing private sector data in the European data economy, accompanying document to the Communication, COM (2018) 232 final, p. 3.

the necessary protection for the commercial interests and secrets of data holders and users.

Fourth, **undistorted competition** has to be ensured through relevant contractual agreements, which should address the necessity of preserving undistorted competition also during the exchange of commercially sensitive data.

Last but not least, it is necessary to **minimise the data dependency of providers**: companies offering a product or service that generates data as a by-product should allow data portability as much as possible. They should also consider if possible and in line with the characteristics of the market they operate on, offering the same product or service without or with only limited data transfers alongside products or services that include such data transfers."

3.3 Expansions

Beyond these legal principles of the EU guideline and in order to prevent disputes from the outset, it is crucial to strengthen cooperation. A connection between economic entities, which collectively use data or work together within the frame of a development, can be strengthened through a knowledge management system and a mutual technical support for data access. Mutual **IT governance** should be established in order to encourage and promote information technology and the strategy and goals of both companies.

If this concerns separate or clearly defined areas, e.g. if a service-oriented company only provides dependent support for another economic entity, it is necessary to determine who gets access to which area and who owns the data.

A possible liability is based upon two constellations: Is it sectoral or consolidated? In the second case, a mutual **security organisation** should be established. In the first case, a clear demarcation in terms of who has to take which measures is mandatory.

4 Data Ownership

4.1 EU Law Reference Points

The previously described EU guideline includes contractual regulations and the question of data access (it is by no means about data ownership). This question is quite important in practice; thus, there should be an **EU regulation**. The protection obligations regarding basic rights and fundamental freedoms would legitimise such legislative efforts. Otherwise, an unnecessary complexity for the transboundary data exchange in the course of the digitisation would arise. Therefore, different legal regimes would regulate the data assignment differently. Every company would try to make the point that the law of that country is applicable, which contains the most beneficial regulations to their interests.

At least the contractual principles laid down in the guideline contain the **principle of shared value creation**. Therewith it is hardly compatible, if a company claims exclusive ownership of data despite combined efforts. The mere transparency and concomitant visibility of who has access to data does not reveal anything about the assignment of data ownership. The assignment can be furthermore softened through mere terms of use for practical reasons. Therefore, many issues remain unsolved.

4.2 National Articles and Copyright Assessment

4.2.1 Essentials in Adjustment with the Federal Constitutional Court

Thereby, according to the **national law of the member states** the question arises: who owns the data? So far, there is also no statutory national regulation, so that only indications can be taken into account. At this, the recourse to property law is common even though ownership of data — in contrast to the ownership of a storage medium—cannot be acquired (§ 903 of the German Civil Code (BGB)) because it is no article in the sense of a physical object in line with § 90 BGB. Nevertheless, an absolute property-like law is acknowledged (Riehm 2018, pp. 76 f., pp. 84 et seq.)—however not continuously (abl. Hoeren/Uphues, article Big Data, in this volume). A **clear contractual regulation** is recommended to avoid insecurities (for the energy industry Frenz 2018, p. 237, also for the following). Apart from that, according to the usual property law and copyright valuations it is decisive who delivers the data and how it is changed by the adaption to the energy economy 4.0. If this kind of processing dominates, the software company or the company, which modifies data, owns it.

Therefore § 950 BGB provides a legal acquisition of ownership in favour of the one, who creates a new moving object by processing or restructuring one or several already existing articles. This is already the case if a higher processing step is obtained (OLG Stuttgart, ruling from 20.03.2001—10 W 33/00, NJW 2001, 2889). Such step is regularly achieved by the development of a comprehensive digitisation solution, which includes the inflow of raw corporate data. Another indication is the value relation in line with § 950 p. 1 a.E. BGB. If the value of the concerning processing act is significantly higher than the value of the initial data, ownership is acquired. This is also valid for the achievement of actually new information with independent value, in contrast to the mere alphabetically sorting (Riehm 2018, pp. 87 f.).

This solution corresponds to the sampling-decisions of the Federal Constitutional Court (BVerfG, verdict from 31 May 2016—1 BvR 1585/13) and the European Court of Justice (ECJ verdict from 29 July 2019—C-476/17) regarding copyright, which allowed the broad advancement of existing elements within the frame of an artwork. This can happen even free of charge—depending on the artist's financial possibilities. The legal basis of the decision was artistic freedom in accordance with Article 5 par. 3 of the Constitution or Article 11 of the European Charter of Fundamental Rights.

At this it is about the **basis of economic life**, even **social life**: Nowadays communication as well as the maintenance and further development of economic procedures and manufacturing processes are barely imaginable without data platforms. Therefore, fundamental rights protection obligations become crucial. This is true both in terms of cyberattacks as well as in the contractual context (specified in Sects. 2.4.3 and 2.4.5). Concerning commercial data processing an underlying contractual relationship and therefore a carried out service in return for payment is common. Thereby, the economic actor who processes data all the more will not suffer from an inflicted economic disadvantage.

4.2.2 Complimentary Usage of Platforms

In the private sector, the **usability of a platform for E-Mails**, WhatsApp-messages etc. is often financed by a commercialisation of the accruing data, especially for promotional purposes. In this respect, too, the question arises if state protection obligations demand the creation of a legal prohibition regarding the continued use of data for promotional purposes. A declaration of consent of the user could be void, at least if it is in the general terms and conditions or not clearly visible and given by implication.

Vice versa, an entire business model and the complimentary private usability of electronic forms of communication would be called into question. If such a usage is necessarily interconnected with the consequence of unwanted advertising and ought to be accepted, everyone has to decide for themselves. Rather, the question of sufficient transparency arises. Thus, the mentioned interrelationship should be well known by now so that everyone is aware of the consequences of the complimentary private use of such lines of communication.

Enhanced clarification can close any information gaps—e.g. through the acting of a data protection officer.

Usage certainly does not change anything about private ownership of data. It stays assigned to the person concerned. This owner will have given at most one contractual usage permit under the law of obligations. On this basis, however, the data are merged and a data set is created. The **entire data occur now in the form of an address set**. Therefore, it is a new compiled convolute and an independent unity with a significantly added value in contrast to the formerly individual data. By the possible synopsis and the sorting into groups, information with autonomous value is created. Consequently, an acquisition of rights is inferred with recourse to § 950 par. 1 p.1 BGB (Riehm 2018, p. 88). Therefore, it is the private actor who assembles the data who is to be seen as owner of the data entirety, at least if one refers back to the storage medium.

4.2.3 Single Constellations

If the collection of data takes place on the basis of a specific usage and there is no further processing of the data, the **user** and not the producer or deliverer, e.g., of an autonomous car or a digitalised machine owns the data. The latter two can, however, contractually grant themselves a right of use in accordance with general terms and conditions (Riehm 2018, p. 87).

The producer can furthermore be granted an exclusive right of use by virtue of an agreement, be it express or provided in the terms and conditions or through an end-user license agreement. In this way, he receives exclusive power of disposal and his right over the data will be generally anticipated. Moreover, the user can be granted a contractual usage permit, under the law of obligations (§ 930 BGB analogue) (Riehm 2018, p. 89).

Usually, **general terms and conditions** are legally valid. If small print is foisted, at most an inequity can be considered, but these benchmarks are not always specific. It is therefore the best to cross out general terms and conditions or to add one's own words so that these fade into the background.

If data is only copied, the submitter keeps the exclusive ownership of the data in line with § 953 BGB analogue—unless he allows the appropriation by the producer of the **copy,** so that the producer can directly acquire a right to the copy of the data in pursuance of § 956 BGB analogue (Riehm 2018, p. 88, with the same result Hoeren 2013, p. 488).

If **employees** reach a higher processing step in a business operation, the employer still is the right holder of the data, insofar as one acted in detail according to his instructions. However, this does not apply if the data in question is the result of one's own creativity, like in the case of a programmer's work. In this particular instance, the employer has to contractually agree to transfer the data to the employee (Riehm 2018, p. 87, with reference to OLG Nuremberg, decree from 23.1.2013—1 Ws 445/12, ZD 2013, 283).

If data is given to the outside world, a **derived data transfer** in line with § 929 BGB analogue takes place. This happens by actually giving someone else access to the data and giving up on one's own access options, for example by deleting, changing the access data or moving it to another server outside the own sphere of influence (Riehm 2018, p. 89). However, if it is a mere copy, nothing changes in the assignment of data.

In general, a work that is protected by copyright and therefore belongs to its creator only exists if it is shaped by the human spirit and reflects the personality of the author (Bullinger 2014, § 2 recital 21 f.). The work **according to paragraph 2 copyright act** must therefore not only be different from the previous one, but also go beyond that and not exhaust itself in the mere routine (Schulze 2018, § 2 recital 18). What is required is an individual result, something that is new because ultimately, individuality is expressed in the fact that something that has not yet existed has been created (Schulze 2018, paragraph 2 recital 16). Only then can there be a creation that is protected by copyright.

If, on the other hand, only processes are adjusted, this must be declined. Even the **installation of an engine** in a conveyor does not allow ownership of it to be passed over (OLG Cologne, ruling from 10 May 1991—19 U 265/89, NJW 1991, 2570).

In the case of mere **adjustments** without a fundamental change in the entire company process, the power of disposal over the data therefore remains with the commissioning company. From this order it can possibly be deduced that the contractor is only working internally to adapt this data and, after the successful implementation of "Energy Industry 4.0", the use of the data made available for this purpose is no longer possible even after processing.

Storing data on a cloud does not lead to a shift of ownership, as the ECJ ruled in a related case: if copies of television programmes stored in a cloud are made available, this must be permitted by the owner of the copyright or related property rights. Ultimately, such service represents the retransmission of the relevant programmes.[10] However, if the permission of the right holder is required, he remains the owner of the data.

5 Results

1. The NIS guideline provides basic rules for data security, which is always problematic in the area of Industry 4.0; in Germany, the IT Security Act applies. It is primarily a matter of corporate precautions—with the state's role as a guarantee.
2. Data protection is comprehensively regulated at the EU level. According to the General Data Protection Regulation, there is a need for a company organisation that limits the processing and storage of personal data to the absolute minimum.
3. Data allocation should be arranged according to Union law. The EU has developed a guideline regarding private-sector data sharing to ensure equal rights in contracts. The ownership of data, if recognised, has not yet been regulated by law and should therefore be clearly contractually agreed upon. Otherwise, ownership of the data can change due to dominant processing, as the sampling decisions of the BVerfG and the ECJ show.
4. The mere collection of data does not establish ownership of them—for example, of the producer in relation to the user of autonomous driving or production systems. It is different with a systematic compilation if the data is sent to the operator of a complimentary platform for electronic communication media (WhatsApp) on the basis of at least a transfer of use or, furthermore, an anticipated transfer of ownership. In this respect, however, transparency must be ensured—for example submitting a declaration of consent or including it in the terms and conditions.

[10] EuGH, Rs. C-265/16, ECLI:EU:C:2017:913—VCAST regarded to proliferation for publicity so that a rendition, which is different from the original rendition takes place. Therefore, no exceptions for private copies become effective.

5. Storing data in a cloud does not lead to a shift of ownership. This results from a Court of Justice of the European Union (CJEU) ruling of 29 November 2017.

References

Bullinger W (2014) In: Wandtke A-A, Bullinger W (eds) UrhR, 4th edn. C.H. Beck

Frenz W (2009) Handbuch Europarecht 4: Europäische Grundrechte. Springer

Frenz W (2011) Handbuch Europarecht 6: Institutionen und Politiken. Springer

Frenz W (2018) Energiewirtschaft 4.0 im Fokus von Daten- und Wettbewerbsrecht. RdE, 237 et seq.

Frenz W (2019) Kohleausstieg en marche, Konsequenzen des Kommissionsberichts vom 22.1.2019 und der Klimakonferenz von Kattowitz, RdE, issue 4

Guckelberger A (2011) Veröffentlichung der Leistungsempfänger von EU-Subventionen und unionsgrundrechtlicher Datenschutz. EuZW, 126 et seq.

Heißl G (2017) Können juristische Personen in ihrem Grundrecht auf Datenschutz verletzt sein? Persönlicher Schutzbereich von Art. 8 GRC. EuR, 561 et seq.

Hoeren T (2013) Dateneigentum, Versuch einer Anwendung von § 303a StGB im Zivilrecht. MMR, 486 et seq.

Holzleitner M-T, Reichl J (2017) European provisions for cyber security in the smart grid – an overview of the NIS-directive. Elektrotechnik & Informationstechnik, pp 14 et seq.

Loukas G (2015) Cyber-physical attacks, how they work and how to protect against them. Elsevier

MacDonnell Ulsch N (2014) Cyber threat! How to manage the growing risk of cyber attacks. Wiley

Paal B, Hennemann M (2017) Big data im Recht. NJW, 1697 et seq.

Piggin R (2018) NIS Directive and the Security of Critical Services. ITNOW, March, p 44

Riehm T (2018) In: Hornung G (eds) Rechtsfragen der Industrie 4.0. Nomos, pp 73 et seq.

Roßnagel A (2015) Das IT-Sicherheitsgesetz. DVBI, 1206 et seq.

Schulze G (2018) In: Dreier T, Schulze G (eds) UrhG, 6th edn. C.H. Beck

Spindler G (2018) In: Hornung G (ed) Rechtsfragen der Industrie 4.0. Nomos, pp 151 et seq.

Relevance of Data Security and Data Protection in Companies from the Perspective of Criminal Law

Carsten Momsen

Contents

C. Momsen (✉)
FU Berlin, Department of Comparative Criminal Law, Criminal Procedural Law and Corporate Criminal Law - Faculty of Law, Berlin, Germany
e-mail: carsten.momsen@fu-berlin.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_3

1 Introduction

Companies can be affected by a variety of classic so-called "**IT crimes**." They can become victims of hacking attacks, just as employees can use the company's resources to commit their own crimes. The same applies to the digital distribution of incriminated content, such as child pornography, racist statements, or statements containing incitement to hatred. In the capital market criminal law, the infrastructure of companies can also be used to commit crimes. The same applies to the spying out of company secrets; here companies can be on the side of the perpetrator (employees) or the victim too. However, these and other manifestations are not exclusively restricted to a corporate context and specific to that extent.

Specific to companies, however, is the situation of being exposed to a collision of not easily compatible interests and duties. If, for example, data protection requirements collide with requirements governing the cooperation with investigating authorities, companies can be threatened with considerable fines (and criminal prosecution of senior employees), as well as if too little information is released. These problems are even more prevalent in internationally active companies.

Therefore, the following analysis shall describe areas that may give rise to specific risks or obligations for companies. These risks and obligations are considerably linked to the protection of the company's own data and infrastructure as well as possible defense measures against attacks committed by third parties. In addition, companies must take precautionary measures, both against becoming the victim of avoidable attacks and against crimes being committed from within the sphere of the corporate entity itself. This addresses the area of prevention and compliance.

If companies are nevertheless involved in criminal proceedings, they are obliged to provide evidence and to cooperate if necessary. They can also become the object of online investigations. On the other hand, there are also rights to protect one's own sensitive information and interests. In any case, documentation of all relevant processes is important here.

The following topics are the focus of the article: data protection and criminal law, own obligations to data protection, e.g., under § 203 of the German Criminal Code (StGB), the criminal relevance of possible defense measures, online investigations against companies, prevention, digital compliance, digital evidence, and the standards of "forensic readiness."

2 Data Protection and Criminal Law

2.1 Violation of Trade and Business Secrets, Criminal Law Obligation to Protect Personal Data in the Company

Trade secrets and personal data are protected by means of criminal law in various respects with regard to the provision of § 203 StGB. Therefore, the company faces the legal obligation to take special care when collecting, processing, storing and especially when passing on such data. In the following, the protection of such data by certain professions and the resulting **consequences for service providers** are discussed, i.e., companies whose business purpose is the provision of data-related services (this section follows in parts Großkopf and Momsen 2018, pp. 98 ff., as well as Momsen and Savic 2017, pp. 301 ff.)

The amendment of § 203 StGB is an example of a current, significantly expanded protection concept. As a result of its revision and the associated adjustment of legal provisions regulating the exercise of occupation, professional secrecy holders and their service providers must readjust their previous operational processes, which is also required by the General Data Protection Regulation (GDPR, Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of individuals with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC, OJ L 119/1 of 4 May 2016), which came into force on 25.05.2018.

Persons and entities that are professionally entrusted with confidential information often use external service providers since they are unable to provide all services necessary for the daily work of the law firm relating to the information they are entrusted with by employees due to lack of time, technical or commercial reasons. Thus, those external service providers come into contact with the commonly highly sensitive data of clients/patients/customers and other persons. However, they were not per se subject to the protection of professional secrecy. The revised appearance of § 203 StGB in connection with the relevant professional (regulatory) legal standards is now to ensure comprehensive protection of secrets. However, the reform of last autumn imposes considerably more far-reaching obligations on the holders of professional secrets, which are presented as comprehensive compliance with the law firm. These obligations are further extended and in part concretized by the **General Data Protection Regulation** (GDPR).

Internal and, above all, external system administrators are in fact forced to inspect all data—including personal data—processed by the programs they offer to eliminate a problem. Therefore, they inevitably have a great deal of "insider knowledge." This applies not only to the processed data itself, but also to the so-called metadata. This specific type of data contains highly sensitive information such as who works when and for how long with which programs and what content. System administrators and similar employees therefore do not only gain access to confidential information within their own companies, but also when they work as independent service providers or look after the IT of a company providing services as a cloud provider (managed services). If outside system administrators work for law firms, notaries, patent attorneys, tax consultants or auditors, for example, i.e., firms whose storage media contain innumerable secrets of their clients, they are naturally required to observe data protection and client confidentiality. Regardless of the practice, which has remained unimpressed by this for a long time, it was legally questionable whether professional secrecy holders in particular are even allowed to grant external system administrators access to the foreign secrets that have been entrusted to them or have become known to them in their professional capacity. With the amendment of § 203 paras. 3 and 4 StGB, the way has now been opened for a clearer regulation of a grey area in the outsourcing of services, including companies in the private health, accident or life insurance sector as well as clearing houses (Act on the New Regulation of the Protection of Secrets in the Involvement of Third Parties in the Professional Practice of Persons Subject to Secrecy, legal materials available at https://www.bundesrat.de/SharedDocs/beratungsvorgaenge/2017/0601-0700/0608-17.html). This means that all companies that offer IT services for these professional groups (i.e., hospitals and administrative and judicial authorities in which members of these professions come into contact with secrets) are affected by the new protection concept.

§ 203 StGB ensures the protection of secrets from unauthorized disclosure, which are entrusted to (professional) secret carriers in the context of their professional activity. First, by the "Law for the new regulation of the protection of secrets in the cooperation of third parties in the exercise of the profession of persons subject to professional secrecy" (BT-Drs. 18/11936, p. 36), § 203 StGB as well as the associated procedural and professional code of conduct law that has been redesigned. With these, a protection of secrets can be ensured even in the age of—not only digital—outsourcing. The change in the law harmonizes forms of outsourcing that are already widely practiced with the requirements of a protection of secrets under criminal law. This is not only about IT outsourcing and the transfer of specific technical tasks by Legal Information Management or Legal Project Management to third parties, but also about other activities such as document destruction, typing, translation or billing work and telephone service. The reformed criminal law and the associated criminal procedural standards, in conjunction with the (GDPR), impose specific and in some cases far-reaching **compliance** requirements on professional secrets and service providers.[1]

In the future, external service providers, e.g., contract data processors will be qualified as assistants and thus newly included in the circle of obligated parties if they are involved in their professional activities in any way and make contributions to them. The reduction in the protection of secrets associated with the involvement of third parties will be compensated for by including in the criminal liability under

[1] This obligation of secrecy is (newly) flanked by § 29 para. 3 BDSG. Within the scope of application of § 203 StGB, the supervisory authorities may not demand or arrange for the release of data. If they nevertheless gain access to appropriately protected data, the scope of application of § 203 StGB is automatically extended to the supervisory authority.

§ 203 StGB persons who, in the course of the proper performance of their activities, have the opportunity to gain knowledge of protected secrets. Now, all persons involved in the exercise of the profession are liable to prosecution if they disclose a secret that has become known to them during their work. The (professional) secrecy carriers are therefore also obliged to instruct the participating persons about the now existing criminal liability (§ 203 para. 4 no. 2 StGB). However, they are not allowed to grant access to the secrets in general, but only as far as this is necessary to use the service (§ 203 para. 3 sentence 2 StGB).

The customer must ensure that an obligation of secrecy is observed when involving **external persons** in the exercise of their profession (§ 203 para. 4 no. 2 StGB). This obligation applies regardless of professional or other legal requirements. Violation of this obligation is punishable by law for all (professional) secrets if the person involved has disclosed a secret without authorization (§ 203 para. 4 no. 2 StGB). An assignment by acclamation, for example in the case of an IT security incident, is therefore not possible, in which as a rule greater damage can only be prevented by rapid action. In the context of the preparation of the continuity and recovery plans, which are also (newly) required under § 64 para. 3 no. 9 Federal Data Privacy Law (BDSG), a sample instruction must therefore be included in the **emergency manual** so that the instruction in case of IT emergencies is reminded.[2] In addition to the instruction on confidentiality, however, instruction on data protection is also required, which must be provided in writing by 24.5.2018 (§ 4a para. 1 s. 3 BDSG old). The same applies in multi-level contractual relationships in which only the contractor is obliged to instruct the subcontractors.

The examination of professional competence should be based on **certifications** and other proof of qualification. Under Art. 42 para. 5 GDPR, certificates under data protection law may only be issued by supervisory authorities or accredited certification bodies. However, there are currently no such certificates under data protection law,[3] as the EU Data Protection Directive 95/46/EC did not accept certifications as proof of compliance, which is why the "Privacy Seal of Approval" of the Independent State Centre for Data Protection of Schleswig-Holstein (see https://www.datenschutzzentrum.de/guetesiegel/register/) and European certificates such as the "European Privacy Seal (EuroPriSe - see https://www.datenschutzzentrum.de/guetesiegel/register/)" have not become established on the market (see ENISA, Annex A: Analysis of existing certifications at a glance, pp. 32 ff.). Other practical

---

[2] A complete emergency management system is described on the website of the German Federal Office for Information Security (BSI) in the BSI Standard 100-4, available at https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/ITGrundschutzstandards/BSI-Standard_1004.html.

[3] See, for example, the research project "European Cloud Service Data Protection Certification (Auditor)" available at: http://auditor-cert.de und European Union Agency for Network and Information Security (ENISA), or the Recommendations on European Data Protection Certification, Version 1.0 November 2017, available at: https://www.enisa.europa.eu/publications/recommendations-on-european-data-protection-certification/at_download/fullReport.

certification procedures have yet to be developed.[4] As in the case of reliability, the service companies themselves will have to offer and guarantee appropriate standards (§ 3 BDSG). The "pre-employment screening" and "in-employment screening" to be carried out under the Money Laundering Act (AMLA) can be used as a benchmark for the necessary checks.[5] For the assessment of reliability, the decisive factor is whether a service provider, considering all relevant circumstances of the individual case, can be expected to perform the service to be provided in a proper and contractually compliant manner and, above all, to maintain the necessary confidentiality. The service provider must have taken concrete technical, organizational, and personnel measures, thus setting up **a compliance management system** and implementing the compliance requirements. These requirements are particularly important when using remote maintenance systems.[6] Services provided **abroad** can also be used. A prerequisite, however, is a level of protection comparable to that in Germany, which, according to the explanatory memorandum to the law, should be available in all EU Member States (BT-Drs. 18/11936, p. 35 with reference to the opinion of Advocate General Juliane Kokott of 29.4.2010, Case C-550/07 P—Akzo Nobel, ECLI:EU:C:2010:229). For the rest of the world, it must be examined on a case-by-case basis whether the necessary protection is guaranteed, unless the protection of secrets does not require comparable protection (§ 43e

[4] See, for example, the proposal for an EU Regulation "on ENISA, the 'EU Cybersecurity Agency', and repealing Regulation (EU) 526/2013, and on Information and Communication Technology cybersecurity certification ('Cybersecurity Act'), COM(2017) 477 final/2 and also ENISA, Overview of the practices of ICT Certification Laboratories in Europe," Version 1.1, January 2018, available at: https://www.enisa.europa.eu/publications/overview-of-the-practices-of-ict-certification-laboratories-in-europe/at_download/fullReport. See also the "Draft Opinion" of the Committee on Civil Liberties, Justice and Home Affairs of the European Parliament (Rapporteur Jan Philipp Albrecht), 2017/0225(COD), available at http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+COMPARL+PE-615.394+02+DOC+PDF+V0//EN.

[5] S. Bundesamt für Verfassungsschutz/Bundesamt für Sicherheit in der Informationstechnik/Bundesverband Allianz für Sicherheit in der Wirtschaft e.V., Wirtschaftsgrundschutz, module MA2 Bewerberprüfung, status July 2017, available at: https://www.wirtschaftsschutz.info/DE/Aktuelles/Wirtschaftsgrundschutz/Bausteine/Bewerberpruefung.pdf.

[6] Before using remote maintenance, check whether the remote maintenance software allows encrypted transmission, which encryption is used, etc. When using cloud platforms such as Dropbox, Google Drive, Microsoft OneDrive, etc., i.e., storing secrets on foreign servers, in addition to transport encryption, user- or group-based encryption at file level must also take place in the company of the secret carrier, because then the service can be provided without knowledge of secrets. The same obviously applies to Infrastructure as a Service (IaaS), where entire computers (servers) are rented, or Platform as a Service (PaaS), where the provider only provides a runtime environment within which users can run their own software. In the case of "Software as a Service" (SaaS), the service provider offers special software that runs on the provider's resources and is made available to the user online, whereby the service provider also takes care of maintenance through updates and upgrades, such as with Microsoft Office 365 and with Google Docs, Sheets, Slides and Forms.

para. 4 BRAO;[7] § 26a para. 4 BNotO;[8] § 39c para. 4 PAO;[9] § 62a para. 4 StBerG;[10] § 50a para. 4 WiPrO[11]).

Since the practice lived up to now can no longer have a legitimizing effect after the amendment of the domestic law and the GDPRs entry into force, the duty of confidentiality and the duty to maintain professional secrecy exist for all information arising at the (professional) secret carriers (concerning the different possibilities to restrict a criminal liability by reference to the idea of social adequacy: Roxin 2006, § 10 Rn. 36 ff.; Rosenau 2018, before §§ 32 ff. Rn. 61 f.). However, this happens, for example, when sending unencrypted e-mails according to the (still) prevailing opinion in the literature (see Degen 2016, § 66 Rn. 109 ff.; Eisele and Lenckner 2014, § 203 Rn. 19; von Lewinski 2004, p. 12; Härting 2005, p. 1248; Sassenberg 2006, p. 196; a.A. Koch 2014, p. 691), although third parties can easily gain knowledge of the e-mails and e-mail providers are undoubtedly outsourcing service providers (explanatory memorandum, BT-Drs. 18/11936, p. 8; see also Dix 2014, § 1 marginal no. 170 m.w.N.). Moreover, in the future the protection of secrets will be flanked by technical and organizational data protection under Art. 5 para. 1 lit. f. and Art. 32 GDPR, and any violation will be punished with fines (Art. 83 GDPR).[12] The processing of personal data without the consent of the person concerned may also constitute a violation of competition law (LG Hamburg, Judgement from 02.03.2017—327 O 148/16, BeckRS 2017, 117352). Finally, § 203 para. 4 s. 2 StGB not only requires a careful selection of service providers and an obligation to maintain secrecy, but also distinctly demands that their activities are monitored.[13] The cooperation must be terminated if the service provider obtains knowledge of third-party secrets which are not necessary for the fulfillment of the contract (§ 43e para. 3 no. 2 BRAO; § 26a para. 3 no. 2 BNotO; § 39c para. 3 no. 2 PAO; § 62a para. 3 no. 2 StBerG; § 50a para. 3 no. 2 WiPrO). Whether the service provider obtains such knowledge can only be determined by constantly monitoring its activities. Such monitoring requires a "**management system for the protection of secrets**," i.e., first of all the establishment of procedures and rules, which are

[7] Federal Statutory Order concerning the Legal Profession.

[8] Federal Statutory Order concerning the Notaryship.

[9] Federal Statutory Order concerning the Practice of Patent Attorneys.

[10] Federal Law concerning the Practice of Tax Accountants.

[11] Federal Statutory Order concerning the Practice of Public Auditors.

[12] S. Statement of the Hamburg Commissioner for Data Protection and Freedom of Information on the business number D42/2017/1114 of 8.1. 2018, available at https://www.datenschutzbeauftragter-info.de/wp-content/uploads/2018/02/schreiben-der-aufsichtsbehoerde.pdf and 8th Activity Report of the Saxon Commissioner for Data Protection, submitted as of 31.3.2017, p. 138.

[13] Page 4 of the BMJV's draft bill for a law on the new regulation of the protection of secrets in the case of the participation of third parties in the professional practice of persons subject to professional secrecy, available at: https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/RefE_Neuregelung_Schutzes_von_Geheimnissen_bei_Mitwirkung_Dritter_an_der_Berufsausuebung_schweigepflichtiger_Personen.pdf.

intended to define and control the protection of secrets (Jahn and Palm 2011, p. 620 on the requirements for a lawyer's secretariat outside the law firm, especially regarding telephone service). The implementation of this system by the service provider must then be monitored and continuously improved. The service provider must therefore be required to provide his service in accordance with the "state of the art," i.e., in accordance with the state of development of advanced procedures, facilities, or modes of operation, which assures the practical suitability of a measure or procedure for the protection of secrets (for impact assessment, Art. 35 DSGVO). For digital archiving, the state of the art is laid down in the "Technical Guidelines" of the Federal Office for Data Security (BSI) (see BSI TR-03138 Replacing Scanning (RESISCAN), available at https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/TechnischeRichtlinien/TR03138/TR-03138.pdf). The BSI guidelines must also be observed for **remote maintenance** (see BSI, IT-Grundschutz, M 5.33 Absicherung von Fernwartung, available at https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzKataloge/Inhalt/_content/m/m05/m05033.html). For the destruction of files and data carriers, the state of the art for (professional) secret carriers is the highest protection class 3 for particularly confidential and secret data as described in DIN 66399 and at least security level 4 (particularly sensitive data reproduction with extraordinary effort). During typing, translation and invoicing work, and, of course, during remote maintenance, the service provider must protect its own IT infrastructure in accordance with the state of the art, i.e., in accordance with the BSI's Catalogue of Requirements M 1 (see BSI, IT-Grundschutz, available at https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzKataloge/Inhalt/_content/m/m01/m01.html). Under § 13 para. 7 Telemediengesetz (TMG), these requirements also apply to services offered online (including Germann and Voigt 2017, p. 93; Michaelis 2016, p. 118). There is no state-of-the art technology specified by independent bodies for telephone services (see Jahn and Palm 2011, p. 613). It is therefore questionable which standards providers must follow. In addition, contracts with service providers and clients will have to be revised or at least adapted. However, it will no longer be possible to simply pass on the risk.

2.2 Hacking Against the Company: The Criminal Relevance of Own Defensive Measures Against Attacks on Corporate IT

Compliance with appropriate security standards is essential not only for the IT service providers and service providers referred to above, but for every company. Nevertheless, attacks on the company's IT can occur, leading to the fear of losing relevant data and protected secrets. To what extent may a company defend itself against such attacks? In what form can **technical prevention** be implemented?

Due to the structure of modern IT attacks, a defense can only be successful if it starts as early as possible and prevents the outflow of data. Preventive or active **defense strategies,** however, are for the most part associated with the commission of criminal offenses under German law. Anyone who has been the victim of an attack on data stocks or hardware will try to put an end to the attack as quickly and sustainably as possible. This means that it is usually too ineffective to make use of state assistance, e.g., to file a criminal complaint and wait for the attacker to be investigated and convicted. In addition to the question of the speed and effectiveness of police measures, another problem is that attacks often have an international dimension and state investigations are therefore relegated to the tough path of legal assistance from other states. The best way to ensure a sustainable defense is through "**active defense**," i.e., active action against attackers with the aim of destroying their infrastructure and preventing further attacks. However, at the same time, this is the basis for a separate act of aggression that is subject to criminal prosecution and can only be justified as self-defense under certain conditions and within narrow limits. If the attack is instigated by government agencies, the scope for "active defense" shrinks even further.

If you look at the mass of hacking attempts on company sites, attacks can be roughly structured:

- The remote system (RS) is owned by the attackers and performs an active attack on the company's system.
- The RS is not owned by the attackers, they are active "users" on the RS.
- The RS has a weakness that the attacker takes advantage of and controls it remotely.
- The RS is used to carry out the attack, the attack is automated.

When planning a "counter strike," the question arises as to what repercussions each of these scenarios will have for the attacked company itself. Due to the structure of criminal behavioral obligations, self-defense in the sense of active defense can only be justified in exceptional cases. At the very least, the original attack must be illicit and there must be a good reason for allowing this form of self-defense by way of exception.

The here relevant justification reason of **self-defense** (§ 32 StGB) requires first an illegal and present attack on protected interests. The defensive action following the attack must be necessary and suitable. In principle, therefore, a defense may only take place against the attacker. Still, in the past there have also been cases in which exceptions to this principle were made, to the extent that actions against objects, etc. belonging to third parties were justified as well, provided that these were used in the attack. The extent of the necessary defense depends on the intensity of the attack. Basically, among several equally suitable defense options, the mildest one is chosen which places the least burden on the attacker. A further limitation of the right of self-defense is based on the criterion of appropriateness. One also speaks here of so-called "social-ethical restrictions" of the otherwise very far-reaching right of self-defense under German criminal law. For example, an evasion will normally not injure the attacker on his part, so that such a passive protective reaction is expected from the attacked person, as far as it is reasonable. Further restrictions can arise in the case of a blatant **disproportion** between the object of preservation and the object of intervention.

These limitations are obviously of great importance for the legitimacy of active defense measures. For example, given the loss of relatively insignificant data in active defense mode, is it permissible to destroy the attacker's entire infrastructure? What applies if the real attacker uses the infrastructure of a third party illegally? Since the constellations described above often elude an abstract legal assessment, decisions are often made that weigh up the various interests involved. Here, the consideration of ethical value decisions can be immensely helpful. For example, uninvolved third parties, which the attacker only takes advantage of, should be spared as far as possible. This can be different, however, if the uninvolved party has nothing in common with the attacker but makes a profit precisely by providing him with resources in good faith.

To weigh up the various basic forms of hacking, which must be combated using active defense measures, hacking must be identified. Then it has to be checked whether the hacking is an attack against which the exercise of the right of self-defense can be considered at all. In a second step, it has to be assessed (if necessary) whether the active defense measure can constitute a permissible act of self-defense under criminal law. Attacks are basically directed against gaps in the system of the attacked person, which are present in every system on a massive scale. At least three different types of attacks can be distinguished:

2.2.1 Domain Name System

This is comparable to the situation when someone anonymously sends a catalog to someone else under a false name.

A (attacker) sends a request to B with the information to send the answer to D (A can freely choose to whom the answer should be sent). B sends a request to C. C sends a reply to B, who sends it to D (attacked). So D is attacked directly by B, but actually by A (the so-called "**attribution problem**").

2.2.2 Attack for Industrial Espionage

Frequently, spy attacks are carried out by injecting a virus into the hardware (keyboard). From there, the virus regularly requests data from the computer. Here, there are hardly any possibilities to detect the attack. Sometimes the attacker also leaves a kind of "back door" through which he can repeatedly gain unnoticed access to the system.

2.2.3 Drive-by-Downloads

The same applies to infection with a virus that encrypts the entire hard disk while a program is downloading.

2.2.4 Attacks via Backups or USB Sticks

USB sticks with a virus, for example, are left in the company car park/office with the inscription "Holiday pictures." Once plugged in, the virus installs itself.

2.2.5 Forms of Attack and Defense

In the case of hacking, data can be affected that are subject to property and assets as well as personal rights. The fact that §§ 202a et seq. of the German Criminal Code (StGB) explicitly criminalize behaviors directed against data (spying, interception, preparatory acts, betrayal of private and business secrets) can serve as an interpretation aid. The regulations protect far-reaching interests in the secrecy of data that are not stored or transmitted in a directly perceptible manner (see Heger 2018, margin no. 1 ff.; Lenckner and Winkelbauer 1986, p. 485; Schmölzer 2011, p. 724; Haß 1993, p. 480 marginal 20; Jessen 2014, p. 37; Schulze-Heiming 1995, p. 37; Schreibauer and Hessel 2007, p. 616; Schumann 2007, p. 675, 676; Eisele 2013, p. 6/1; critical: Hilgendorf 2015 p. 8/49-54; diff. Haft 1987, p. 9). The right to store and use data is therefore a legally protected interest (*Rechtsgut*) and its defense can therefore constitute a justified act of self-defense. One can speak thereby of a comprehensive criminally protected area of stored data. If the target of the attack is data stored by an external provider, such as a cloud provider, this data would have to be defended by the company and constituted an act of defense of a third party.

In principle, intrusion into another computer by means of a specially created program can be evaluated as an attack. To spy or sabotage a computer, one must first circumvent certain protective devices of the computer one is "attacking." Technically speaking, this "bypassing" could already be seen as a kind of "attack." Additionally, attacks can often be linked or imputed to human behavior.

Although a high degree of automatization can be observed when a computer is intruded the creation of a program itself that is used to carry out the attack always requires a certain human behavior (creation/input of a code, etc.). In addition, the activation sequence must be triggered at a certain time.

But indeed, it is problematic that self-defense is only permissible against attacks that are currently taking place (BGH NJW 1979, 2053). Due to the speed of data loss in attacks on corporate IT, the question of when an attack is imminent in such a way that it "already" constitutes a present attack is of particular relevance. The decisive factor is whether further waiting will (significantly) worsen the chances of preserving the good. Based on the assumption that in the case of many attacks, a reaction is nevertheless only possible after the intrusion has already taken place, it must also be clarified whether the attack is still ongoing, i.e., "still" present, since otherwise the legal property could no longer be saved, but at best be restored. Furthermore, an attack is not finished as long as the success of the attack is increased, intensified, or a repetition is feared. An attack is terminated not only when it is finally carried out, but also when it has failed or has been abandoned. If a former aggressor who is already in retreat is attacked, it is an attack against him which may be averted by him through self-defense. If the victim has a kind of "**automatic recoil program**" that triggers a counterblow but does so only after a "reconnaissance phase," it becomes difficult to assume a "presence." Depending on the case, it depends on how long a reconnaissance phase typically lasts and whether or which technical possibilities exist to carry it out "quickly" (i.e., in direct temporal connection to the attack).

2.2.6 Misattribution

On the so-called "attribution problem," i.e., translated into criminal law categories, the difficulty of attributing the attack, e.g., attributing a DDoS attack to the real initiator, the necessity of the defense particularly becomes a considerable problem (Momsen and Savic 2018, § 32, para. 25–25.3). Often it is not possible to identify exactly where a hacker attack comes from in case the hacker covers his tracks by using other systems for the actual attack. Conversely, this means that under certain circumstances there may be no suitable measure to actively counter an attack. In this case only passive measures such as setting up a firewall would be permissible. Thus, the question which measures are suitable at all to counter an attack and are therefore subject to a justification under § 32 StGB is governed by the technical parameters. In the remaining cases, the selection of the means of defense is of particular importance if it cannot be ruled out or is even likely that the active defense will damage the infrastructure of an uninvolved or at best negligently supporting third party. The "attribution problem" leads to an identification problem (especially in professional attacks). The attacked person can only identify the compromised computer as the attacker, but not where the attack, which was already directed against the attacking computer, originated.

2.2.7 Automatic Reaction to Attacks

The problem is that automatic programs that are supposed to actively defend against attacks often only strike after the attack has already passed. Furthermore, the attacks are awfully specific—it is (mathematically) not possible to write a program that is prepared for every conceivable attack. By close observation, it is only possible to find out that the system has changed. If this change was caused by an attack or just by a system error, it has to be investigated additionally. This leads to the fact that the attacks are noticed much too late. Once the attack is complete, however, only the gap is visible, but not whether and which data was stolen.

2.2.8 Suction "Honeypot" Method

Promising is therefore the possibility to install software on one's own system that is "contaminated," so that the hacker who copies and opens it on his system can access and identify the system via the "remote shell" method. Thus, it is possible to, for example, use the hacker's camera and microphone (GPS coordinates, etc.). One problem with this method is that the attack must first be noticed. Additionally, professional hackers are particularly familiar with this procedure and can protect themselves against it.

2.2.9 Summary

If the Active Defense measure merely disrupts the infrastructure of the compromised system, the question of suitability arises. If the attack is essentially complete and can be repeated at any time via another compromised computer, the Active Defense measure becomes unjustifiable due to lack of suitability for defense. The measure would then itself be an illegal attack and could be punished under criminal law. In the context of the obligation to choose an appropriate means of defense, the considerations already outlined must be made on whether and in what way the operator of the compromised computer is (partly) responsible for the fact that his computer was used for the attack.

A glaring imbalance, leading to the exclusion of self-defense by Active Defense, could finally exist if a complex infrastructure is destroyed to prevent the theft of insignificant data, possibly of general importance.

3 State-Conducted Hacking Attacks: New Online Investigation Tools

Another form of hacking of corporate IT can occur in the context of criminal investigations if the company itself or individual employees are suspected of having committed a crime and the law enforcement authorities make use of new online investigation methods introduced by the German legislator in 2017. The German legislator introduced modernized methods of lawfully intercepting telecommunication (Sect. 3.1) and methods that allow a remote search for evidence stored on digital devices (Sect. 3.2).

3.1 Source Telecommunication Surveillance

Conventional methods of surveilling telecommunication under § 100a para. 1 s. 1 StPO (German Code of Criminal Procedure) are indeed sufficient for tapping telephone calls or reading "traditional" text messages. Nevertheless, the law enforcement authorities quickly reach their (technical) limits in messenger communication via social networks such as WhatsApp, Telegram, or the transmission of telephone calls by means of "voice over Internet" (see: BT-Drs. 18/12785, p. 51). The technology used to encrypt the information that is communicated in this area cannot be circumvented while relying on the traditional means.

This is where the so-called "source telecommunication surveillance" comes into play. Unlike conventional methods of surveilling telecommunication, it enables the law enforcement authorities to intercept and route out the content that is about to be communicated *before* it is encrypted and sent by the sender's information technology device. The same applies to the recipient's device, where the communicated content is displayed and stored in a decrypted format again. Irrespective of all encryption efforts, all communicated content can be displayed and accessed by the investigating authorities.

However, this interference not only affects the **secrecy of telecommunications** under Article 10 of the German Basic Law (GG) (see Schiemann 2017, p. 341) but also infiltrates the information technology system of the person concerned and thus encroaches on his or her basic right to the guarantee of the confidentiality and integrity of information technology systems under Article 2 para. 1 in conjunction with Article 1, para. 1 of the Basic Law (see BVerfGE 120, 274). Furthermore, it is difficult to deny a contradiction between the obligation of the state to work towards the promotion of the security of information technology systems, as laid down in § 3 para. 1 p. 1 BSIG, and its interest in keeping open **security gaps** in these systems, which make official infiltration possible (Roggan 2017, p. 829).

Under § 100a para. 1 s. 2 StPO the source TKÜ is permissible if certain facts give rise to the suspicion of a serious criminal offense furtherly specified in § 100a para. 2 StPO, which also weighs heavily in the individual case that is subject to the investigation. Moreover, it is required that the investigation of the facts or the determination of the whereabouts of the accused would be significantly more difficult or hopeless when other investigatory methods are chosen. The Federal Constitutional Court considers the qualification of an offense as "serious" within the meaning of § 100a para. 2 StPO to be covered by the legislator's scope of action in view of the legal interests protected by the respective offense, such as the functionality of the state or its institutions (BVerfG NJW 2011, 833, 836). With reference to the lack of a plausible dogmatic structure the catalog of offenses is nevertheless partly described as "partially disproportionate" (Eschelbach 2017, § 100a para. 10).

Under § 100e para. 1 p. 1 StPO (German Code of Criminal Procedure), only the competent court is authorized to order the source TKÜ. However, under para. 1 p. 2, in case of imminent danger, the public prosecutor's office is also authorized to give the respective order. Each access possibility created is restricted under § 100a para. 5 s. 1 StPO in such a way that only communications-related content *can* be recorded which *should be* recorded according to the ratio of § 100a StPO. Any changes made must be automatically reversed after the measure has been completed—as far as technically possible. The obligation to protect the program used against unauthorized use and information by third parties, as set out in paragraph 5 sentence 2, is also limited to the *state of the art.* Even the phrase "in accordance with the state of the art" acknowledges that opened security gaps can, under certain circumstances, expose the person concerned to the increased risk of such unauthorized intrusion, for example by criminals.

3.2 Online Search

In 2017, the online search was also introduced with § 100b StPO. Under the legal definition of § 100b para. 1 s. 1 StPO, it is carried out by intervention in and extraction of data from an information technology system with technical means without the knowledge of the person concerned. It goes beyond the scope of the source TKÜ, which is limited to communication-related content, with a technically almost identical procedure and covers "all content stored on an IT system," i.e., "stored e-mails regardless of the time of receipt, SMS and WhatsApp messages, photo files, social media contacts, etc." (Roggan 2017, p. 825).

The general requirements under which an online search is permissible are harmonized with the provisions concerning the source telecommunication surveillance presented above. However, the specific offenses that can be investigated on by making use of the online search has been harmonized with the catalog concerning electronic eavesdropping measures directed at the residential property of a suspect, see § 100c para. 1 no. 1 StPO. This parallelism can be explained by the fact that the online search interferes with the **fundamental right of the person concerned to the integrity and confidentiality of information technology systems** under Art. 2 para. 1 in conjunction with Art. 1 para. 1 GG (Graf 2018, § 100b marginal no. 8) to a degree that is equivalent to the surveillance of living space (BT-Drs. 18/12785, p. 54 with reference to BVerfG, Urt. v. 27.2.2008—1 BvR 370/07, marginal 200; Eschelbach 2017, § 100b marginal 3). The catalog of causes of action is composed of offenses which weigh particularly heavily in view of the legal interest concerned and the threat of punishment, but also of those whose prosecution typically encounters greater difficulties in obtaining reliable evidence (Eschelbach 2017 § 100b marginal no. 11). Online searches must be technically restricted and logged.

Although the online search under § 100b para. 3 s. 1 StPO may only be directed against the accused, third parties may also be affected. On the one hand, this is the case if certain facts suggest that the accused uses an information technology system belonging to them or if an intervention in their information technology system is not sufficient (para. 3 s. 2). On the other hand, however, third parties may also be affected within the scope of a permissible measure under § 100b para. 3 s. 3 if this is "unavoidable." Although the unavoidability must be examined in each individual case (Graf 2018, § 100b marginal no. 24), the online search can therefore also collect information from completely uninvolved third parties who are in contact with a defendant (Soiné 2018, p. 502). If the measure is directed against a company employee, content information from communication with other employees, customers, and business partners, for example, may be affected too.

It is also technically possible without further ado to use a "**keylogger**" (so-called "Federal Trojan"), which is now also legally permissible, to spy out the passwords that may be required to take over communication without the knowledge of the person concerned. For communication partners, there is practically no way to determine with whom they are actually communicating, apart from recourse to classic analogous techniques such as the use of voice codes or similar means. For companies, this creates a tension between the obligation to effectively protect personal data and confidential information.

4 Digital Evidence

When companies become the subject of a criminal investigation on the accused or injured side, it becomes important to store digital information in a form that allows them to be used as evidence in procedure. The practical relevance of this aspect derives from the fact that these pieces of evidence can then either be used to exonerate the defendant and/or when it is to be made available to the investigating authorities (usually after prior internal company investigation) to achieve a mitigation of threatened sanctions through cooperation (Momsen 2015, pp. 1234 et seq.; Momsen and Tween 2015, pp. 1027 et seq.; Momsen and Savic 2017; Momsen and Grützner 2017, pp. 242 et seq.) In addition, the digital information must be stored and secured in such a way that business operations can continue as far as possible even if parts of the IT infrastructure and data carriers are seized.

4.1 Characteristics of Digital Evidence

In this sense, digital data cannot be directly used as evidence since the content embodied by them is not directly perceptible. Furthermore, they cannot be compared with documentary evidence. Since, unlike written language, they do not represent ciphers that can be understood by everyone, they require mediation. This, however, creates the risk of information selection and thus of interpretation and reduction.

The **process of transforming** digital data into a process-related piece of evidence has a comparable **potential for interpretation and reduction**. If the conversion to evidence is done by automated processes, the reduction problem and by that the imminent loss of evidence is particularly grave. If the conversion is done by human interpretation processes, the risk of manipulation of the content of the evidence prevails. On the one hand, this applies to the treatment of digital data secured for evidentiary purposes by law enforcement agencies. On the other hand, as described above, data secured for evidential purposes can also become the target of hacking attacks without further ado.

4.2 Importance of Digital Evidence

Leaving aside witness statements, digital evidence has become the most significant form of evidence from a quantitative perspective (see Endicott-Popovsky and Frincke 2007, pp. 364 ff.). Content and information are additionally, but increasingly also exclusively, created and distributed digitally. Business is conducted online; IT systems are found in almost all companies. Text documents as well as photo, video and audio recordings are now mostly created and stored digitally. This results in a considerable enlargement of the data volume as well as the quantity of devices involved in the exchange of information. The "Internet of Things," i.e., the integration of many devices not primarily used for communication, such as cars or "smart" household appliances, leads to an immense number and variety of possible sources of information about the behavior of individual persons. This flood of information must be filtered by algorithm-based so-called "Big Data" concepts and made effectively manageable. In addition, cloud storage concepts lead to the fact that data can be accessed by users from various terminal devices, allowing for changes to be made to the data. The increase in digital information is inevitably accompanied by the growing importance of digital evidence. Alibis can be verified or falsified, motives can possibly be traced and connections between people can also be traced.

Digital data has a considerable potential to change various forms of communication (see Rudolph 2013 for examples). Since criminal proceedings are of course nothing more than a specific platform of communication (Wassermann 1996, Einl. II, Rn. 10 ff.), these changes also have an impact here. Not only in more complex criminal proceedings has the evaluation of e-mails and communication data, so-called "social networks," become the central object of the hearing of evidence on the one hand, and on the other hand the initial collection of such evidence is not carried out by the criminal prosecution authorities themselves, but to a considerable extent within the framework of internal investigations by the companies concerned in accordance with the requirements of their IT/data compliance (the Hanover police force, for example, has made the evaluation of e-mails and communication data a central part of the collection of evidence). The Hanover police have launched a pilot project called "Facebook search" (http://www.handelsblatt.com/politics/germany/pilot project-how-the-police-in-Hanover-search-for-witnesses/7382618.html). The evaluation of the collected data is also often outsourced to private service providers due to a lack of own resources.

Nevertheless, from the perspective of procedural law, the same fundamental questions will have to be asked as for non-digital evidence (the counterpart of digital evidence is not only the analog evidence, but also all other types of evidence not based on perpetuated information) (witnesses, other submissions by natural persons). In addition to the legal framework for the collection and use of evidence, these include questions about the value and quality of the evidence and the scope of the evidence. Like a DNA profile, a cell site analysis can only prove certain facts: a terminal device was active in a certain district at a certain time.

4.3 Contextualization and Misinterpretation

Nevertheless, digital evidence has some distinctive features that must be considered in criminal proceedings. Due to the digital form of the stored information, modifications in various ways are relatively easily (Gercke 2012, p. 713). It is possible for anyone to modify texts created with a computer or to edit pictures and videos without any major difficulties. The corresponding programs are partly available free of charge and offer possibilities for subsequent alterations, which would not be possible so easily with a handwritten document or a photo developed from a negative. This creates a specific uncertainty factor on the correctness of a fact that has been proven with the respective file. The assignment of digital information to a person can also cause problems.

If, for example, a computer that has been used to commit a cybercrime is used by several employees in a company, identifying the individual that is responsible for the crime may be difficult (see Casey 2002, p. 2; Chaski 2005, pp. 1 ff.). It is appropriate that case law on Internet-related crimes does not base probable cause for an arraignment against the connection owner solely on the assignment to his or her IP address. The concrete perpetration of a certain person cannot be determined in this way (see LG Karlsruhe MMR 2010, 68; LG Köln, decision of 20.10 2008—106-5/08, juris; MMR 2009, 291; LG Saarbrücken K&R 2008, 320; mostly cases concerning requests for inspection of records). The difficulties increase even more if the users use anonymization software that is frequently and urgently recommended from the perspective of IT security (see above Sect. 1), potentially leading to a systematic devaluation of the digital evidence (Meier 2012, p. 198).

Furthermore, the initial gathering of evidence, and in some cases even the "creation" of digital evidence (on the process of evidence creation (evidence creation/fabricating the evidence) Marshall 2008, pp. 55 et seq.), is often not in the hands of law enforcement agencies. This creates a more complex and less obvious, but nevertheless significantly higher risk of manipulation and loss of evidence-relevant information than present with most conventional forms of evidence (see Geschonneck 2004, pp. 243 ff.). It has a comparable effect that the parties to the proceedings relying on digital evidence are frequently confronted with such a high volume of data that a selection process in the sense of a reduction to information relevant to the proceedings must take place early on in the preliminary proceedings, usually on the part of the investigating authorities. This step must be comprehensible and verifiable for the other parties to the proceedings, insofar as they are entitled to inspect the files. Of course, this often fails in practice due to two circumstances: The volume of raw data can overtax the data processing capacities of the defense. The investigating authorities themselves also face considerable difficulties. Up to now, the investigating authorities have hardly any elaborate and comprehensible instruments for dealing with "Big Data," and the selection is usually based on subjective criteria, which makes it extremely difficult to reconstruct this process despite its constitutive nature for the hearing of evidence.

A further distinctive feature is the circumstance that the digital file, like the individual digital datum itself, apart from very few special constellations, is in itself completely unsuitable as evidence for criminal proceedings governed by the principles of orality and immediacy. For digital data must necessarily be made visualizable or otherwise perceptible by means of a process of transformation and editing. This process is obviously an overly critical circumstance for evidence. Processing is nothing other than manipulation (in a value-neutral sense) so under certain circumstances, one could even speak of the "production" of the evidence. Both terms are extremely problematic terms from the perspective of procedural law in connection with evidence.

4.4 Digital Data and Evidence Standards

The confiscated DVD or hard disk is merely an object of (judicial) inspection and has no probative value beyond the fact of its existence. Even the circumstances of its discovery are usually reserved for witness evidence. The digitized or digitally stored information requires processing to be usable as evidence in criminal proceedings (see Sect. 2.2 above). For digital evidence, however, this processing procedure has a crucial unique feature: The exploitable information always remains in the context of its digital storage.

For example, the printed e-mail resembles the conventionally developed photo in that in both cases a technical process creates a potential object of judicial inspection (possibly documentary evidence). On the authenticity of the evidence, it is primarily important that the conversion process (digital data to readable text, negative to true-exposure and true-color image (depicted with risks of manipulation or distortion in Marshall 2008, pp. 75 ff.) is technically flawless. To prove this, the person who carried out the processing would have to be called upon as a witness, possibly also as an expert witness. In the case of the text file, however, the raw data and metadata are usually still available (referred to in the following as "context data" for the sake of simplicity (based on "data context," see Marshall 2008, p. 83 in distinction to the data itself or the information content it embodies ("data content," loc. cit. pp. 69 ff.).

Without this, a soundproof of authenticity cannot be provided. They must therefore always be collected in the background of the "perceptible" information, which represents the evidence in itself and in the sense of the law. This creates a need for standards of evidence suitability for these contextual data as well (overview in Casey 2002, pp. 25 ff.; Geschonneck 2004, pp. 64 ff.; Rowlingson 2014, pp. 11 ff.) that shall be presented below (Sect. 5). However, if the context data are indirectly relevant to evidence in that sense, the question arises whether and to what extent their isolated authenticity can or must be verified.

5 "Forensic Readiness" and Digital Compliance

5.1 Concept

With the term "forensic readiness" that has widely been transplanted into the German legal and technological terminology from the common law tradition, the practical implications of the peculiarities of digital evidence with special regards to precautionary measures are addressed. Other than the terminology may suggest, it cannot be understood in the sense of a "readiness for court or trial" but rather as a general term for measures aimed at creating and preserving the "suitability and quality of digital evidence" that may later be used in a procedural context. Therefore, "Forensic Readiness" has a broader, proactive content in the sense that information is created, collected, archived, and documented in such a way that it can be used in the hypothetical case of a later trial. The background to this is, among other things, the endeavor of the persons or institutions involved in this process to ensure that the general and professional standards governing their actions have been observed, effectively eliminating a conceivable liability for errors. "Forensic Readiness is defined as the ability of an organisation to maximise its potential to use digital evidence while minimising the costs of an investigation."[14] "Forensic Readiness" thus primarily fits into the framework of preventive compliance, which is intricately linked to aspects of criminal procedure law (see also the objectives of Compliance Bock 2011, pp. 19 et seq.; Rotsch 2013, pp. 3 et seq.)

5.2 Standards of Evidential Value

Even if not in a criminal procedure-specific framework, "Forensic Readiness" sets requirements, compliance with which significantly increases the evidential value of the information presented. Also, in the context of a German criminal proceedings the so-called "Daubert-Criteria" (Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993), on the case in National Research Council 2009, p. 90) can be helpful. In the absence of a specific test that could be used to determine whether

[14] Rowlingson (2014), p. 1: "A forensic investigation of digital evidence is commonly employed as a post event response to a serious information security incident. In fact, there are many circumstances where an organization may benefit from an ability to gather and preserve digital evidence before an incident occurs" (a.a.O.). Tan (Fn.1), S. 1 definiert wie folgt: "Forensic Readiness" has two objectives: "Maximalizing an environments ability to collect credible digital evidence; and 2. Minimalizing the costs of forensics in an incident response."

(digital) evidence is of the required scientific quality, the U.S. Supreme Court suggested in the Daubert decision (Daubert v. Merrell Dow Pharmaceuticals, Inc, 509 U.S. 579 (1993), pp. 593 et seq.) that several factors should be considered in each case (Casey 2002, pp. 73 et seq.; Ryan and Shpantzer 2008, p. 2) to avoid false positive results (Detailed analysis in National Research Council 2009, pp. 90 et seq. with reference to the commentary on the Fed. R. Evid. 702 Reference to General Electric, 522 U.S. (at 146): "that there is simply too great an analytical gap between the data and the opinion proffered").

These factors are not exhaustive and do not constitute a checklist or a final evaluation standard in the sense of a "final test" (Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993), pp. 593 et seq.) The Daubert jurisprudence represents a further development of the traditional Frye decision of the "District of Columbia Court of Appeals" of 1923 on the handling of "scientific evidence," which outlines the requirements for scientific experts to this day[15] (experience, training, anchoring in generally recognized methods and procedures; cf. Frye v. United States, 54 App. D. C. 46, 293 F. 1013 (1923). These criteria can easily be reconciled with the principles concerning the admissibility of expert evidence in § 244 para. 4 StPO in the German criminal proceedings (cf. Meyer-Goßner 2018, § 244 marginal no. 75).

In criminal proceedings, these criteria can become relevant in two ways. On the one hand, the digital evidence could be presented in such a way that it is already shown when it is introduced into the proceedings that the data collection was carried out in accordance with the afore-mentioned requirements. This would be the ideal conception associated with "forensic readiness." However, to effectively check a digital piece of evidence for its probative value, it is essential to know its potential weak spots. This is because the corresponding request for evidence must meet the requirements of § 219 para. 1 s. 1 StPO and § 244 para. 3 StPO and the respective case law (BGHSt 1, 29 (31); 6, 128 (129); StV 2000, 180; Meyer-Goßner 2018, § 244 marginal 18 ff.; Sättele 2017, § 244 marginal 82 ff.)

In view of the specifics of digital evidence as described above, the history of the origin of the evidence is of primary interest. Evidence relevant to criminal proceedings usually arises in connection with specific incidents, i.e., any criminal offense possibly including crimes committed within a corporate entity. After a corresponding incident, evidence often arises in different places and in different forms. Only some places are known at the beginning of the investigation. Digital evidence can be stored in different media, be it physical storage media such as DVDs or hard disks or immaterial ones such as social networks or cloud storage. Often, the complete picture is only revealed when different storage locations of a piece of information are compared (see Marshall 2008, pp. 19 ff., 85 ff. for details).

[15] National Research Council (2009), p. 93 (m.w.n.)—"... that an expert's testimony is reliable where the discipline itself lacks reliability (...)." In view of the rapidly developing fields of "digital forensics," this is of importance not to be underestimated in terms of the admission of experts. If necessary, this may be a reason for an additional expert within the meaning of § 244 (4) StPO.

5.3 Integrity, Authenticity, Reproducibility

Depending on the place of storage and the knowledge of possible further storage locations, assessments can be made about the integrity and authenticity of the digital evidence. "Integrity" means that evidence must be and remain "unchanged." The degree of integrity should be as high as possible. Doubts may arise, for example, if the collection of evidence was not carried out by law enforcement agencies or if the amount of data was drastically reduced (Marshall 2008, pp. 19 ff., 43 ff.). In both cases, the question arises as to whether the raw data was submitted to check whether corruption or contamination of the data may have occurred (Marshall 2008, pp. 40 et seq.). Authenticity means that the individual piece of evidence is immediately the same as the one originally obtained. Here, problems can arise if digital data has been moved to other media or locations before or after the evidence was obtained. While this does not necessarily mean that the evidence loses quality, nevertheless, one will often have to use context data (meta or raw data) in the above-mentioned sense to ensure authenticity. If the integrity is lost or the degree of integrity falls below a certain level, conclusions can only be drawn to a very limited extent. The same applies to authenticity; guaranteeing this criterion also depends on the possibility of identifying the origin of information. This makes it necessary to evaluate derivative information or contextual data. For example, if there is evidence that a hard disk has been "cleaned," then the flash memory cache on the hard disk will be relevant to the proof in addition to the copy of the hard disk.

Furthermore, the digital evidence must be reproducible. "Reproducibility" means traceability in the sense of a derivable logical chain of complex proofs or information from simpler proofs or data. Closely related to this is the verifiability in the sense of the internal consistency of the digital evidence: pieces of evidence that are by their nature susceptible to manipulation (like many stored data) gain a higher evidentiary value if there are parallel strands of evidence that are altogether consistent with each other. If, for example, a piece of information is stored identically at different independent storage locations, a case can be made for a high evidential value, since it is unlikely that all storage locations have been manipulated at the same time. If you, for example, take a document that may have been edited by several users simultaneously on one platform (e.g., "Google Drive"), the probative value increases if different users have stored the document in an identical manner on their end devices. Furthermore, possible end uses by devices or persons (entities), the environment, restrictions and controls (environment), organization of the relevant IT (organization), infrastructure of buildings, networks, etc. must also be considered. (Infrastructure), workflows (Activities), (data) processing (Procedures) and the data (Data) itself (based on the "Seven-Element Security Model" by Marshall 2008, pp. 56 ff.).

The probative value described by the criteria can be divided relatively well into categories, as *Casey* has shown with the "Levels of Certainty" developed by him (Casey, http://flylib.com/books/en/2.57.1.74/1/—"Levels of Certainty"; see also Casey 2002, p. 70). It should be noted that the problems shown are intensified when the digital evidence was originally collected as part of an internal investigation. On the private nature of internal investigations, this follows the principles concerning a criminal procedure led by the state only to a very limited extent (see Momsen 2015, § 6 B II 2 a, 2021a).

5.4 Compliance and Documentation of IT Forensic Standards

If a company has initially conducted an investigation itself, the quality of the resulting evidence will largely depend on whether the standards of IT forensics have been observed. The same applies, of course, to external IT services that are commissioned by law enforcement agencies to secure and evaluate evidence. The German Federal Office for Information Security (BSI) has summarized in a guideline the requirements for a forensic investigation process in the IT sector, which essentially consider the already described peculiarities of digital evidence (however, these requirements must be put into perspective from a technical point of view, in more detail Rudolph 2013; Momsen and Hercher 2014, pp. 173 ff.). As such these standards are comparable to the "Daubert standards" in terms of methodology. Acceptance of the applied methods and steps is required; these must be described and generally accepted in the professional world. When new methods are used, their propriety must be proven. To ensure credibility, the robustness and functionality of methods must be verifiably given. Repeatability must be possible. If third parties make use of the tools and methods used, the same results must be achieved with the same source material. Secured digital evidence must not be altered unnoticed by the investigation itself. It must be possible to prove the integrity of the evidence. By selecting the methods, it must be possible to establish logically comprehensible connections between events and traces of evidence and also to persons; only in this way a cause and its effect can be linked. Finally, a complete documentation must be created for every single step of the investigation process. In addition, there must be complete proof of the whereabouts of digital traces and the results of the investigations carried out on them, i.e., the traceability of the so-called "chain of custody" (IT forensics guide (footnote 47), pp. 24, 87ff.; see also Casey 2002, pp. 21f.) After all, the most common starting point for questioning the quality of digital evidence is incomplete or missing documentation.

6 Perspectives: The Use of AI and Evaluation of Big Data in Criminal Proceedings and Preparatory Internal Investigations

6.1 Consequences of the Extended Use of Big Data, Algorithms, and AI

New technologies harbor specific risks. These risks can be identified by the key words lack of understanding of the processes, lack of transparency, lack of individual fairness, promotion and reinforcement of existing inequality, lack of valuation and trust-based decisions, and a variety of individual and structural biases by the people involved in the design, process, and evaluation of data processing.

In the field of internal investigations and criminal justice, this can lead to serious misinterpretations and misjudgments, such as the surveillance and prosecution of innocent people or the violation of elementary principles, e.g., the presumption of innocence. Of course, the national and international rules of data protection have to be considered.

A specific problem is the risk of discrimination and reinforcement when attributing possible responsibility through data analysis. This often affects the presumption of innocence. It may be considered as openly discussed if the presumption has to be observed in internal investigations, however, as long as the products of such inquiries shall be used as evidence in later court proceedings, such core rights should better be followed. The former international Secretary General of Amnesty International, *Shetty*, saw the presumption of innocence from Article 6 II ECHR, Article 20 III in conjunction with Article 28 I 1 GC. In a bigger picture, he warned that discrimination against certain groups, e.g., ethnic and religious minorities, can be increased.

The question arises as to who is considered suspicious. Stigmatizing indicators such as foreign appearance or a kind of "typecast" should not be used (see Monroy 2017).

There is thus also the fear that not only potential offenders will be targeted, but also other persons, with the consequence that investigative measures and associated encroachments on fundamental rights might be directed against persons who do not pose a danger. In this context, the question also arises as to the responsibility for forecasts that turn out to be incorrect. Can the investigators or manufacturers exculpate themselves by saying that the software has made a false prediction?

6.2 Quality of Data

The **quality of** the **data** collected is crucial for the quality of the probability statements. The performance therefore depends decisively on what data is used for the probability calculation.

The completeness, correctness, reliability, accuracy, and topicality of the processed information are essential.

This is particularly important because data errors inevitably lead to misinterpretations. Such misinterpretations are sometimes not even noticed because they may correspond to stigmatizing or conventional patterns of thought. However, algorithms are only as objective as the programmers who created them; as the criminological assumptions on which they are based; as the data they use (see Završnik 2018, p. 3).

This is particularly true where the techniques are based on crime statistics because crime statistics do not necessarily reflect the reality of crime, but rather police registration behavior.

With the use of increasingly digitally generated and retained data, the police are often operating close to the limits of what is permitted in the digital domain under European data protection law.[16] The automation of crime prevention through the introduction of forecasting software should reinforce this trend. Already more than 2 years ago, the conference of German federal and state data protection commissioners warned against a "further shift of the police intervention threshold in the forefront of dangers and crimes" (see Monroy 2017—https://datenschutz-berlin.de/attachments/). It is completely unclear today which crimes will be automatically detected in the future and which data sources will be included. With digital investigations, the rule of thumb is that the haystack has to be enlarged to find the needle. There is also a risk of incorrect prognoses, which, according to the data protection officers, is expected especially with the increasing number of preliminary analyses and is associated with significant consequences for the persons suspected in this process.

6.3 Privatizing Surveillance Through Social Networks

Finally, publicly available information from social networks can also be integrated. Corresponding, already pre-filtered data could be supplied by the police authorities themselves. A modern police operations-room nowadays has functions for evaluating trends on Twitter, Facebook, or Instagram. This would enable the police to track hashtags or geodata on Twitter during an operation. For example, it would be advantageous for the situation assessment to have tweets from soccer fans or demonstrators displayed in a geo-referenced manner to draw conclusions about soon necessary operational measures.[17]

[16] Monroy (2017)—"Examples are the rapid use of radio cell queries or the sending of silent SMS as a standard measure in investigations."

[17] Momsen and Bruckmann (2019), S. 20 ff.

6.4 Personalized Data

Since the quality and mass of data is so important, this will lead to a quest to make more data usable for software solutions. The development in the USA makes it clear where the journey is heading.

In Germany, however, there are limits to the collection and processing of personal data under the constitution—and the concept of **personal self-determination** under Article 2 I in conjunction with Article 1 I GC.

The right of informational self-determination is a manifestation of the general right of personality and was recognized as a fundamental right by the Federal Constitutional Court in the so-called census judgement in 1983.[18] The starting point for the Federal Constitutional Court is the so-called general right of personality (APR), i.e., Article 2.1 of the Constitution (Grundgesetz—GG) in conjunction with Article 1.1 GG.

Self-determination in the free development of the personality is endangered by the conditions of modern data processing. Those who do not know or cannot influence what information concerning their behavior is stored and kept in stock will adapt their behavior out of caution. This would not only impair individual freedom of action, but also the common good, since a liberal democratic community requires the self-determined participation of its citizens. "A social order and a legal system enabling this would not be compatible with the right to informational self-determination, in which citizens can no longer know who knows what, when and on what occasion about them."[19]

In the view of the European Parliament, the right to informational self-determination also derives from Article 8(1) of the European Convention on Human Rights.[20]

"Correspondence" as well covers any IT-, online- or web-based communication.

Hence, the Federal Constitutional Court (BVerfG) emphasizes that the use of such systems is basically only permissible for an objectively determined and limited

[18] German Federal Constitutional Court (BVerfG), judgment of the First Senate, 15 December 1983, 1 BvR 209/83 and others—Census—BVerfGE 65, 1.

[19] BVerfG: Judgment of the First Senate of 15 December 1983 (1 BvR 209/83, marginal no. 146). Federal Constitutional Court. 14 December 1983.

[20] Article 8—Right to respect for private and family life.

1. Everyone has the right to respect for his private and family life, his home and his correspondence.

2. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

reason, and that they can only be used without any effort if they are used in response to a dangerous or risky action.[21]

6.5 Differences and Limitations of Prognostic Decisions in Investigations

Using Big Data as well for security purposes and in the field of investigative measures turns out to be the great challenge for the future. A balance between public safety and the personal rights of the individual will have to be found.

Particular attention should be paid to compliance with the threshold to suspicion of a crime and the rights of the accused. In criminal trials, e.g., fairness, is an important issue. The idea of procedural fairness not just means to respect the presumption of innocence but refers to the whole design of the investigation and trial. It is very much related to a balance of powers reflected in several points such as the "Brady Rule," the "Miranda Warnings," and much more or even in the GCP or both the German and the US Constitution and its amendments.

Still, it is an unsolved problem how to translate this idea of fairness into mathematical terms. As *Chelsea Barabas* mentioned (see Barabas 2019, pp. 2–3), bias-based conceptions of validity and **fairness** fail to interrogate the deeper normative, theoretical, and methodological premises of these tools, which often rely on arrest and conviction data to predict future criminal activity and dangerousness. These data directly reflect the allocation of law enforcement resources and priorities, rather than rates of criminal activity.[22]

Another fairness-related problem shown by *Kleinberg* et al. occurs more statistically:[23] A risk score could either be equally predictive or equally wrong for all races or groups with different numbers of preconditions like criminal records or convictions — but not both. The reason was the difference in the frequency with which blacks and whites were charged with new crimes. "If you have two populations that have unequal base rates," *Kleinberg* said, "then you can't satisfy both definitions of fairness at the same time."

The researchers constructed a mathematical proof that the two notions of fairness are incompatible. Especially in the criminal justice context, false findings can have far-reaching effects on the lives of people charged with crimes. Judges, prosecutors, and parole boards use the scores to help decide whether defendants can be sent to

[21] BVerfG, Order of the First Senate of 4 April 2006, 1 BvR 518/02—dragnet investigation—BVerfGE 115, 320. BVerfG, Judgment of the First Senate of 27 February 2008, 1 BvR 370/07 and others—Online search/computer fundamental right—BVerfGE 120, 274.

[22] Elliott (1995); Barabas (2019), pp. 2–3.

[23] Kleinberg et al. (2016).

rehab programs instead of prison or be given shorter sentences. Concerning future predictive measures, a self-fulfilling prophecy is nearly inevitable.[24]

Trust is another important decision criterion maybe impossible to express as algorithmic function. To find a proper and individually just decision, it must be considered if and how much the human being object of the decision can be trusted.

However, as *Cathy O'Neil* wrote, from a mathematical point of view, trust is hard to quantify (O'Neil 2018, pp. 102–104). Because trust can only be earned by personality and character. Those are just individual expectations not to be analyzed by typical Big Data. If so, usually the outcome would be to trust somebody more ore even less because he belongs to a certain group with a percentage XY failing to be compliant. Highly biased and kind of algorithm's cognitive dissonance (see Momsen and Washington 2019, pp. 453 ff.). However, as the German debate on "Forensic DNA Phenotyping (FDP) showed, there is the risk that the whole political approach is biased even when it comes to preventive measures.[25]

6.6 The Presumption of Innocence: Consequences to Big Data and Algorithm/AI-Based Investigations

The collection of data in general is problematic when the fundamental rights of (legally seen as) uninvolved parties are inevitably restricted. the name suggests—it is only through quantity that data quality is created.

When collecting data, a balancing of interests is relevant to the presumption of innocence insofar as many innocent people—and suspects—suddenly find their fundamental rights curtailed, sometimes without knowing anything about it.

The **presumption of innocence dictates** that action may only be taken if there is a concrete suspicion. In particular, the possibilities of technical surveillance involving a large number of innocent people and surveillance without a specific objective should be criticized, although it should be proportionate to the presumption of innocence only in the case of serious offenses.

Otherwise, the use of algorithms on Big Data highly likely would create something like a general suspicion on everyone or even some parts of the community. This not just describes the road to discrimination according to prevalent factors but as well it tends to narrow the civic space in a specific way. As for example, it would become safe to behave in a way that does not match certain criteria the algorithm operates with. Or even better not to make extensive use of civil liberties.

Thus, the presumption of innocence will be conflicted twice—in a more specific procedural understanding but as well in a broader meaning of unsuspiciously using constitutional liberties even extensively without becoming a subject of policing and probable criminal investigation. Although these ideas are neither specifically tied to

[24] Angwin and Larson (2016).

[25] Momsen and Weichert (2018).

AI nor to human rights, we will soon see that the presumption of innocence is strongly connected to both, particularly when it comes to the question of legitimizing or delegitimizing the intrusion of individual rights by administrative and private entities (see Berk 2019, pp. 116 ff., 128 ff.).

Self-assessing companies as governments must also exercise restraint in the sense of proportionality to protect the civil liberties of those concerned (more in detail: Momsen 2021a, b; Momsen and Rennert 2020; Momsen et al. 2020).

References

Angwin J, Larson J (2016) Bias in criminal risk scores is mathematically inevitable. Researchers Say, ProPublica, 30 December 2016. https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say

Barabas C (2019) Beyond bias: re-imagining the terms of ethical AI in criminal law, pp 2–3. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3377921

Berk R (2019) Machine learning risk assessments in criminal justice settings, pp 116 ff.

Bock D (2011) Criminal compliance, 1st edn. Baden-Baden

Casey E (2002) Error, uncertainty, and loss in digital evidence. Int J Digital Evid 1(2):2 ff.

Chaski C (2005) Who's at the keyboard? - Authorship attribution in digital evidence investigations. Int J Digital Evid 4(1):1 ff.

Degen A (2016) § 66. In: Heussen/Hamm (eds) Beck'sches Rechtsanwalts-Handbuch, 11th edn. Munich

Dix A (2014) § 1. In: Simitis (ed) Federal Data Protection Act, 8th edn. Munich

Eisele J (2013) Computer and media criminal law. Munich

Eisele J, Lenckner T (2014) § 203. In: Schönke/Schröder (eds) StGB, 29th edn. Munich

Elliott DS (1995) Lies, damn lies, and arrest statistics. Center for the Study and Prevention of Violence, Boulder

Endicott-Popovsky B, Frincke D (2007) In: Schmorrow/Reeves (hrsg.) Augmented cognition, HCII 2007, S. 364 ff., Berlin

Eschelbach G (2017) § 100a. In: Satzger/Schluckebier/Widmaier (eds) Code of Criminal Procedure: StPO with GVG and EMRK, Commentary, 3rd edn. Cologne

Gercke M (2012) The omitted step from computer to internet criminal law. AnwBl, 709 ff.

Germann M, Voigt P (2017) IT security - not an issue only for operators of critical infrastructures. CR, 93 ff.

Geschonneck A (2004) Computer forensics, 1st edn, Heidelberg

Graf J-P (2018) § 100b. In: Count (ed) BeckOK-StPO, 31st edn

Großkopf L, Momsen C (2018) Outsourcing of professional secrets - criminal obligation to comply? CCZ, 98 ff.

Haft F (1987) The Second Act to Combat Economic Crime (2. WiKG) - Part 2: Computer Crimes, NStZ 6 ff.

Härting N (2005) IT-Security in the Law Firm - The Attorney's Secrecy in the Age of Information Technology. NJW, 1248 ff.

Haß G (1993) In: Lehmann (ed) Legal protection and exploitation of computer programs, 2nd edn. Cologne

Heger M (2018) § 202a. In: Lackner/Kühl (eds) StGB, 29th edn. Munich

Hilgendorf E (2015) In: Doctor/Weber/Heinrich/Hilgendorf (eds) Criminal law special part, 3rd edn. Bielefeld

Jahn M, Palm J (2011) Outsourcing in the law firm: breach of private secrets? The criminal and professional law assessment of a "lawyer's secretariat" outside the law firm. AnwBl, 613 ff.

Jessen E (2014) Access authorization and special security as defined by § 202 a StGB. Frankfurt
Kleinberg J, Mullainathan S, Raghavan M (2016) Inherent trade-offs in the fair determination of risk scores. Cornell University. https://arxiv.org/pdf/1609.05807.pdf
Koch FA (2014) Legal and ethical encryption requirements? Using the example of the legal profession. DuD, 691 ff.
Lenckner T, Winkelbauer W (1986) Computer crime - possibilities and limits of the 2nd WiKG (I), (II), (III). CR, 483 ff., 654 ff., 824 ff.
Marshall AM (2008) Digital forensics - digital evidence in criminal investigation
Meier B-D (2012) Sicherheit im Internet. Neue Herausforderungen für Kriminologie und Kriminalpolitik, MschrKrim, S. 184–204
Meyer-Goßner L (2018) § 244. In: Meyer-Goßner/Schmitt (eds) Code of Criminal Procedure: StPO, 61st edn. Munich
Michaelis P (2016) Cybersecurity: technical requirements of the "measure" according to Section 13 (7) TMG - challenge "state of the art." ITRB, 118 ff.
Momsen C (2015) Internal investigations from a criminal law perspective. In: Rotsch (ed) Criminal compliance. Baden-Baden, pp 1234 ff.
Momsen C (2021a) Digital evidence and criminal defense – how international standards apply in German criminal proceedings. John Jay College of Criminal Justice, The Center for International Human Rights, CIHR-Publications, New York
Momsen C (2021b) Implications and limitations of the use of AI in criminal justice in Germany. John Jay College of Criminal Justice, The Center for International Human Rights, CIHR-Publications, New York
Momsen C, Bruckmann P (2019) Soziale Netzwerke als Ort der Kriminalität und Ort von Ermittlungen - Wie wirken sich Online-Durchsuchung und Quellen-TKÜ auf die Nutzung sozialer Netzwerke aus? KriPoZ, pp 20 ff.
Momsen C, Grützner T (2017) Legal regulation of internal company investigations - gain in the rule of law or unnecessary complication? CCZ, 242 ff.
Momsen C, Hercher N (2014) Digital evidence in criminal proceedings Suitability, extraction, exploitation, revisibility. In: The acceptance of the rule of law in the judiciary, Volume of Material for the 37th Defense Attorney Day, Freiburg 2013, pp 173 ff.
Momsen C, Rennert C (2020) Big data-based predictive policing and the changing nature of criminal justice – consequences of the extended use of big data, algorithms and AI in the area of criminal law enforcement. In: KriPoZ, pp 160–172
Momsen C, Savic L (2017) Explanation of several terms in the fields of digital forensics, digital evidence, cybercrime and cyberinvestigations. In: Grützner/Jakob (eds) Compliance and governance from A-Z, 2nd edn. Vienna
Momsen C, Savic L (2018) § 32. In: von Heintschel-Heinegg (ed) Beck-OK/StGB, 40th edn
Momsen C, Tween D (2015) Criminal compliance in the USA. In: Rotsch (ed) Criminal compliance. Baden-Baden, pp 1027 ff.
Momsen C, Washington SL (2019) Wahrnehmungsverzerrungen im Strafprozess - die Beweisprüfung im Zwischenverfahren der StPO und US-amerikanische Alternativen (Perception bias in criminal proceedings - the examination of evidence in interim proceedings in the German Code of criminal procedure and US-American alternatives). In: Goeckenjan/Puschke/Singelnstein (eds) Festschrift für Ulrich Eisenberg, Berlin, pp 453 ff.
Momsen C, Weichert T (2018) From DNA tracing to DNA phenotyping – open legal issues and risks in the new Bavarian Police Task Act (PAG) and beyond, Verfassungsblog, 2018. https://verfassungsblog.de/from-dna-tracing-to-dna-phenotyping-open-legal-issues-and-risks-in-the-new-bavarian-police-task-act-pag-and-beyond/
Momsen C, Rennert C, Willumat M (2020) Security, populism, human rights and the underestimated role of AI and big data – a review of "security and human rights". In: Goold BJ, Lazarus L (eds) KriPoZ, 2nd edn, pp 183–189

Monroy M (2017) Soziale Kontrolle per Software: Zur Kritik an der vorhersagenden Polizeiarbeit, Cilip, Oktober, 2017. https://www.cilip.de/2017/10/11/soziale-kontrolle-per-software-zur-kritik-an-der-vorhersagenden-polizeiarbeit/
National Research Council (2009) Strengthening forensic science in the United States: a path forward, pp 90 ff.
O'Neil C (2018) Weapons of math destruction – how big data increases inequality and threatens democracy, pp 102–104
Roggan F (2017) Die strafprozessuale Quellen-TKÜ und Online-Searchung: Electronic surveillance measures with risks for the accused and the general public. StV, 821 ff.
Rosenau H (2018) Before §§ 32 ff. In: Satzger/Schluckebier/Widmaier (eds) Criminal Code, 4th edn. Cologne
Rotsch T (2013) In: Rotsch T (ed) Criminal compliance before the tasks of the future. Cologne, pp 3 ff.
Rowlingson T (2014) Digital evidence. Fed Evid Rev, 1 ff.
Roxin C (2006) Criminal law general part, vol 1, 4th edn. Munich
Rudolph C (2013) Forensic readiness in: acceptance of the rule of law in the judiciary, 2013, material tape for the 37th Criminal Defense Attorney Day, Freiburg
Ryan DJ, Shpantzer G (2008). Legal aspects of digital forensics, abrufbar unter: http://euro.ecom.cmu.edu/program/law/08-732/Evidence/RyanShpantzer.pdf
Sassenberg T (2006) Dealing with IT in the law firm. AnwBl., 196 ff.
Sättele A (2017) § 244. In: Satzger/Schluckebier/Widmaier (eds) Code of Criminal Procedure: StPO with GVG and EMRK, 3rd edn. Cologne
Schiemann A (2017) More effective and practical design of criminal proceedings? What is left of the major reform of the StPO. KriPoZ, 338 ff.
Schmölzer G (2011) Criminal offences on the Internet: a substantive consideration. ZStW, 709 ff.
Schreibauer M, Hessel T (2007) The 41st Criminal Law Amendment Act to combat computer crime. K&R, 616 ff.
Schulze-Heiming I (1995) The criminal law protection of computer data against the forms of attack of espionage. Sabotage and Time Theft, Münster
Schumann K (2007) The 41st StrÄndG on combating computer crime. NStZ, 675 ff.
Soiné M (2018). The criminal procedural online search. NStZ, 497 ff.
von Lewinski K (2004) Attorney-client confidentiality and e-mail. BRAK-Mitteilungen 1:12
Wassermann R (1996) Insert II, recital 10 ff. In: Wassermann (ed) StPO Commentary on the Code of Criminal Procedure, vol 1. Munich
Završnik A (2018) Big data, crime and social control. Routledge Frontiers of Criminal Justice

Cyberattacks on Incident Relevant Facilities

Hans-Jürgen Müggenborg

Contents

Under Art. 5 of the Seveso III Directive (Directive 2012/18/EU of the European Parliament and of the Council of 4 July 2012 on the control of major-accident hazards involving dangerous substances, amending and subsequently repealing Council Directive 96/82/EC, OJ EU No. L 197, p. 1), the operator of a hazardous **incident facility** must prove to the authorities that he has taken all necessary measures to prevent serious accidents and to limit their consequences to human health as well as the environment. These basic accident law obligations have been implemented in Section 3 Paragraph 1 and Paragraph 3 of the **Hazardous Incident Ordinance** (Twelfth Ordinance Implementing the Federal Immission Control Act (Incident Ordinance—12. BImSchV) in the version of the notice of March 15, 2017, BGBl. I p. 483, last amended by Article 1a Paragraph 5 of the Ordinance of December 8, 2017, Federal Law Gazette I p. 3882), one based on Section 7 Paragraph 1 and 4, Section 23 Paragraph 1, Section 23b Paragraph 5, Section 48a

H.-J. Müggenborg (✉)
RWTH Aachen, Aachen, Germany

University Kassel, Aachen, Germany
e-mail: info@rechtsanwalt-mueggenborg.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_4

Paragraph 3 BImSchG,[1] and Section 19 Paragraph 1 and 3 of the ChemG-based ordinance of the federal government. The German legislator has retained the term "accident," which corresponds to a serious accident within the meaning of the Seveso III Directive (Hansmann and König 2018, § 2 12. BImSchV marginal number 18). This article examines what this means for **defending against cyberattacks**.

1 Incident Operating Areas

Only the lower and upper class operating areas (Section 1 Paragraph 1 Hazardous Incident Ordinance) are covered by the **accident law**. These are operating areas in which hazardous substances are present in quantities that are reaching or exceeding quantity thresholds listed in column 4 of the list of substances in Annex I of the Incident Ordinance (operating areas of the lower class) as well as the operating areas with hazardous substances in quantities that are greater than the quantity thresholds specified in column 5 of the list of substances in Annex 1 (operating areas of the upper class). Operating areas are often plants that require approval in the sense of the 4. BImSchV including other areas that are subject to the same operator such as warehouses and research laboratories. However, a permit under immission control law is not always required. There are facilities that do not require approval under immission control law (the need for a permit of facilities for the storage, loading, and unloading of substances and mixtures is specified in No. 9 of Annex 1 of the 4. BImSchV dependent on reaching certain storage quantities) but are nevertheless subject to the Hazardous Incident Ordinance. Often these are warehouses under No. 9 of Annex 1 of the 4. BImSchV (4th ordinance for the implementation of the Federal Immission Control Act (ordinance on plants requiring approval—4th BImSchV) as amended by the notice of May 31st, 2017, BGBl. I p. 1440) below the storage quantities specified there, from which the requirement for approval begins.

The accident law leaves the area of individual plants and relates to operational areas. Under Section 3 Paragraph 5a BImSchG, the operating area is the entire area under the supervision of an operator in which hazardous substances in the specified quantity thresholds are actually present or intended for storage in one or more facilities. This includes shared or connected infrastructures or activities or a process that has gotten out of hand and applies, for example, to a fire or an unintended chemical reaction.

[1] Law on the protection against harmful environmental effects from air pollution, noises, vibrations and similar processes (Federal Immission Control Act—BImSchG) of May 17th, 2013, Federal Law Gazette I p. 1274, last amended by Art 3 of 3rd December, 2020, p. 2694.

2 Avoidance of Interference by Unauthorized Persons

When fulfilling the obligation to prevent incidents, **interventions by unauthorized persons** must be avoided in line with Section 3 Paragraph 2 No. 3 Incident Ordinance, unless they can reasonably be ruled out. Moreover, not every conceivable hazard is to be ruled out. Section 3 Paragraph 1 of the Hazardous Incident Ordinance only requires the "precautions required according to the type and extent of the possible hazards." From this it can be deduced that nothing impossible or only achievable with unreasonable efforts is required of the facility operator (Hansmann and König 2018, Section 3 12. BImSchV, margin number 7; LG Trier, ruling of March 21st, 2013—5 K 1021/12, BeckRS 2013, 48552). The required protective measures can be of a technical, personal, or organizational nature (Hansmann and König 2018, Section 3 12. BImSchV, margin number 5).

Interventions by unauthorized persons are not only a factual threat, for example, in the form of external acts of sabotage on the facility, but increasingly also electronically, even if this has so far found virtually no expression in the commentary literature. Cybersecurity is not only required to protect corporate property (Frenz 2016, pp. 121 et seqq.) and also not only affects critical infrastructures within the meaning of the BSI Act (Act on the Federal Office for Information Security (BSI Act) of August 14th, 2009, Federal Law Gazette I p. 2821, as amended by Art. 1 of the Act of June 23rd, 2017, Federal Law Gazette I p. 1885) and the BSI Kritis Ordinance,[2] but also the maintenance of the security of production facilities in particular. At this point it becomes clear that technological progress has made things easier, but also increases the vulnerability of incident relevant facilities.

The **wireless monitoring** and diagnosis of facilities using mobile communications, WLAN, and Bluetooth is on the one hand a benefit, but on the other hand creates weak spots. Fill levels, aggregate states of facilities and plant sections are nowadays subject to remote monitoring and maintenance, also for reasons of cost. Radio data transmissions can no longer be thought of in industrial plants today. Technology allows measured values from facilities, storage tanks, and reaction vessels to be transmitted cost-effectively over long distances. Plant control centers today consist of a collection of computers on which the required data can be called up and the plants can be started up, controlled, and switched off. In this way, it is technically not a problem to operate even complex chemical plants remotely, whereby the techniques of wireless data transmission are sometimes used. Electronic technologies have long made their contribution to increasing plant productivity. Wireless local area networks based on the IE802 industry standard have been

[2] Ordinance on the determination of critical infrastructures under the BSI Act (BSI Kritis Ordinance) of April 22, 2016, Federal Law Gazette I p. 958, amended by Article 1 of the ordinance of Sept. 6th, 2021, Federal Law Gazette I p. 4163. Under the law, facilities from the areas of energy, water, nutrition, information technology and telecommunications, health, finance and insurance, as well as transport and traffic are covered. As a rule, these are not incidents. They are only there, if they have or can develop dangerous substances in the relevant quantity threshold of Annex I Column 4 of the Hazardous Incident Ordinance.

established in GSM/GPRS field devices for some time. It is not only possible to exchange voice, but also graphic and image data, which makes the technology ideal for monitoring, controlling, and adjusting systems from a distance.

As many private individuals have already learned, the **wireless transmission of data** harbors risks. **Hacker attacks**, with which private computers have already been infected in large numbers, can also occur via the Internet. It is therefore important to protect the corresponding process and machine data as well as analysis results from unauthorized access, particularly from any influence on the data. In principle, data can be read by unauthorized mobile devices. It must therefore at first be considered internally who is allowed to view and use which data. This is usually ensured by granting access rights.

Attacks on the data from the outside, also known as **cyberattacks**, are more problematic. For example, terrorists with the appropriate technical knowledge could not only read unauthorized data, but even intervene to control the facilities and cause targeted incidents. In this respect, one can distinguish between three categories of attacks:

- The mere reading of facility data would be the least problematic.
- The influencing of measurement data is more serious, as it can lead to incorrect decisions by the plant operator, similar to the way in which incorrectly displayed measurement data in aircrafts lead to serious mistakes by pilots up to plane crashes.
- The most serious situation is a takeover of the complete plant control by hackers which comprises the possibility of switching off the system or even deliberately causing a malfunction.

2.1 Cyberattacks Known to Date

These scenarios are not just theoretical. The cyberattack on the Iranian nuclear program by the sophisticated malware called *Stuxnet* gained a certain degree of awareness in 2010. The malware manipulated the control components for centrifuges and caused physical damage to the facility. *Stuxnet* was the first cyberattack on an industrial plant and received a corresponding media response.

Another malware called *Havex* appeared in 2014. This was a malware with a modular structure that was particularly aimed at facilities in Germany. Initially, *Havex* was spread via spear phishing. Employees of the attacked company received emails that looked authentic and were intended to induce the recipient to click on a link that was directly infected with the malware.

The so-called **watering hole attacks** occurred later. Attackers observe which websites are frequently used by certain user groups and infect them with malware. In the area of production facilities, this works primarily because the download links provided on the home pages of manufacturers of certain industrial components are infected, so that cyberattacks on incidents and plants can be successful in the way that the facility operator who clicks on the link catches malware. Every facility operator who carried out such a download was damaged in this way. In December 2014, hackers succeeded in bringing a heavily secured blast furnace in the Ruhr area (North Rhine Westphalia, Germany) under their control. The blast furnace could no longer be switched off because of software manipulation, which led to considerable damage. The hackers had gained access to office computers through so-called spear phishing, particularly through targeted fake emails tailored to employees, to penetrate successively into the network and finally gain access to the software control unit of the blast furnace. At first, individual control components at the steelworks failed until the blast furnace could no longer be shut down. Eventually the plant was badly damaged by the cyberattack.

In December 2015, hackers hijacked a power grid in Ukraine and paralyzed substations and switchgear (report on this in Aachener Nachrichten of June 16th, 2018, p. 3). The result was a power outage that lasted for hours.

In 2016, students from the Universities of Illinois and Michigan laid out USB sticks loaded with spy software in prominent places. The finders of the USB sticks connected and opened them on their devices. More than half of the USB sticks then spread malware on the affected computers.

At the end of 2017, hackers attacked a power plant in Saudi Arabia with the presumed aim of destroying it. The attack was only noticed and repelled because the malware accidentally triggered a safety shutdown of the power plant. The fact that the hackers almost succeeded in intervening in the control of the plant should give cause for concern, although the power plant has a Triconex hardware security architecture that is also used in many German industrial plants.

2.2 Cyber Insurance and Firewalls

Companies can now protect themselves against the consequences of a hacker attack with **cyber insurance**. To do this, however, they have to demonstrate technical and organizational security requirements in the company, without which an insurance cannot be obtained. In these cases, the insurer always carries out a risk assessment and provides, among other things, questions about the protection of the IT systems by firewalls. The firewalls set up electronic obstacles that are intended to prevent hackers from overcoming the system barrier. However, the hackers also work at the highest technical level and try to overcome the barriers as if it were a sport. Therefore, firewalls must always be adapted. In particular, they need to be strengthened and kept up-to-date, which is not always the case, especially with medium-sized companies. While hackers used to attack PCs and servers, industrial plants are now increasingly their target. This was also the topic of the world's largest IT security conference, the Black Hat Conference in Las Vegas, in 2017. The question now is what requirements must be placed on facility operators to prevent cyberattacks via electronic paths.

3 Recommendations of the Commission for Plant Safety (KAS)

The Commission for Plant Safety (KAS), which advises the federal government under Section 51a BImSchG, has issued the KAS-29 guideline (Guideline "Special Requirements" for safety technology and safety organization to support plant personnel in emergency situations with special consideration of the KAS-20 guideline, KAS-29 from February 2014). One requirement is to ensure that electrical safety devices are decoupled from one another (No. 5.1 KAS 29). This also applies in the interest of limiting the effects of accidents, because overvoltage after a short circuit or lightning strike trigger less consequential damage.

The Commission for Plant Security has dealt with the protection against cyber-physical attacks in more detail in the guideline KAS-44 (Guidelines of the Commission for Plant Security for Protection against Cyber-Physical Attacks (KAS-44) of November 23rd, 2017). There, too, the networking of existing IT and OT systems (OT systems = Operational Technology Systems, i.e., systems of industrial engineering) in operational areas is identified as weak point for deliberate disruptions to normal operation. It is required that the measures to guarantee IT security in the security management system, based on the security concept under Section 8 Paragraph 1 of the Hazardous Incident Ordinance, must be documented and implemented as well.

The management of the company operating the plant is generally responsible for IT security (No. 2 KAS-44). Therefore, the management must develop an IT security guideline for the organization and regularly adaption to the constantly changing framework conditions. To achieve the goal of IT security, clear organizational structures and processes must be created.

According to No. 3 KAS-44, an essential element for this is sensitization and instruction of all employees and third parties who can directly influence IT security precautions. Employees must be trained in the dangers of cyberattacks so that they can respond appropriately. The effectiveness of the measures must be also checked regularly.

The Commission for Plant Safety also points out in No. 4 KAS-44 that, under Section 3 Paragraph 2 No. 3 of the Hazardous Incident Ordinance, the operator must consider unauthorized interference when determining precautions to prevent hazardous incidents. Relevant **for IT security** are all parts and components of facilities, the manipulation of which by a cybercriminal has an indirect or direct effect on the functional safety of the plant (No. 4 KAS-44). For example, the following are mentioned: security-relevant facility parts, components, parts, security-relevant software, all network entry and exit points to other networks, all IT systems outside the production area from which a communication relationship can be established in the production area, and all the operating area appreciable safety-relevant documentation. To record these parts of the facility and components, a network architecture image must be created in which all transmission protocols are considered in the representation of the communication relationships.

Furthermore, in No. 5 KAS-44, the Commission for Plant Security demands that IT security must also be an integral part of all plant construction phases and their integration into the operational area up to commissioning by the operator is crucial. IT security requirements are already formulated in the concept phase and detailed and implemented in the following phases by the system integrator, who is responsible for setting up the facility.

In addition, in line with No. 6 KAS-44, a risk management system based on, for example, the risk management according to ISO 27005 (currently: ISO/IEC 27005: 2018 (E)—Information technology—Security techniques—Information security—Risk management, July 2018) must be set up to ensure permanent IT security. This essentially consists of a risk identification, a risk analysis, and a risk assessment. The basis of the risk identification is based on the register according to No. 4 KAS-44, the currently existing hazards for the operational area. The risk assessment evaluates the effectiveness of existing protective measures in relation to current risks. If the protective measures prove to be ineffective, appropriate measures must be taken to reduce them effectively. Since the avenues of attacks are constantly evolving, the risk assessment must be repeated regularly without a deadline being specified in the guidelines.

Furthermore, in No. 7 KAS-44, reference is made to the fact that it is necessary to identify and evaluate IT security incidents early enough to be able to take the technical and organizational protective measures required to prevent incidents under Section 3 Paragraph 1 of the Incident Ordinance. In this way, suitable measures can be taken to avoid incidents in the future and incorporated into the risk management system.

Moreover, according to No. 8 KAS-44, suitable measures for restoring IT security after an incident are to be specified. It is shown to be important that the employees are educated accordingly and, if technically possible, trained in simulated scenarios. The effectiveness of these measures must be checked regularly as part of risk management.

4 Further Considerations

With the guidelines in KAS-44, the Commission for Plant Safety has submitted relevant general statements and clearly pointed out that the avoidance of cyberattacks is part of the basic accident law obligation to prevent unauthorized interference (Section 3 Paragraph 2 No. 3 Hazardous Incident Ordinance). Only with the help of a continuous review and implementation of security measures an operator can prevent hacker attacks of the type described and thus attacks by unauthorized persons. However, the statements are kept quite general, which is understandable against the background of the constant advancement of technical possibilities, which are being further developed by the hacker scene. The security measures to be required are not just technical measures, because cyber security is not a ready-to-buy product on the market.

4.1 Raising Employee Awareness

The training and **sensitization of one's own workforce** is crucial for die success of security measures. Of course, external storage devices must not be used. However, this requires a wide range of organizational measures, especially in the area of **security management**, awareness-raising and contingency planning. In addition, every employee must be informed about spear phishing so that they are aware that emails that look deceptively real can already represent an attack. Under no circumstances may attachments or links to such e-mails be opened or clicked. However, since the attackers are constantly upgrading their technology, training courses and instructions must be, as already mentioned, repeated regularly, at least once a year, and raised to the latest technical standard. Ultimately, it is an eternal cat-and-mouse game between plant operators and hackers, which the plant operator can only win, if he always works on the technical and organizational front of cyber defense. Cybersecurity is therefore not a temporary project in which only technical precautions for firewalls and the likes play their part, but also an ongoing task of employee training and instruction. The organizational measures must then also be regularly trained and checked if they are to have a lasting effect. In the event of misconduct, appropriate action must be taken.

4.2 Review of Employees in the Hiring and Onboarding Phase

A general **organizational measure** should be to get a precise picture of future employees, which mainly concerns the recruitment phase, but does not exclude later reviews. A review according to the SÜG3 by the Federal Ministry for Economic Affairs and Energy is only possible outside the public sector in the cases of Section 24 Paragraph 1 SÜG.

What is always possible and reasonable for the operator in his function as an employer is, for example, checking the employee on social networks such as Facebook, Twitter, and Instagram. Insofar as the employee's publications suggest a violent or radical attitude on the part of the employee, that employee may not be deployed in areas that allow an influence on the safety of the operating area. An employer who neglects audits and uses people who are recognizable as radical or ready to use violence in security-relevant areas of a company exposes himself to the charge of organizational fault (Mehrbrey and Schreibauer 2016, pp. 75 et seqq.). Organizational negligence cannot only result in criminal liability, for example, due to negligent homicide, negligent bodily harm or negligently caused environmental pollution (Sections 324, 324a, 325, 325a, 329, 330a StGB), but also to liability of the company for resulting damage (e.g., under Section 823 et seqq. BGB, UHaftG, etc.).

4.3 Technical Protective Measures, Especially DMZ

It goes without saying that the isolation against attacks from the outside plays a decisive role in IT security. If necessary, devices and IT systems must be equipped with hardware in such a way that external storage media such as floppy disks, external hard drives, and USB sticks cannot be connected because the slots required are dispensed right from the start.

With the right combination of technical and organizational measures, many attacks can be averted. However, as long as there is an electronic connection to the outside world, be it in the form of a remote-control system or only in the form of the transmission of essential system data via the Internet, WLAN or Bluetooth, the necessary security against cyberattacks cannot be guaranteed.

Since the main route of attack for cyberattacks is the Internet, it should be or become standard that a **DMZ (demilitarized zone)** between the public network and the intranet is set up with the system control. The DMZ is an independent network that is connected in-between the system operator's internal network, which includes the system control and monitoring, and the external network (World Wide Web). All servers such as web servers, mail servers, and the application gateways are located in the DMZ. The DMZ thus functions as a buffer zone that separates the internal and external networks from one another through firewalls. With that said there are two technical solutions to this. With the simpler one, there is only one firewall between the public network under DMZ control. With the more sophisticated there is a second firewall between the DMZ and the intranet. If it is ensured here that both firewalls come from different manufacturers, it becomes very difficult for external attackers to access the intranet and thus the system control through both firewalls. Since firewalls can also have security gaps, this prevents the entire system from being overcome by a manufacturing flaw in the event of a security gap. The second firewall would then fend off the attack. Such a protection is not yet available in all accident systems, but it should become standard.

Of course, it is important to ensure that the firewalls are always kept up to date with the latest security technology. There are now system solutions available on the market for this purpose.

5 Conclusion

To prevent cyberattacks on incident operating areas, which is one of the basic legal obligations under Section 3 Paragraph 2 No. 3 of the Incident Ordinance, a careful IT security infrastructure must be set up and maintained. For this purpose, certain minimum technical requirements must be provided, and the operating areas must be secured internally as well as externally, above all by a DMZ with, ideally, a double firewall. In addition, there is a corresponding sensitization and training of the own workforce and regularly involved employees from third-party companies, who must develop a feeling for not falling for deceptively real imitation websites and e-mails and not opening them to avoid installing malware as far as possible.

References

Frenz W (2016) Industrie 4.0 und Datenschutz im fairen Wettbewerb. EuZW, 121 f.

Hansmann K., König JM (2018) In: von Landmann R, Rohmer G (eds) Umweltrecht, Loseblatt, Effective: 07/2018

Mehrbrey KL, Schreibauer M (2016) Haftungsverhältnisse bei Cyber-Angriffen – Ansprüche und Haftungsrisiken von Unternehmen und Organen. MMR, 75 et seq.

Industry 4.0: Prosecution Practice

Markus Hartmann

Contents

1 The Digital Crime Scene

Crime is always a reflection of social and economic conditions. This applies to the realm of analogue life as well as to the digital one. Business processes and many dimensions of business and private life are increasingly shifting to the Internet or are

Translated by Michael Quandt, RWTH Aachen University

M. Hartmann (✉)
Department of Cybercrime (Central and Contact) in North Rhine-Westphalia, North Rhine-Westphalia, Germany
e-mail: markus.hartmann@sta-koeln.nrw.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_5

significantly influenced by network-based infrastructures. Crime follows this trend. Thus, modern information and communication technologies are used extensively to commit crimes. First, **manifestations of analogue crime** simply diffuse into **cyberspace**. From the point of view of a commercial criminal, robbing a bank is far less attractive today in a simple cost-benefit calculation than manipulating the disposal limits of electronic payment cards, in order to then withdraw the loot in a "bank robbery 2.0" at the ATM safely like any ordinary customer. Even everyday images of crime are increasingly materializing online. Investigative practice in the general criminal law department leaves room for the hypothesis that classic face-to-face insults are becoming less common than defamatory statements on social media.

In the area of crime to the detriment of companies, classic fraud has recently experienced a relevant renaissance with the so-called **CEO fraud**. This is a form of offense in which the perpetrators gain information in advance, not only about the forms and channels of communication but also about the authority to sign and dispose. They use this information to approach an authorized person in the company using a plausible legend precisely timed in order to induce them to transfer large amounts of capital internationally. During a managing director's trip abroad, for example, it is precisely asserted that there might be an opportunity to acquire a company, but this is still to be treated with the utmost confidentiality. Therefore, the addressed employee, as the "only trustworthy person," can promote within the company by responding to the telephone call of an alleged legal advisor of the managing director to transfer millions of cash to non-European countries. Depending on the corporate culture and the safeguards implemented in the operational processes, such attacks are astonishingly successful. The FBI has collected cases that have become known worldwide and records the damage potential that can be assigned to them. The current statistics record more than 78,000 individual cases worldwide with a potential damage of over US $ 12 billion for the period October 2013 to May 2018 (FBI Public Service Announcement I-071218-PSA). Looking back at the crime phenomenon, doubts about its assignment to the area of cybercrime appear justified. However, without the high availability of company-specific information on the Internet, building a plausible legend from the perpetrator's point of view would hardly be affordable. Undoubtedly, however, the area of "CEO fraud" can be assigned to the area of organized crime in view of the extraordinarily high sums of damage and the extremely professional approach of the perpetrators.

In addition to this **digital-analogue form of crime**, an area of specific information and communication crime is emerging which, by exploiting the technical and communication infrastructures of the Internet, creates new crime phenomena. The threat from **botnets**, **ransomware,** and **cyber espionage** has a specific character thanks to its highly technical components. These technical manifestations of cybercrime harbor a particularly threatening potential for the digital economy. The catchphrase of "Industry 4.0" can be traced back from a criminalistic point of view to an increased exposure of companies in the digital world, which is a result of networking within all company resources. Although this form of networked digitization undeniably brings considerable resources and advantages in efficiency with it, it also opens up new attack vectors that pose considerable challenges to both companies as potentially affected parties and law enforcement authorities.

From a societal perspective, cybercrime has become a relevant threat in the meantime. What all statistics have in common is that they can at best put a spotlight on reality with a view to the considerable **dark field of cybercrime**. But even in the cone of light of this limited knowledge, figures of almost **€ 55 billion in economic damage** per year just for the Federal Republic of Germany (published by Bitkom, Wirtschaftsschutz Study 2017) appear suitable to adequately assess the importance of the topic of "combating cybercrime." Cybercrime **ranks third** behind **corruption and drug offenses** worldwide in terms of the financial damage it causes (Lewis 2018). This finding is supported by the PwC Global Economic Crime and Fraud Survey 2018 (Lavion 2018). In a survey from a wide range of industries, companies named cybercrime as one of the three most relevant manifestations of white-collar crime. If companies are asked how they are individually affected, more than 75% of companies in Germany consider themselves affected by cyberattacks (Bitkom Research, Status Quo 2017).

These findings, which are dramatic from a legal-political point of view, result mainly from the increased level of **organisation among cyber criminals**. In the meantime, they have come together to form an underground economy that is networked, based on division of labour and it enables criminals with little technical means to commit high-quality cyberattacks. The still widespread assumption that a cybercriminal requires a high level of IT knowledge to be successful in this area of crime is put into perspective in law enforcement practice. "**Cybercrime as a Service,**" meaning an infrastructure to purchase any tools and services for committing cybercrime online, is the technical basis of life in many areas of digital crime.

"Cybercrime as a Service" is more than just marketing-appropriate packaging. A network economy based on the division of labour between the perpetrators can only develop if the underlying criminal market is able to generate a critical mass of revenue opportunities. For the understanding of modern cybercrime, according to the experience of investigative practice, the understanding of the **economization of the form of crime** is of decisive importance. An economic cost and benefit calculation between income opportunities, risk of discovery, and impending criminal sanctions is unfortunately all too often in the interests of the perpetrators at the moment. If, for example, the implementation of a so-called DDoS attack, meaning the creation of an overload situation in a network-connected resource, requires a financial investment of rental costs for the attack infrastructure of the amount of only US $ 66 per attack (Schwarz 2016), the perpetrator has an effective lever to commit digital blackmailing. The actual act amounts to a digital adaptation of the classic extortion of protection money. Compared to the threat of potential fire hazards, the cybercriminal carries out a short-term, effective and, for him, occasional overload attack on the digital resources of a company that causes little effort. He then demands a manageable amount of capital in relation to the damage potential to refrain from further attacks. From the point of view of the attacked victim, this gives rise to serious considerations regarding this trade-off. From the perpetrator's point of view, the limited monetary claim also pays off in individual cases since the attacks scale well without significant additional effort. It is so easily done because the networks used for this factual scenario consist of remote-controlled computers of uninvolved third parties hijacked by malware, so-called botnets. These are available for rent on the Internet in almost any bandwidth.[1]

The prosecution of **botnet crime** is a particular challenge for law enforcement officers. Botnets are a technically complex network of compromised computers, the so-called bots, and regularly multi-layered control levels, the so-called command and control infrastructures. If you add legal difficulties, such as the example from the international nature of a botnet, which often has technical infrastructure components in numerous countries, then the high costs of successfully penetrating such an attack infrastructure becomes clear.

"Cybercrime as a Service" manifests itself particularly impressively on **digital marketplaces on the Darknet**. All kinds of tools and services for committing digital crimes are offered here, with the operators of the platforms themselves limiting their contribution to the operation of the infrastructure. While the number of registered users of such darknet marketplaces can sometimes be in the tens of thousands, the UK National Cybercrime Unit estimates that only approximately 100 to 200 people are actually responsible for the development, provision and sale of the particular tools.[2] The operators of the platforms benefit regularly in form of commissions for criminal transactions carried out through their marketplaces. From a legal policy point of view, it seems worth considering criminalizing the operation of an incriminated platform in and of itself, because if the infrastructure is destroyed, large parts of cybercrime would be deprived of its breeding ground.

Ransomware is another current form of digital crime. After infiltrating a target system, this blackmail software encrypts all user data there with a system-specifically generated key and demands a ransom for revealing it, usually in a virtual payment unit such as Bitcoin. Ransomware is a mass phenomenon on the Internet and affects individuals and companies alike. If ransomware successfully infects a company, it can cause enormous financial damage. For example, the ransomware "wannacry" is said to have caused damage of up to US $ 4 billion in 2017. Worldwide, companies are said to have suffered **losses from ransomware of more than US $ 8 billion in 2017** (BSI situation picture 2018).

While the spread of ransomware is like a digital scatter shot in the dark of the Internet, companies are increasingly threatened by targeted attacks. Permanent and technically high-quality compromises of corporate networks with the aim of anchoring the attacker in the long term for the purpose of tapping the data communicated in these networks are referred to as "**Advanced Persistent Threat,**" or **APT** for short.

[1] Overall, very instructive on the DDoS image of offenses: "Insight into the Global Threat Landscape - NETSCOUT Arbor's 13th Annual Worldwide Infrastructure Security Report" (report in its entirety available at https://www.netscout.com/report/). Accessed on February 13, 2019.

[2] See https://www.databreachtoday.com/how-do-we-catch-cybercrime-kingpins-a-8283. Accessed on February 13, 2019.

What these attacks have in common is an inspection method that is specifically tailored to the individual structure of the attacked company and the technical resources used there, with particularly high-quality attack instruments. Based on the analysis of the attack tools used, investigative practice shows a high degree of professionalism on the part of the perpetrators, which makes an assignment to the area of professionally organized or third-country-induced perpetrator structures appear possible.

The particular danger of APT attacks lies in the long period between the intrusion of the perpetrator and the discovery of the incident. In this time window, which current studies for the European region indicate as almost six months (FireEye M-Trends 2018), the perpetrators can steal the digital assets of the attacked companies without hindrance. Here it is a particular challenge for companies to upgrade their own infrastructures and security mechanisms for the time-critical detection of attack traffic. For their part, the criminal prosecutors can only hope for a concrete success of the investigation if they have a technical competence that is not inferior to the perpetrators, especially in the analysis of the course of events.

The area of particular digital crime to the detriment of economic actors has recently been increasingly supplemented by attacks on the **digital supply chain**. It stands to reason that this is a reaction of the perpetrators to the increased technical security precautions of companies. The direct attack is avoided. However, penetration into the network infrastructures of software suppliers exploits the relationship of mutual trust between the software manufacturer and the company using it. The software update appears to the common detection mechanisms as legitimate network traffic. If the update server is compromised, the attack software can often anchor itself in the target network without any further defense mechanisms. The "**Petya/notPetya**" case, in which the update server of a company software provider was successfully infected with a malicious update, demonstrated the devastating effects of such an attack scenario. Some companies claim to have sustained losses of up to US $ 300 million because of this incident alone.[3]

With that said, no list of common manifestations of cybercrime can claim to be exhaustive. The most prominent overall tendency of cybercrime is its constant **drive to innovate**. Attack scenarios that are successful today may be ineffective tomorrow or be replaced by even better and more effective attack options. Digital corporate security can therefore never be more than a current status but should be a goal that must be pursued through constantly renewed defense mechanisms. The high speed of innovation poses particular challenges for law enforcement officers. It encounters a persistent set of rules of substantive law and criminal procedural requirements, the renewal of which can hardly be carried out with even approximately the same rate of adjustment. It is even more astonishing that large parts of the criminal procedural rules, on the basis of which cyber investigators work, come from the pre-digital age. This raises considerable and so far unresolved questions of legal policy.

[3] See http://www.spiegel.de/netzwelt/netzpolitik/moller-m-rsk-cyberangriff-kosten-reederei-hundderte-millionen-a-1163111.html. Accessed on February 13, 2019.

In addition to the already mentioned criminality of the operation of criminal infrastructures in the network and the peculiarities of botnet crime, above all **criminal procedural difficulties** are to be named. The "technology offenses" in the narrower sense (for example §§ 202a ff., 303a ff. StGB), which justify the initial suspicion in an early phase of the investigation, do not leave room for high-quality technical investigative measures according to the legislative concept of §§ 100a ff. StPO. Without a comprehensive and timely backup of all computer forensic traces of a digital crime, it is not possible to clarify the course of the crime or to identify the suspect responsible. Creating a basis for encroachment in accordance with fundamental rights is not accessible to simple solutions and will determine the legal and political discussion in the near future. This basis has to comply with criminal procedures and simultaneously meet the requirements of the investigation as well as the particular sensitivity of digital personal data.

2 Digital Defense Strategies

Despite the complexity of the threat situation, concrete, easily to be implemented, albeit not cost-neutral conclusions can be drawn based on practical experience in criminal investigations.

2.1 Digital Leadership Culture

IT security must be considered as a priority of any company today. As shown, the damage potential of current IT attacks is of such magnitude that there is a risk of significant damage. IT security is initially a cost factor, but a lack of it can mean the extinction of a successfully attacked company. Cyber compliance, which is well understood in this sense and aggressively implemented by company managements, therefore sees IT security as one of the main priorities of corporate entities, the importance of which increases with the degree of digitization.

2.2 The Threats Posed by CyberCrime Work Across Industries, Sectors, and Company Structures

Due to the easy availability of high-quality attack tools and the low cost of a cyberattack, from the perspective of a cybercriminal not only scenarios in the context of "global players" and large corporations pay off, but in **medium-sized businesses** as well. If such companies are technology leaders in their market segment, the element of **industrial espionage and competitive spying** through IT compromise is added as a risk vector (Bollhöfer and Jäger 2018). Cyber awareness in the sense of a comprehensive awareness of the vulnerability of one's own IT systems and the security mechanisms required for protection is therefore of crucial importance in every company form and size.

2.3 Awareness Alone Does Not Save a Company

While a few years ago the companies affected by cybercrime were not sufficiently aware of the dangers of it, a significant change in awareness has taken place since then. Today there is more of a **deficit in implementation and action**. In representative surveys, more than 50% of companies state that they have already been attacked by typical manifestations of cybercrime and a further quarter of them think this is possible (Bitkom, Wirtschaftsschutz Study 2017). At the same time, technical security measures are only actually implemented in a fraction of the existing companies. Almost every company is familiar with the classic IT security triad consisting of password, firewall and virus scanner. However, even measures that should be part of the standard set of instruments of every corporate communication—such as encrypted e-mailing—are only actually used by some of the companies (Bitkom Research, Status Quo 2017). It looks even worse with **high-quality measures**, such as the use of pentesting, meaning a simulated attack on one's own network for the purpose of discovering and eliminating weak points. Less than half of the companies use such high-value measures (Bitkom Research, Status Quo 2017; Bitkom, Wirtschaftsschutz study 2018). It must therefore be stated, that only projected or planned IT security is not yet a functional one.

2.4 There Is No Total IT-Security

More than half of the companies are of the opinion that the design of their IT security measures can completely prevent cyberattacks (Bitkom Research, Status Quo 2017). From the practice of law enforcement in the field of cybercrime, however, it is easy to deduce without any restrictions that **what can be hacked will be hacked**. IT security measures always only reflect the state of knowledge of their respective implementation. Nothing is as ineffective as yesterday's virus scanner. Each security appliance, in turn, is susceptible to specific security risks. According to the general experience of law enforcement in this area, only the foreseeable time when a company is compromised has not yet been determined. It is therefore crucial to take precautions to identify successful attacks in good time. Against this background, particularly the long periods spent in APT scenarios (FireEye M-Trends 2018, see 1. above) harbor an urgent appeal for structural reforms.

2.5 The Crisis Is Certain

The inevitability of a cyber security incident necessitates the timely establishment of a crisis management system. There is a lot of catching up to do here, as less than half of all companies can fall back on **emergency management** (Bitkom, Wirtschaftsschutz study 2018).

In particular, the high dynamics of a cyber crisis and its unpredictability in terms of content lead to a considerable burden on the **crisis management structures** in companies. First, clear authority to issue instructions is critical to success: The spread of ransomware can be effectively limited by switching off the company's own network. This will only happen if those responsible in the IT department or the operational network administrators themselves are granted the decision-making authority to order such far-reaching measures, even if this entails considerable restrictions in the availability of the IT systems.

Hardly any company is able to deal with a comprehensive cyber security incident with its own resources and employees alone. It is therefore advisable to define the group of external service providers and service companies that will be called in advance to provide support in the event of a cyber crisis.

2.6 Repression Is Prevention

The commercialization of cybercrime (Lewis 2018) is the engine of its growth. At the same time, however, it also offers an effective lever for combating it. Only effective law enforcement can make a significant contribution to increasing security in the digital space. Merely eliminating security gaps identified during incident processing and not **filing a criminal complaint** in the event of an attack is equivalent to an invitation to repetitively compromise the company with the next security gap found. On the other hand, if the support infrastructure, which is essential for "Cybercrime as a Service," can be effectively contained, the entire field of cybercrime offenses will be deprived of its breeding ground. In this respect, the division of labour in cybercrime is also an effective starting point for combating it.

3 Of Myths and Legends

In less than a third of cases do companies affected by cybercrime involve government agencies in dealing with a cyber security incident (Bitkom, Wirtschaftsschutz Study 2017). This suggests a profound **crisis of confidence**, the causes of which are often based on reservations about the law enforcement authorities, which, however, do not stand up to critical and practical consideration.

3.1 Criminal Charges Expose the Company to the Public

Cybercrime acts—like all technical issues—are extremely complex to embed in a public **communication strategy**. In this respect, the concern of many companies is understandable that being mentioned in the context of a cyber security incident could give the public the devastating impression of inadequate IT security in the company. However, this view overlooks the many possibilities of a clear and effective communication strategy that can result from close consultation with law enforcement authorities. Law enforcement authorities are specialists in classifying facts and describing technical events. They all have professional press departments that adequately take corporate matters into account, also in the context of an authority's obligation to reveal the truth. A coordinated communication strategy helps to avoid misunderstandings and supposed contradictions. The assumption of many injured parties that they can keep a lid on an issue is hardly reliable in times of high public sensitivity to cyber security incidents, when there are extensive reporting obligations to supervisory authorities and those affected, and not least because of the communication behaviour of the perpetrators themselves. A responsible communication strategy must therefore start from the proliferation of meaningful information.

3.2 Investigations Cause More Collateral Damage Than Good

Although it is true that law enforcement authorities are obliged by law to clarify the facts of the matter and to investigate the course of events and suspects (Paragraph 152 Section 2 of the Code of Criminal Procedure), the **principle of proportionality** ensures that the interests of the injured party are effectively considered when implementing measures to preserve evidence. The still widespread assumption that the preservation of evidence in a company necessarily equates to the physical safeguarding of entire server parks has no basis in view of the high level of professionalism of the police stations that carry out such measures. The usual case of the investigation procedure is the IT security measure that has been agreed on and implemented in cooperation with the companies concerned.

3.3 Nobody Likes To Have the Prosecutor at Their Footstep

The uncomfortable feeling that, in the course of their investigations into cyber security incidents, law enforcement authorities also use the IT resources available to them in the course of their investigations into cyber security incidents to investigate the company's previously unknown misconduct, is unfounded. The prosecutors focus on the facts on which the investigation is based. The "intended **accidental**

discovery" is a popular but incorrect legend. The principle of legality forces the law enforcement authorities to prosecute these offenses in cases in which the suspicion-based investigations reveal evidence of other criminal offenses. However, this should also be in the well-understood interests of the companies concerned. If, for example, child pornography images are found during the evaluation of a computer infected by malware, the company affected by the initial investigation will also want to know whether and, if so, which employee is responsible for storing these image files.

3.4 Light Into the Darkness

If a criminal complaint is filed in less than a third of the cases (Bitkom, Wirtschaftsschutz Study 2017, cf. above 3.), this inevitably leads to the upkeep of a considerably high number of undetected crimes in this area.[4] Dark fields favour the **misallocation** of governmental resources. If the real levels of cybercrime are not revealed, there will not be enough police officers, prosecutors, and judges to combat it effectively. The same applies to the allocation of financial and other resources in the corporate context. Maintaining the dark field paves the way for the future. Therefore, every criminal complaint is a contribution to more IT security, even if it does not lead to a concrete success in the investigation by arresting or convicting a suspect.

3.5 And We Will Get Them After All[5]

The international nature of cybercrime is undoubtedly a major challenge for law enforcement. The cooperation with international authorities and organisations is demanding and—at least in a non-European context—time and resource consuming. In the meantime, however, many countries have come to realize that cybercrime can only be effectively prosecuted in international or multinational investigation groups because of the international nature of its commission. The specialisation of judicial and police services to combat cybercrime has made a significant contribution to this. In the field of transnational data access, too, legislative projects give hope that

[4] Other figures see even lower advertisement rates, about just over 13% for the UK, see https://www.telegraph.co.uk/news/2016/11/01/how-much-of-a-problem-is-cyber-crisis-in-the-uk/. Accessed on February 13, 2019.

[5] The motto of ZAC Talk No. 3 (2019), an information and discussion event organized by the ZAC NRW for large companies, said: "And we will get them after all! – Why law enforcement on the Internet is not futile."

international cooperation will be further strengthened.[6] Notwithstanding the numerous detailed questions to be discussed, the willingness to promote the effectiveness of international cooperation by streamlining and simplifying procedures is to be welcomed. It thus takes effect in practice of **trusting cooperation** within the international law enforcement community which has already established a tendency towards pragmatic cooperation. As a result of these developments, the probability of successful investigations is also increasing in international matters.

At the same time, it is worthwhile to assess the concept of the success of an investigation in the field of cybercrime not only based on the probability of conviction, but also to look at the **systemic contribution of law enforcement to cyber security**. The clarification of the course of events, the formulation of certain signs of successful attacks (so-called *indicators of compromise*), the analysis of existing weak points and the attribution of cyberattacks contribute significantly to protection against future threats.

4 Prosecution As a Service

The judiciary has recognized and accepted the challenge posed by cybercrime. Almost all federal states have set up **specialised offices**—albeit in different organizational forms—that provide the legal and technical expertise to be able to successfully conduct complex investigative procedures with a high degree of technicality. In North Rhine-Westphalia, the **Central and Contact Point Cybercrime (ZAC NRW)**, which is located at the Cologne Public Prosecutor's Office, is a judicial institution that is responsible for the prominent cybercrime proceedings in the entire state. This concentration of responsibilities enables the formation of phenomenon-specific specialist departments, for example for attacks on critical infrastructures or trading of incriminated goods and services on the darknet. The public prosecutors working in the central office, with their personal expertise, are available to companies as an anchor of trust over the long term and enable the forward-looking coordination of a coordinated approach even in advance of individual concern.

This form of "**proactive repression**" avoids the otherwise inevitable loss of efficiency in the cyber crisis and enables comprehensive preservation of evidence even in time-critical investigative situations.

Countering the actions of cyber criminals with an interconnected control strategy is one of the essential requirements of successful digital law enforcement. Although restrictions might occur from the tasks of a law enforcement agency, the concept of a

[6] See Proposals of the European Commission for a regulation of the European Parliament and of the Council on European production orders and preservation orders for electronic evidence in criminal matters and for a directive of the European Parliament and of the Council laying down uniform rules for the appointment of representatives for the purposes of Gathering of evidence in criminal proceedings, from April 17, 2018.

"shared mission"[7] between law enforcement, business and society can do more for cyber security than the competence of each individual sphere.

Industry 4.0 is supported by **Justice 4.0**.

References[8]

Bitkom e. V., study "Spionage, Sabotage und Datendiebstahl – Wirtschaftsschutz in der Industrie" (2018) (Espionage, sabotage and data theft – economy protection in the industry), https://www.bitkom.org/sites/default/files/file/import/181008-Bitkom-Studie-Wirtschaftsschutz-2018-NEU.pdf

Bitkom e. V., study "Wirtschaftsschutz in der digitalen Welt" (2017) (Economy protection in the digital world), https://www.bitkom.org/sites/default/files/pdf/Presse/Anhaenge-an-PIs/2017/07-Juli/Bitkom-Charts-Wirtschaftsschutz-in-der-digitalen-Welt-21-07-2017.pdf

Bitkom Research, Study "Status Quo IT-Sicherheit in Deutschen Unternehmen" (2017) (Status quo IT-security in German business companies)

Bollhöfer E, Jäger A (2018) Wirtschaftsspionage und Konkurrenzausspähung - Vorfälle und Prävention bei KMU im Zeitalter der Digitalisierung (economy espionage and competition espionage – incidents and prevention at KMU in the time of digitization), Available at https://wiskos.de/de/publikationen_vortraege/publikationen.html

FBI Internet Crime Complaint Center, Public Service Announcement I-071218-PSA from 12.07.2018., https://www.ic3.gov/media/2018/180712.aspx

Federal Office for Information, Die Lage der IT-Sicherheit in Deutschland 2018 (the state of IT-security in Germany 2018), https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2018.pdf

FireEye Mandiant M-Trends (2018). https://www.fireeye.com/content/dam/collateral/en/mtrends-2018.pdf

Lavion D (2018) PWC Global Economic Crime and Fraud Survey 2018. https://www.pwc.com/gx/en/forensics/global-economic-crime-and-fraud-survey-2018.pdf

Lewis J (2018) The Economic Impact of Cybercrime - No Slowing Down. McAfee/Center for Strategic and International Studies (CSIS)

Schwarz D (2016) Estimating the Revenue of a Russian DDoS Booter. https://asert.arbornetworks.com/estimating-the-revenue-of-a-russian-ddos-booter/

[7]Former US President Obama stated at the "Cyber Security and Consumer Protection Summit" on February 13th, 2015 that this can only be achieved in a "shared mission", see https://fsi.stanford.edu/node/218828. Accessed on February 13, 2019.

[8](Internet sources, last reviewed 13.02.2019).

Big Data and Artificial Intelligence: Law and Industry 4.0

Thomas Hoeren and Stefan Pinelli

Contents

T. Hoeren (✉)
University of Münster, Institute for Information, Telecommunication and Media Law, Münster, Germany
e-mail: hoeren@uni-muenster.de

S. Pinelli (✉)
Volkswagen AG, Wolfsburg, Germany

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_6

1 Approaching the Term Big Data

Again and again, there are ground-breaking innovations in the way mankind performs work, which make existing approaches obsolete and lead to the creation of entirely new business models. A first innovation of revolutionary dimensions was the invention of the steam engine. After the division of labour in assembly line production (second revolution), which based on the use of electricity, and the automation of work processes using electronics and information technology (third revolution) we are currently on the brink of the next (fourth) revolution: the networking of production facilities and machines (Staffler 2018, p. 269; Bräutigam and Klindt 2015, p. 1137). This networking produces vast amounts of data, but much of this data is initially compiled in an unstructured way. This creates enormous potential for the analysis of these data sets. An effective analysis can no longer be reached with the methods of information technology used to date. At this point, enormous computer capacities and intelligent analysis systems are required. The reasons why Big Data is essential and profitable for a target-oriented analysis of data can be explained with the help of the (famous) Vs: In 2001, *Laney* first described the 3-V model, focusing on volume, velocity, and variety (Laney 2001, p. 1 ff.).

The term **volume** refers to the continuous increase of available information and data. This development was also driven, among other things, by the countless data generated by sensors, for example. At this point, traditional database systems do not offer sufficient capacity to store the large amount of data. In addition, the maintenance costs would be disproportionately high.

Velocity refers to the rapid growth rate of data. This increases the challenge for companies to perform error-free data analysis in a short time to react to certain patterns quickly or faster than the competition. In this context, the expression "real-time analysis" is often used. This is precisely what is required to be able to intervene in current logistics processes and to take rapid and effective countermeasures against occurring issues (e.g., if a truck used for transport gets into an unexpected traffic jam) (Spangenberg et al. 2017, p. 43).

Variety means the variety of data sources and data types. Previous analytical methods were basically supplied with structured data. However, there is no such thing as a consistent and recognizable structure of texts, images, or videos. The information contained in this unstructured data can be of considerable interest to companies. At this point, Big Data analyses create the added value of providing the ability to analyze unstructured data.

The added business value of data is captured by the term (business) **value.** It is (probably) not possible to attribute a certain general economic value to a date. This value is rather formed by supply and demand, as it depends on too many factors and on the parties involved in the possible data trade (Fraunhofer Institute 2016, p. 11). A date in itself can become completely useless for a company, but in combination with certain other data it can be of great value. Furthermore, data of the same type can have different benefits and, consequently, different values depending on who is requesting it.

2 The Importance of Big Data in Industry 4.0

Detached from pictorial language such as "Data is the new oil" (Wandtke 2017, p. 6) or "Data is the new humus" (Markl 2018, p. 30), it can be seen that effective handling of data for industrial companies is of essential importance for the success of a company.

Minimized delivery times in all connections of the value chain, an effective organization of inbound inventories or the determination of the actual order sizes required in purchasing are just a few examples that show the great benefits that companies can derive from Big Data applications (Siestrup and Zebb 2017, p. 62). In the view of the European Commission, data represents a production factor or economic good in its own right (EU Commission 2015, p. 59). Data are key economic drivers and tools of special significance for innovation and global competitiveness (The European Data Strategy 2020, p. 1). Data on the machine itself as well as data generated during the use of the machine can be valuable (Plattform Industrie 4.0 2016b, p. 21). The systematic analysis of such data can sometimes lead to the development of completely new business models. In this respect, a particularly pressing question in the industrial context is to whom machine data are to be assigned and how companies can subsequently "secure" the knowledge gained from the analysis (Plattform Industrie 4.0 2016b, p. 21; Schlinkert 2017, p. 224).

Politicians have been increasingly concerned with the topic of Industry 4.0, for example, the Federal Ministry of Education and Research with its paper on "The New High-Tech Strategy - Innovations for Germany" and the Federal Ministry of Economics and Energy with its "Industry 4.0" platform or the "SME Digital" program. Additionally, the German Federal Ministry for Economic Affairs and Energy has recently hosted the virtual "Conference on Shaping a globally secure industry 4.0 Ecosystem: Enabling international interoperable security policies" to which international experts from industry, ministries and regulatory authorities had been invited. At the European level, the topic was brought up by the EU Commission, among others, which discussed numerous relevant issues in its initiative to digitize European industry.

Especially in the area of production, vast amounts of data are accumulated, which can help to improve the quality of products or reduce development and production time (Bachmann et al. 2014, p. 169). In this context, "machine-to-machine communication" (M2M communication) is increasingly coming into focus (Börnsen and Büllingen 2015, p. 43). It is of vital importance, whether personal or non-personal data (such as pure machine data) are involved in the process. The legal requirements relating to data processing are determined by this very classification, because personal data are subject to data protection law, above all the GDPR (Article 2(1) GDPR).

Big Data is also able to redesign the value chains in Industry 4.0. For example, a vehicle manufacturer of a smart car will not only act as a manufacturer, but with a view to the networked functionalities ("Connected Car") will also remain the customer's contact person and offer services (Henseler-Unger 2017, p. 9). On "car sharing," cooperation between vehicle manufacturers and intermediary platforms can be useful. In any case, it would appear that the time of classic, linear value chains has passed, and new dynamic value chains are emerging (Roland Berger Strategy Consultants 2015, p. 18).

3 Legal Basis

Not surprisingly, there are no legal regulations that match new business models at the time they are created. Contractual regulations are therefore of particular importance (Piltz and Zwerschke 2021, p. 11). In addition, it must be critically questioned to what extent the current legal framework should be adapted to the digitalized realities of life (Plattform Industrie 4.0 2016a; Faust 2016, p. 29 ff.). A first approach to creating an adaptation—particularly to digital platforms and the network defects associated with them—can be seen in the 10th amendment to the German Act against Restraints of Competition (GWB), which entered into force on 19 January 2021. In the following it will be illustrated to what extent the existing legal system covers the facts of Industry 4.0 with a big-data reference.

3.1 Generation and Protection of Data

Machine data can help in the development of new business models. This is another reason why regulations are needed that assign this data and protect companies from access by third parties. First, the already existing copyright, manufacturer, and know-how protection rules should be considered. Other approaches to deriving a comprehensive right to "digital data" range from analogy to property ownership to considerations of adapting individual intangible property regulations, such as those of copyright law (Hornung and Hofmann 2018, p. 15). The protection of data in Industry 4.0 is often dependent on concrete circumstances of collection and processing that only arise from the context (Schlinkert 2017, p. 224).

3.1.1 Assignment According to Property Law Regulations

De lege lata, under the prevailing view, there is no comprehensive absolute right to data itself (for discussion: Zech 2015a, p. 137 ff; Zech 2015b, p. 1151 ff; Specht and Kerber 2018, p. 70 ff; Eichberger 2019, p. 709 ff; Adam 2020, p. 2063 ff.; Hoeren 2019, p. 5 ff; Hoeren 2018, p. 58 ff). An analogy to the comprehensive provision of Section 903 sentence 1 of the German Civil Code is not appropriate due to the lack of comparability of the interests involved (Sattler 2017, p. 43). The approach, which has also been raised by the EU Commission in the meantime, to legally standardize property-type rights in the form of service protection rights with regard to the creation of machine data is clearly rejected by large parts of the literature (Schlinkert 2017, p. 224). Economic evidence of a lack of incentives to collect and produce data has so far not been provided (Specht and Kerber 2018, p. 198 f.). Rather, the flexibility of companies is being jeopardized, which could hamper innovation. The allocation should be contractually regulated.

In this respect, a revision of the law on general terms and conditions with the aim of improving the standardization of contracts is appropriate for the B2B sector. Industrial companies would be sufficiently sensitized with regard to operational data so that self-regulation of the market in the form of data use agreements concluded between companies would be possible (Drexl et al. 2016, p. 915). Unlike "data ownership," which is to be rejected, data possession is certainly a possibility. The lack of material quality of data does not prevent such a classification (Markendorf 2018, p. 410 f.). With possession, a weaker right would be chosen in comparison to ownership, which, for example, with the regulations from Sections 861 ff. German Civil Code, brings with it some useful connection rights (Hoeren 2019, p. 7 f.).

3.1.2 Allocation According to Regulations on Intellectual Property

3.1.2.1 Patent Protection

Section 1(1) of the Patent Act provides that a patent shall be granted for inventions subject to certain conditions, resulting in protection. The term "invention" is defined as "doctrine of planned action which is realizable and repeatable and represents the solution of technical tasks by technical means." (Sattler 2017, p. 30 f.). The classification of machine-generated data as an invention is regularly not considered for two reasons. On the one hand, their qualification as doctrine of planned action is problematic (Sattler 2017, p. 30 f.). On the other hand, it fails because of the requirement of a technical solution. Section 1(3) of the Patent Act explicitly excludes the simple reproduction of information (No. 4) as well as the algorithm as a mathematical method (No. 1) from classification as an invention under patent law. In view of the developments in the course of Industry 4.0, it is discussed whether an algorithm should be protected by patent law under certain circumstances (Federal Court of Justice GRUR 2015, p. 983); however, this protection would only concern the algorithm and would not cover the machine-generated data per se. Utility model law also requires an invention in Section 1(1) GebrMG to trigger industrial property rights, so that no protection for machine-generated data can be derived from this either (Goebel and Engel 2015, Section 1 GebrMG, para. 3).

3.1.2.2 Legal Protection for Computer Software

Legal protection for computer software under Section 69a(1) of the German Act on Copyright and Related Rights presupposes a personal intellectual creation (cf. (3) and Section 2(2) Copyright Act). It is true that machine-generated data are coded information which is present in readable characters and are thus similar to computer programs (Sattler 2017, p. 32 f.). However, unlike computer programs, which can represent an intellectual creation of a linguistic nature (Dreier 2018, Section 69a Copyright Act, para. 1), it is not human performance within the meaning of copyright law (Loewenheim/Leistner 2020, Section 2 Copyright Act, para. 38 f.; Grützmacher 2019, Section 69a Copyright Act, para. 34).

3.1.2.3 Protection as Database Producer

A further possibility of protection under the Act on Copyright and Related Rights is the protection of the performance in the creation of a database under Section 87a (1) since no personal intellectual creation in the sense of copyright law is required in this respect. The decisive factor for the protection of machine-generated data here is that the characteristic of substantial investment is fulfilled. A substantial investment within the meaning of Section 87a(1) Act on Copyright and Related Rights presupposes that this is intended to enable the procurement of data. The investment in the creation and systematic maintenance of a database is thus protected. According to case law, the generation of data is not covered by the meaning and purpose of the provision, so that investments in this regard are therefore irrelevant (Court of Justice of the European Union NJW 2005, p. 1263; Court of Justice of the European Union GRUR 2005, p. 252, p. 254). Sometimes it is hardly possible to clearly distinguish between data procurement and data generation (Federal Court of Justice GRUR 2005, p. 857 f.; Wiebe 2017, p. 340 f.).

3.1.2.4 Protection As a Corporate Secret

Machine-generated data can be protected as a trade or business secret by the regulations of the Law on the Protection of Trade Secrets (GeschGehG), which came into force in April 2019 as an implementation of the Directive (EU) 2016/943 on the Protection of Trade Secrets. For companies, the increasing networking of machines and production facilities opens up new sources of danger for unintentional disclosure of information worthy of protection (Staffler 2018, p. 271; Chirco 2016, p. 12 f.). It becomes particularly opaque if the production facilities are located in different countries that have different requirements in terms of protection of trade secrets (Staffler 2018, p. 271). The networking of different production sites and different production processes creates numerous interfaces at which such data can potentially be accessed that are highly informative for the strategy and working methods of a company (Chirco 2016, p. 12 f.).

Besides the protection in Sections 203, 204 of the German Criminal Code the economic secrets of an enterprise are primarily protected under German law in the relatively new Law on the Protection of Trade Secrets. In Section 2 No. 1, trade secrets are defined as information

(a) that is not generally known or readily accessible, either as a whole or in the precise arrangement and composition of its component parts, to persons in the circles that normally handle that type of information, and is therefore of economic value; and
(b) which is the subject of secrecy measures appropriate under the circumstances by its rightful owner; and
(c) for which there is a legitimate interest in secrecy.

The term *trade secret* can also include data which companies collect by using machines or structured compilations of information such as data pools (Alexander 2021, Section 2 GeschGehG, para. 25; Krüger et al. 2020, p. 580). The information must have economic value because of its non-disclosure (Ohly 2019, p. 443). Machine-generated data can be used commercially for various purposes and thus have an economic value (Hessel and Leffel 2020, p. 649).

Although not explicitly included in the definition, the term *trade secret* reveals that the information must have a business or company connection (Alexander 2021, Section 2 GeschGehG, para. 83). A reference to a company is not possible if the data does not concern the sphere of the company that requires protection, i.e., data that can be attributed to private individuals or other companies. All data that is related to a business activity that has already been or will be conducted thus has a business connection (Alexander 2021, Section 2 GeschGehG, para. 83).

Machine-generated data is considered to be in the public domain and therefore not protected if it is generally known or if it is possible for third parties to gain access to it without significant hurdles (Alexander 2021, Section 2 GeschGehG, para 83). Such easy access is given, for example, if free access to the data is made possible via a website (Federal Court of Justice GRUR 2006, p. 1046—the decision was made based on the German Act against Unfair Competition (UWG), before the Law on the Protection of Trade Secrets came into force). If—especially in value chains—third parties are granted access to the data, it is extremely important to conclude a non-disclosure agreement with them. If such an agreement exists and external persons are bound to secrecy, the criterion of disclosure is not fulfilled by the fact that these persons have access to the data (Federal Court of Justice GRUR 2012, p. 1049; Maaßen 2019, p. 360). Under the new law, a breach of the agreement also constitutes a tortious act under Section 4(2) No. 2 and 3 (Harte-Bavendamm 2020, Section 2 GeschGehG, para. 64).

The data must be the subject of secrecy measures appropriate under the circumstances by its rightful owner. The rightful owner will usually be the owner of the machine and thus the manufacturing company (Hessel and Leffel 2020, p. 650). When protecting machine-generated data, technical protection measures such as data encryption can be considered in addition to contractual non-disclosure agreements as already mentioned (Hessel and Leffel 2020, p. 649). The requirement of a legitimate interest in confidentiality is still legally disputed, as this implementation of the EU Trade Secrets Directive only exists in German law (Alexander 2021, Section 2 GeschGehG, para. 74). Apart from this point worth mentioning, there are no special considerations for the machine-generated data (Hessel and Leffel 2020, p. 650).

Consequently, the qualification of machine-generated data as a trade secret can provide far-reaching protections, as it can give rise to civil claims and criminal sanctions in the event of infringement by third parties (Hessel and Leffel 2020, p. 650).

3.2 Purchase of Data (Data Use Agreement)

When passing on machine-generated data, the parties involved can include the conditions for this process in a general contract (such as a purchase or maintenance contract) or conclude a separate contract on data transfer or data use (Sattler 2017, p. 48). Apart from the restrictions which may exist due to the legal regulations on general terms and conditions, the contracting parties are entitled to determine the contractual provisions autonomously in the light of their private autonomy. Especially on the lack of exclusive rights to data, it is reasonable to make use of this option. The following aspects should be particularly clarified (in detail Sattler 2017, p. 49 ff.):

- Regulation of contractual penalties for breaches of contractual obligations
- Definition of "machine-generated data" and rules concerning the allocation of data in the relationship between the Contracting Parties
- Record the purpose of the data use and on this basis grant exploitation rights
- Agreement on the interfaces of data transmission (including technical processes and duties of cooperation, etc.)
- Regulation concerning the rights to the results of data processing and data use
- Prerequisites for termination of contract and (reverse) processing after termination of contract
- Agreement concerning the protection of data security against internal and external access and abuse
- Regulations on liability (reasons for liability, limitations, or exclusions of liability etc.)

3.3 The Use of Autonomous Systems

3.3.1 Conclusion of Contract

In Industry 4.0 there is an increasing number of contracts which are concluded on one or partly on both sides by using autonomous software agents. These software agents can react independently to changing circumstances, their range of functions therefore goes beyond systems that merely execute predetermined behavioral commands by programming (Groß 2018, p. 5). The question of the extent to which the use of autonomous software agents can be attributed to the respective user when concluding a contract is still unclear if the software agent—who has its own legal personality due to the scope of its functions (Gleß and Weigend 2014, p. 570)—goes too far and is not compatible with the applicable law. The distinction between messenger and deputy, which is also controversial when using human intermediaries, is rather subject of the current scientific discussion (Krebs 2021, preface Section 48 German Commercial Code, para. 81). Some argue that the software agent should be treated as a messenger (Riehm 2014, p. 113 f.). The messenger does not have to have its own legal personality, he merely transmits a foreign declaration of intent (Schubert 2018, Section 164 German Civil Code, para. 71). The software agent sometimes bears considerable responsibility for the content of the contract and can, for example, determine the terms of delivery or even the price (Groß 2018, p. 5). In this respect, the classification as a messenger does not seem appropriate (Specht and Herold 2018, p. 43). According to a contrary view in the literature, it is much more appropriate to put the software agent on an equal footing with a deputy due to the range of functions (Schirmer 2016, p. 664). This is in contradiction to the fact that the role of the deputy requires a separate legal personality, which the software agent does not have based on a possible human will formation (alternative considerations on this: Specht and Herold 2018, p. 43 f.). An analogous application of the regulations for the deputy with limited legal capacity, which was considered for this reason (Cornelius 2002, p. 353; Sorge 2006, p. 25), raises problems through a possible liability as falsus procurator. The software agent does not have its own liability assets after all, so that the liability risk would be transferred to the user's contractual partner, such an outcome cannot be accepted (Groß 2018, p. 5). It seems reasonable to attribute the software agent's declaration to the respective user (Krebs 2021, preface Section 48 German Civil Code, para. 82). By using the software agent, the user creates a state of trust which leads to the fact that the declaration is to be assigned to this user according to the objective recipient horizon (Higher Regional Court of Frankfurt MMR 2003, p. 406; Regional Court of Cologne MMR 2003, p. 482; Groß 2018, p. 5).

3.3.2 Inclusion of General Terms and Conditions (GTC)

In the context of contracts that are concluded by autonomous software agents on both sides, the question arises to what extent general terms and conditions can be included in the contract. In this context, the following must be particularly considered:

3.3.2.1 Establishment of the GTC

The party that initiates the inclusion of the clauses in the contract *provides* the general terms and conditions and thereby acts as the clause user (Stadler 2021, Section 305 German Civil Code, para. 6). The origin of this provision is that consumers are to be protected from the superior negotiating position of entrepreneurs, as they usually have a clear lack of understanding for the provisions formulated in the contract. With regard to a B2B contract, some argue that there is no comparable imbalance and that the characteristic of the introduction of GTC is therefore not suitable in this context for the assessment of inclusion; in addition, in these contracts it is usually determined by chance from which side the GTC are introduced and who is subsequently considered to be the user of the clause (Groß 2018, p. 8). The situation is similar for contracts concluded by autonomously acting software agents. These agents can read in the content of the contract in real time and compare it with the user's previously programmed ideas and requirements (Groß 2018, p. 8).

3.3.2.2 Enabling Information

For GTCs to be effective, Section 305(2) No. 2 German Civil Code requires that the user of the clause offers the other party a reasonable opportunity to take note of the GTCs. On software agents, the question arises as they are configured in such a way that they can act autonomously, and the users are no longer aware of any actions. When the GTC are stored on the user's servers, the clause user can reasonably assume that his negotiating partner is able to inspect them at the latest (Groß 2018, p. 8).

3.3.2.3 Compliance With the Written Form

It follows from Section 305(1) sentence 2 German Civil Code that GTCs must be in written form. During contract negotiations by software agents, however, there will usually be no correspondence between the two parties. In this respect, it is recommended not only to exchange the program code of the software agents, but also to transmit the GTC in text form (Groß 2018, p. 8 f.).

3.3.3 Liability Law

In the context of Industry 4.0 the question of who must assume liability for autonomous systems arises in several constellations. On the one hand, there is the question of liability law with regard to autonomous vehicles for which basically a very detailed liability especially under the German Road Traffic Act (StVG) has been regulated in 2021 (Autonomous Driving Act from 27 July 2021, BGBl. I p. 3108). There are, however, several applications of autonomous systems in which such a special liability cannot be applied. There are also discussions about a strict liability for faulty big-data predictions (Kirchner 2018, p. 20; Bräutigam and Klindt 2015, p. 1139; Horner and Kaulartz 2016, p. 7). Furthermore, there are questions on the liability for erroneous data based on contractual agreements or warranty rights.

3.3.3.1 Tort Law

Liability under tort law, particularly under Sections 823 ff. German Civil Code, usually presupposes fault. Particularly on the behavior that an artificial intelligence acquires by self-learning, the fault of the user is considered to be out of the question.

The only thing that appears to be reproachable is a considerable violation of due diligence obligations in the operation of the AI (Zech 2019, p. 210; Bräutigam and Klindt 2015, p. 1139). Parts of the literature argue in favor of considering a strict liability for the operators of AIs (Bräutigam and Klindt 2015, p. 1139). Such strict liability is intended for cases in which one person opens up a source of risk and, consequently, another person suffers damage caused by the operation that cannot be prevented (Sprau 2021, preface Section 823, para. 11). The risk in the operation of AI lies in the fact that, due to the self-learning process, the mode of operation cannot be completely controlled by the operator. On employment relationships, it is argued that the employer makes use of autonomous systems to maximize profits and that this creates an operational risk which the employee can sometimes only control with difficulty (Groß and Gressel 2016, p. 996).

Additionally, a discussion emerged about product liability for autonomous systems (Steege 2021, p. 6; Seehafer and Kohler 2020, p. 213; Zech 2019, p. 212).

There is also a discussion about strict liability independent of fault for erroneous big-data predictions (Bräutigam and Klindt 2015, p. 1139; Horner and Kaulartz 2016, p. 7). This must be countered by the fact that a probability of error is immanent in the predictions. In the case of erroneous predictions, it will often be the case that the injured party could have been expected to deal with the factors of the analysis or simply not trust the analysis.

3.3.3.2 Contractual Liability for Incorrect Data

Due to the wording of Section 433(1) German Civil Code, for the application of Sections 433 ff. German Civil Code it is necessary that the object of purchase are goods (Hoeren and Völkel 2014, p. 31). Under Section 453(1) Alt. 2 German Civil Code, the regulations on the purchase of goods are also applicable to other objects. This term also includes data, so that the contractual warranty law is applicable to a data sale (Berger 2021, Section 453 German Civil Code, para. 11). Under Sections 453, 433(1) sentence 2 German Civil Code, the seller is obliged to provide this data free of material and legal defects in the case of a data sale. The existence of a defect is determined by the recently introduced equally subjective and objective concept of error (Lorenz 2021, p. 2065 f), i.e., by the agreements of the parties to the contract (Kirchner 2018, p. 21).

If an industrial enterprise 4.0 provides a service provider with the data generated in the production process for Big Data analysis, it must be discussed how incorrect data affects the contractual obligations. In this case, the provision of error-free data represents (only) an obligation or, depending on the contractual agreement, an ancillary service obligation under Section 241(1) German Civil Code (Kirchner

2018, p. 23). However, such a circumstance is suitable to subsequently exclude the liability of the data processing service provider for such cases in which errors in the results—which were made available after the analysis—are based on the supplied incorrect data (Peschel and Rockstroh 2014, p. 576).

3.4 Data As Essential Facility

Data are generally not exclusive, so that in most constellations it is not apparent to what extent competitors can be excluded from downstream markets by a refusal to deal (Bundeskartellamt 2017, p. 10; Nuys 2016, p. 516). However, this is different for machine-generated data. Here, data (sets) are obtained from analytical methods and can be combined into exclusive data sets, which can then constitute an essential facility. Acknowledging this European doctrine, data has been added as one example in Section 19(2) No. 4 Act against Restraints of Competition (GWB) (Paal and Kumkar 2021, p. 814). Furthermore, a rule for access to data was also added in Section 20(1a) Act against Restraints of Competition, which does not require a dominant position of the other party on the market, but only a dependency on the data, controlled by the other party. While both changes give some clarification and recognize the general need for such rules, they do not give detailed guidance as to the exact requirements for access to data (Paal and Kumkar 2021, p. 814). On the instruction to pass on data on this basis, there are, however, considerable data protection concerns in connection with personal data, since this process will not normally satisfy a legality requirement of Article 6(1) of the GDPR. Machine-generated data, on the other hand, is mostly non-personal data, so that it is possible to qualify it as an essential facility under the conditions described above (Schweitzer and Peitz 2017, pp. 81 et seq.).

References

Adam S (2020) Daten als Rechtsobjekte. NJW 2063 ff

Alexander (2021) In: Köhler H, Bornkamm J, Feddersen J (eds) UWG, 39th edn. C.H. Beck, München

Bachmann R et al (2014) Big Data – Fluch oder Segen? mitp, Heidelberg

Berger C (2021) In: Jauernig O, Stürner R (eds) Bürgerliches Gesetzbuch, 18th edn. C.H. Beck, München

Börnsen S, Büllingen F (2015) Marktorganisation und Marktrealität von Machine-to-Machine Kommunikation mit Blick auf Industrie 4.0 und die Vergabe von IPv6-Nummern, WIK-Diskussionsbeitrag Nr. 400

Bräutigam P, Klindt T (2015) Industrie 4.0, das Internet der Dinge und das Recht. NJW 1137 ff

Bundeskartellamt (2017) Big Data und Wettbewerb. https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Schriftenreihe_Digitales/Schriftenreihe_Digitales_1.pdf?__blob=publicationFile&v=3 (last accessed 21 April 2021)

Chirco C (2016) Industrie 4.0 in der Praxis. InTeR, 11 ff

Cornelius K (2002) Vertragsabschluss durch autonome elektronische Agenten. MMR 353 ff
Court of Justice of the European Union (2005) Kein Investitionserfordernis für Fußballspielplan. GRUR 252 ff
Court of Justice of the European Union (2005) Schutz der Daten einer Sportdatenbank gegen Nutzung Dritter. NJW 1263
Dreier T (2018) In: Dreier T, Schulze G (eds) Urheberrechtsgesetz, 6th edn. C.H. Beck, München
Drexl J et al (2016) Ausschließlichkeits- und Zugangsrechte an Daten. GRUR Int. 914 ff
Eichberger M (2019) Rechte an Daten. VersR 709 ff
EU-Kommission (2015) Commission Staff Working Document, A Digital Single Market Strategy for Europe, SWD (2015) 100 final. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52015SC0100&from=EN (last accessed 21 April 2021)
The European Data Strategy (2020) Shaping Europe's digital future. https://ec.europa.eu/commission/presscorner/api/files/attachment/862109/European_data_strategy_en.pdf. Accessed 4 Jan 2022
Faust F (2016) Digitale Wirtschaft – Analoges Recht: Braucht das BGB ein Update? NJW-Beil. 29 ff
Federal Court of Justice (2005) Verletzung des Datenbankherstellerrechts bei Neuzusammenstellung von Daten aus Chartlisten. GRUR 857 ff
Federal Court of Justice (2006) Aneignung von Kundendaten als Verletzung des Geschäftsgeheimnisses. GRUR 1044 ff
Federal Court of Justice (2012) Mitnahme eines Betriebsgeheimnisse enthaltenden Dokuments beim Ausscheiden aus Dienstverhältnis - MOVICOL-Zulassungsantrag. GRUR 1048 ff
Federal Court of Justice (2015) Patentfähigkeit von mathematischen Methoden - Flugzeugzustand. GRUR 983 ff
Fraunhofer-Institut (2016) Industrial Data Space - Digitale Souveränität über Daten. https://www.fraunhofer.de/content/dam/zv/de/Forschungsfelder/industrial-data-space/Industrial-Data-Space_whitepaper.pdf (last accessed 21 April 2021)
Gleß S, Weigend T (2014) Intelligente Agenten und das Strafrecht. ZSTw 561 ff
Goebel F, Engel F-W (2015) In: Benkard G (ed) Patentgesetz, 11th edn. C.H. Beck, München
Groß J (2018) AGB 4.0: Allgemeine Geschäftsbedingungen im Rahmen autonomer Vertragsschlüsse. InTeR 4 ff
Groß N, Gressel J (2016) Entpersonalisierte Arbeitsverhältnisse als rechtliche Herausforderung – Wenn Roboter zu Kollegen und Vorgesetzten werden. NZA 990 ff
Grützmacher M (2019) In: Wandtke A, Bullinger W (eds) Praxiskommentar zum Urheberrecht, 5th edn. C.H. Beck, München
Harte-Bavendamm H (2020) In: Harte-Bavendamm H, Ohly A, Kalbfus B (eds) GeschGehG. C.H. Beck, München
Henseler-Unger I (2017) In: Sassenberg T, Faber T (eds) Rechtshandbuch Industrie 4.0 und Internet of Things. C.H. Beck, München
Hessel S, Leffel L (2020) Rechtlicher Schutz maschinengenerierter Daten. MMR 647 ff
Higher Regional Court of Frankfurt (2003) Übermittlungsfehler des providers. MMR 405 f
Hoeren T (2018) A new approach to data property? AMI 58 ff
Hoeren T (2019) Datenbesitz statt Dateneigentum. MMR 5 ff
Hoeren T, Völkel J (2014) In: Hoeren T (ed) Big Data und Recht. C.H. Beck, München
Horner S, Kaulartz M (2016) Haftung 4.0. CR 7 ff
Hornung G, Hofmann K (2018) In: Hornung G (ed) Rechtsfragen der Industrie 4.0. Nomos, Baden-Baden
Kirchner G (2018) Big Data Management: Die Haftung des Big Data-Anwenders für Datenfehler (Vertragsrecht – Teil 1). InTeR 19 ff
Krebs P (2021) In: Münchener Kommentar, Schmidt K, Ebke W (eds) HGB, vol 1, 5th edn. C.H. Beck, München
Krüger S, Wiencke J, Koch A (2020) Der Datenpool als Geschäftsgeheimnis. GRUR 578 ff

Laney D (2001) 3D data management: controlling data volume, velocity, and variety. https://studylib.net/doc/8647594/3d-data-management%2D%2Dcontrolling-data-volume%2D%2Dvelocity%2D%2Dan (last accessed 21 April 2021)
Leistner M (2020) In: Loewenheim U, Leistner M, Ohly A (eds) Urheberrecht, 6th edn. C.H. Beck, München
Lorenz S (2021) Die Umsetzung der EU-Warenkaufrichtlinie in deutsches Recht. NJW 2065 ff
Maaßen S (2019) "Angemessene Geheimhaltungsmaßnahmen" für Geschäftsgeheimnisse. GRUR 352 ff
Markendorf M (2018) Recht an Daten in der deutschen Rechtsordnung. ZD 409 ff
Markl V (2018) Daten sind der neue Humus. In Big Data, Smart Data, next? pp 28–31. https://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/Smart-Data/2018_08_19_Big_Data_Smart_Data_next.pdf?__blob=publicationFile&v=5 (last accessed 21 April 2021)
Nuys M (2016) "Big Data" - Die Bedeutung von Daten im Kartellrecht. WuW 512 ff
Ohly A (2019) Das neue Geschäftsgeheimnisgesetz im Überblick. GRUR 441 ff
Paal B, Kumkar P (2021) Wettbewerbsschutz in der Digitalwirtschaft. NJW 809
Peschel C, Rockstroh S (2014) Big Data in der Industrie – Chancen und Risiken neuer datenbasierter Dienste. MMR 571 ff
Piltz C, Zwerschke J (2021) "Rechte an Daten": Neuere Entwicklungen und Haftungsfragen. GRUR-Prax 11 ff
Plattform Industrie 4.0 (2016a) Ergebnispapier Digitalisierte Industrie – Analoges Recht? https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/digitalisierte-industrie-analoges-recht.pdf?__blob=publicationFile&v=8 (last accessed 21 April 2021)
Plattform Industrie 4.0 (2016b) Ergebnispapier Industrie 4.0 – wie das Recht Schritt hält. https://www.bmwi.de/Redaktion/DE/Publikationen/Industrie/industrie-4-0-wie-das-recht-schritt-haelt.pdf?__blob=publicationFile&v=6 (last accessed 21 April 2021)
Regional Court of Cologne (2003) Willenserklärung durch "Auto-Reply". MMR 481 f
Riehm T (2014) Von Drohnen, Google-Cars und Software-Agenten. ITRB 113 ff
Roland Berger Strategy Consultants (2015) Die digitale Transformation der Industrie. https://bdi.eu/media/presse/publikationen/information-und-telekommunikation/Digitale_Transformation.pdf (last accessed 21 April 2021)
Sattler A (2017) In: Sassenberg T, Faber T (eds) Rechtshandbuch Industrie 4.0 und Internet of Things. C.H. Beck, München
Schirmer J-E (2016) Rechtsfähige Roboter? JZ 660 ff
Schlinkert H-J (2017) Industrie 4.0 – wie das Recht Schritt hält. ZRP 222 ff
Schubert C (2018) In: Münchener Kommentar, Säcker F et al (eds) BGB, vol 1, 8th edn. C.H. Beck, München
Schweitzer H, Peitz M (2017) Datenmärkte in der digitalisierten Wirtschaft: Funktionsdefizite und Regelungsbedarf? http://ftp.zew.de/pub/zew-docs/dp/dp17043.pdf (last accessed 21 April 2021)
Seehafer A, Kohler J (2020) Künstliche Intelligenz: Updates für das Produkthaftungsrecht? EuZW 213
Siestrup G, Zebb D (2017) Reifegradbestimmung: der Weg zur Supply Chain 4.0. Industrie 4.0 Management 59 ff
Sorge C (2006) Softwareagenten. Vertragsschluss, Vertragsstrafe, Reugeld. https://doi.org/10.5445/KSP/1000003768 (last accessed 21 April 2021)
Spangenberg N et al (2017) Big Data in der Logistik. Industrie 4.0 Management 43 ff
Specht L, Herold S (2018) Roboter als Vertragspartner? MMR 40 ff
Specht L, Kerber W (2018) Datenrechte – Eine rechts- und sozialwissenschaftliche Analyse im Vergleich Deutschland – USA. http://www.abida.de/sites/default/files/ABIDA_Gutachten_Datenrechte.pdf (last accessed 21 April 2021)
Sprau H (2021) In: Palandt O (ed) Bürgerliches Gesetzbuch, 80th edn. C.H. Beck, München
Stadler A (2021) In: Jauernig O, Stürner R (eds) Bürgerliches Gesetzbuch, 18th edn. C.H. Beck, München

Staffler L (2018) Industrie 4.0 und wirtschaftlicher Geheimnisschutz. NZWiSt 269 ff
Steege H (2021) Auswirkungen von künstlicher Intelligenz auf die Produzentenhaftung in Verkehr und Mobilität NZV 6 ff
Wandtke A-A (2017) Ökonomischer Wert von persönlichen Daten. MMR 6 ff
Wiebe A (2017) Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken. GRUR 338 ff
Zech H (2015a) Daten als Wirtschaftsgut – Überlegungen zu einem "Recht des Datenerzeugers". CR 137 ff
Zech H (2015b) "Industrie 4.0" – Rechtsrahmen für eine Datenwirtschaft im digitalen Binnenmarkt. GRUR 1151 ff
Zech H (2019) Künstliche Intelligenz und Haftungsfragen. ZfPW 198 ff

Media Law 4.0

Frank Fechner and Johannes Arnhold

Contents

F. Fechner · J. Arnhold (✉)
Ilmenau University of Technology, Ilmenau, Germany
e-mail: frank.fechner@tu-ilmenau.de; johannes.arnhold@tu-ilmenau.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_7

1 Introduction

'Media Law 4.0' as a legal concept does not exist. Nevertheless, when it comes to media law, this label can consider the way in which the media has changed over the last few years—moving away from the conventional 'classic media' distinct from each other towards a conglomerate of the most varied media services. This 'convergence of the media' results from the technical possibility of digitalising all media contents and exchanging and disseminating them without any loss of quality. Given the legal circumstances actually prevailing, it seems hardly in keeping with the times to distinguish, in an analysis of 'Media 4.0', between different forms of media, as is largely done by national and European legal provisions; instead, it appears useful to classify according to the persons or players working in the media. This attempt will be made below whereby the fact that a player may be active in various fields of activity should not be overlooked.

If one attempts to describe a 'Media Law 4.0', it is less the change in the law itself than the further development of the world of media which becomes the focal point of the analysis. The law responds to technical and societal developments, so that it would be more appropriate to speak of **'Media law in the age of Media 4.0'**.

The first noticeable change is that in the **forms of media**. Newspapers are increasingly supplemented or completely replaced by audio-visual media, and radio and television by online media. Conventional mass media pale beside mixed forms of communication; instead of media forms covering several topics, search engines and language assistants convey targeted but possibly also biased information; social media with interactive functions outperform the mass media aimed at the recipients.

A second change manifests itself in the blurring of the distinction between **media creators and recipients**. Whereas the law distinguishes between linear and non-linear dissemination and regulates linear services to a greater extent, the distinction is of almost no importance to the user. This results in radical changes, such as the possibility of having greater direct influence on political decision-making processes. In the positive sense, this results in 'more democracy', a participative effect which is of particular importance in the representative democracy. In the negative sense, these possibilities are frequently associated with brutal language, viral dissemination of contents which breach privacy, incite hatred or are deliberately false, as far as fake news.

In the process of change are, thirdly, **media contents** and **forms of use**. Instead of the written and spoken word, ever more hybrid forms of text, image, audio and video files are emerging. Media contents are increasingly used on mobile devices such as smartphones and tablets (Schütz 2018, p. 36), which affect the scope of the media contents. Short texts and contributions in the form of film snippets take the place of detailed presentations. Moreover, it is increasingly difficult to discern between editorial and advertising contents. This is especially true of the sector of influencers but also on traditional journalist contents within the framework of newly emerged business models such as content marketing or native advertising. Additionally, many contents are published on the Internet and remain retrievable over a longer time: 'The Internet never forgets'.

The changed use of the media gives rise, fourthly, to a transformed media context with new **forms of financing.** Alongside the conventional financing of the press primarily by subscriptions or advertisements and the financing of the dual broadcasting system by licence fees or advertising with its distinction between public-law broadcasting with its state-guaranteed financing and private broadcasting financed by advertising, other forms of financing have emerged in the online sector such as diverse models of paid content, personalised advertising and crowdfunding.

New technical possibilities and new business ideas have given rise to hitherto unknown **competitive situations** between the players. Media competition is the result, inter alia, of cross-mediality in the national sector. An example are the competitive relations between original broadcasting organisations, in particular organisations under public law, and private press publishers on the Internet, as was discussed in connection with the **Tagesschau-App** (TV news broadcasting app) (Cologne Higher Regional Court *(Oberlandesgericht)*, judgment of 30 April 2015 – I ZR 13/14, Wimmer and Nawrath 2016, p. 132 f.).

In fifth place is the internationalisation or **globalisation** of the media market. This is shown explicitly by international players increasingly pushing their way onto national media markets with new and largely unregulated business models (e.g. Netflix, Amazon-Prime or Spotify as providers of streaming services). They compete with the traditional classic broadcasting providers, which are strongly regulated in national law by comparison. There is also competition with regard to contents which are licensed and, therefore, associated with costs. An economically significant example are 'sports broadcasting rights', e.g. to the Champions League for football, where, owing to highest bids by US or British media groups (Perform Group, Sky), provision at national level can no longer be financed or cannot apparently be reconciled with the functional remit—neither for public-law nor private broadcasters.

Competition between the media affects not only content in the battle for the attention of users and, therefore, the quality or quantity of the media contents; the 'speed' at which media contents are conveyed has also considerably increased. The possibility of directly 'going online' correlates with the risk of sacrificing careful journalistic research to be 'first'. Moreover, attention-grabbing topics in the political mainstream are frequently preferred and campaigns which rapidly develop their own momentum after being fired up in the media and may also lead to breaches of privacy (Ulbrich and Frey 2017, p. 33), as was observed in the cases of Kachelmann (Gounalakis 2016, p. 738; Hofmann and Fries 2017, p. 2371) and '#metoo'

(Mafi-Gudarzi 2018, p. 521; Rodenbeck 2018, p. 1227). In the extreme case, reports are invented as done by the *Der Spiegel* journalist Relotius. Fastest and most direct is a livestream from an event location by a random person at the scene. Even if this is the most authentic form of conveying media contents, showing a section of reality without explaining the overall circumstances can frequently give recipients a false—or at least biased—impression.

One of the currently most exciting questions of media law arising from these developments is to what extent the **state** may or must steer or even suppress such developments. How can the state manage to find a system that does not restrict economic and communicative freedoms on the Internet and that simultaneously guarantees sufficient protection for both content providers and their recipients? To be considered are differing aspects and interests such as the protection of young persons, data protection, consumer protection, possibly also protection against hate speech and fake news. Can the state separate the truth from lies? Does it have to protect the population against lies or would it be encroaching on the freedom of formation of opinion by doing so? Should anonymity be guaranteed on the Internet or should it rather be prevented? Does the state have to ensure a balance between media contents by obliging 'gatekeepers' to convey certain contents regardless of their economic relevance? How can the state ensure fair competition between players in a media world 4.0? Should standards be lowered in favour of the more strongly regulated sectors such as broadcasting or, the opposite, more pressure be exercised on intermediaries who strongly influence the markets? Given the increasing market power of intermediaries such as Google or Facebook, the question is whether the intermediaries should be subject to greater regulation or, vice-versa, if less regulation would benefit broadcasting (Beaujean 2018, p. 3).

The approach in the pages that follow is to look at individual **players in the media sector** and examine in what way the law affects their positions and will possibly change the same with regard to the challenges of 'Media law in the Media 4.0 Age'.

A first step involves the legal and actual classification of media creators as journalists and whether media freedoms or journalistic privileges also apply to non-professional players such as 'amateur journalists'. A second step involves analysis of the media enterprises. How are they financed; how can the contents they provide be protected? How are competitive relations between the media providers to be structured? For example, are streaming services to be treated as broadcasts and are public-law broadcasters to be permitted to offer press-like services? And, finally, are the regulatory provisions from the analogue sector still appropriate to the competitive situation between the 'classic media' and Internet providers and intermediaries? Ultimately, during a final stage, the rights of recipients are to be shown and, above all, of those affected by media reports, because their lives are reported on or photos or videos of them are published.

2 Legal Principles of a Media Law 4.0

2.1 Conventional Regulation of the Media

Media law is, and will remain, a **cross-sectional subject matter;** this means that the most diverse range of statutes may apply to media-law issues (Fechner 2021, p. 3). The existing provisions and statutes in the media sector continue to apply in the age of digitalisation. Despite the need for reform, existing statutes are also a good basis for the Media 4.0 Age; not least, fundamental rights have proven themselves through all changes.

2.2 Framework Conditions

The regulatory matter of media law will remain divided in the Basic Law in the future between the federal government and the federal states. Whereas the federal government can claim **legislative jurisdiction** with regard to the law governing the economy, the protection of young persons, telecommunications and other special subjects, it is the federal states that primarily have jurisdiction for media contents (fall-back jurisdiction of the federal states Art. 70 (1) Basic Law *(Grundgesetz, 'GG')*. To be mentioned, first, are a few special statutes which have sector-specific application, e.g. the Telecommunications Act *(Telekommunikationsgesetz, 'TKG')* to transmission processes and the Telemedia Act *(Telemediengesetz, 'TMG')* to the contents of telecommunications. In contrast to these statutes passed by the federal government, there are important media-law regulations of the federal states such as the federal-state press statutes *(Landes-Pressegesetze)* or the State Treaty on Media *(Medienstaatsvertrag)*.

Additionally applicable are **general statutes** which apply in full to the media at least within the German sovereign territory. A person who posts criminally relevant media content on the Internet may have committed an offence; copyrights must also be observed on the Internet, equally youth-protection and data-protection regulations and the right to privacy. The regulations of the federal states may not conflict with the statutory provisions of the federal government, and both must be kept within the framework of the Basic Law. They must ensure the freedom of the media, the freedom of opinion and information (Art. 5 (1) Basic Law) and the freedom of the arts (Art. 5 (3) Basic Law). Finally, the federal legislature is also bound to fundamental rights and European law provisions.

2.3 State Treaty on Media

The legislature is aware of the convergence of the media as is also shown by the introduction of a **State Treaty on Media** *(Medienstaatsvertrag, 'MStV')*. A comparison of the provisions of the former State Treaty on Broadcasting with those of the State Treaty on Media indicates that, although the State Treaty in Broadcasting is supplemented and updated, there can be no talk of any comprehensive regulation of 'the media'. The changes tend to lie in the details.

The State Treaty on Media, which supersedes the former State Treaty on Broadcasting, has as its major innovation introduced primarily provisions for platforms, user interfaces and intermediaries and, therefore, also includes non-linear, on-demand media in the broadcasting law provisions, because they too as 'gatekeepers' can influence the formation of opinion among the population by selecting what they offer. Other media in addition to broadcasting are to be subject to regulation, i.e. on non-discriminatory access of users to media and transparency with regard to the selection of offerings. This becomes clear, for example, when search engine operators who, on the one hand, undertake ranking of offerings with regard to the needs of the user may, however, also give preference to offerings for which they have received payment.

But even the State Treaty on Media is not a uniform regulation in the sense of a 'Media Law 4.0'. In fact, there will be no 'Media Law 4.0 Act' worthy of the name in the near future.

3 Media Creators

3.1 Journalists

There is a traditional understanding that the contents of the mass media are researched and published by journalists. Although journalism as a profession is not subject to any mandatory vocational training or legal access control, the legislature nevertheless assumes the traditional idea of a 'serious' journalist acting responsibly and guided by ethics and the truth. The special features of 'Media Law 4.0' have hitherto been largely ignored by the legislature. Thus, it is not easy for the journalist to find answers to the pressing questions of the digital age. Not just the photojournalists are uncertain with regard to the taking and use of photographic material in the face of the General Data Protection Regulation (GDPR); editors are also unsure whether they may publish contents of social networks in their medium.

Given the great importance to democracy of having media that is free and independent of the state, journalists are given considerable **privileges** under constitutional law and ordinary law. Under constitutional law, they are protected by press and broadcasting freedom which can also be summarised as 'media freedom' and which has found its expression in the press statutes of the federal states or in the State

Treaty on Broadcasting (now incorporated into the State Treaty on Media). Most important are editorial secrecy and the protection of informers which are intended to prevent the influence of the state on the process of creation of media contents. Added to these are the prohibition of pre-censorship and extensive provisions giving protection against searches and confiscation, also partly incorporated in other statutes such as the German Code of Criminal Procedure *(Strafprozessordnung, 'StPO')*. Special features for journalists also arise with access to events such as major trials (Section 6 (2) German Public Meetings Act *(Versammlungsgesetz, 'VersG')*; see also the Federal Constitutional Court decision *(Bundesverfassungsgericht, 'BVerfG')* of 12.4.2013 - 1 BvR 990/13). Complementary provisions include those such as the journalists' right to refuse to give evidence in court so that the protection of informers cannot be undermined by duties to testify as witnesses in court proceedings.

On the other hand, the legislature also makes certain **requirements of the work of journalists**. The 'duty to tell the truth' written into some federal states press statutes makes this requirement particularly clear. Since, however, the 'truthfulness' of media contents can hardly be decided by a state entity, the journalistic duty of care concept used in other press statutes is almost certainly more appropriate. This requires the journalist to verify news prior to its dissemination 'with the most extreme care required by the circumstances regarding content, source and objective accuracy' (thus, e.g. Section 5 Thuringia Press Act *(Thüringer Pressegesetz, 'TPG')*). Also to be considered in this context is the fact that journalists normally have to work under great pressure, and they do not always have the facilities available to verify truthfulness such as, e.g. those of the court. Nevertheless, this is not a duty that could be mandatorily enforced by the state or even just be sanctioned by a punishment or as an administrative offence. That in itself would represent excessive state interference in the freedom of journalistic work.

The privileges and duties of press law apply via the State Treaty on Media not only to radio/TV journalists but also to journalists who work on the Internet. Thus, Section 54 (2) MStV requires that 'telemedia with offerings of a journalistic-editorial nature in which, in particular, complete or partial contents of periodic printed products are reproduced in text or image' satisfy the recognised journalistic principles and that the provider must verify news prior to dissemination with the due diligence appropriate to the circumstances with regard to their content, source and truthfulness. Consequently, they have the same information requirements as radio/TV journalists (Section 19 (4) MStV). Nevertheless, this privilege only applies to such journalistic-editorial offerings 'in which, in particular, complete or partial contents of periodic printed products are reproduced in text or image'. Consideration is given here to the convergence of the media obviously only to a limited extent. It also remains uncertain when exactly a content creator can invoke the privileges of journalists and whether this demarcation is compatible with constitutional law, which is higher-ranking.

3.2 Bloggers and Other 'Amateur Journalists'

Journalists working full-time for a mass medium are faced with ever more competition from **bloggers** and other 'amateur journalists' who have no journalistic training whatsoever and who do not belong to a classic media undertaking, but nevertheless produce media contents and, in case of doubt, are able to report far more quickly from the scene of the event. This raises the question of whether the privileges of journalists can, or must, also be applied to 'amateur journalists'.

Whereas it is occasionally argued that only the professional journalist can invoke such privileges, those who hold this opinion must be asked where they wish to draw the line and whether freedom of the media does not oppose such an exclusion. Can a blogger with a large group of recipients not have greater influence on opinion than a trainee journalist of a local newspaper? How many journalists are required to establish an 'editorial team' that can invoke the right not to reveal sources? What significance can a press credential have for this purpose? The dispute between jurists continues into constitutional law. Can an online editorial team invoke freedom of opinion only or also—like the editorial team of a printed newspaper—press freedom which guarantees additional institutional freedoms such as the protection of sources or the right to refuse to testify? It is quite obvious that statutory law is not yet geared in every respect to the 'Media 4.0' context. Moreover, neither the courts nor the legal advisers agree on the interpretation of the existing laws.

3.3 Influencers

Another phenomenon that can be attributed to 'Media Law 4.0' are 'influencers'. Within a short period of time, they have acquired great influence on the advertising market (Willems 2018, p. 707). Described as **influencers** in the media are players, normally in social networks, on YouTube and Instagram especially, who report from their everyday life and, therefore, give the impression of being impartial and independent. In fact, they **advertise** by mentioning products or services, showing or using them, or pitching their advantages in their videos or postings. Influencers normally receive consideration from their advertising partners which can take the form of remuneration or the supply free of charge of a product or service to be advertised. Of particular interest to companies are those influencers who have large numbers of subscribers or followers and the associated coverage and a high degree of authenticity. They enable the advertising partner to undertake marketing tailored to the specific target group. In this business model, the requirement that the media keep advertising and editorial content separate runs the risk of not being adequately observed.

Nevertheless, influencer marketing does not operate within a legal vacuum (Leeb and Maisch 2019, p. 40). Legislative action at EU level has already taken place within the framework of the amendment of the Audiovisual Media Services

Directive (AVMSD). In particular, the so-called video-sharing platform services (as defined in Art. 1 (1), (aa) of the new version of AVMSD), including YouTube and other social networks, are to be subjected to greater regulation. The focus is primarily on the protection of minors against violence, hate, terrorism and other harmful contents.

The question that arises primarily in the national sector is whether, and normally also how, advertising contents are to be **identified** in the social-media sector. Recourse is to be made to the applicable regulatory framework. For telemedia, Section 6, No. 1 of the Telemedia Act (TMG) and Section 30 of the State Treaty on Media specify that advertising in telemedia must be clearly and unequivocally recognisable as such and must be separate from other content (so-called separation requirement *(Trennungsgebot)*, Heins 2018, p. 795). Failure to identify advertising or commercial contents also represents an unfair act under competition law, i.e. under Section 5a (6) Act against Unfair Competition (*Gesetz gegen den unlauteren Wettbewerb, 'UWG').* At the same time, this raises problems because of the wide range of different functionalities offered by the relevant platforms, which may be subject to rapid transformation.

Until recently there was almost no clarification of questions such as: Is a direct link in the accompanying text to the account of a contractual partner on Instagram to be evaluated differently for advertising law purposes from a mere mention with hashtag?[1] What is the case with other forms of tagging? Moreover, account must be taken of the selected form (image, video, text) of posting and the consideration received by the advertising partner or of the intention or main emphasis of posting. This requires further clarification by case law. There is and was an equal lack of clarity until now about the specific form or minimum degree of recognisability as part of a possible duty of identification, i.e. where exactly is any identification to be inserted on the relevant platform and made clear by what specific wording. (Celle Higher Regional Court *(Oberlandesgericht, 'OLG')*, judgment of 8.6.2017 – 13 U 53/17; Berlin Court of Appeal *(Kammergericht, 'KG'),* decision of 17.10.2017 – 5 W 233/17; Hagen Regional Court *(Landgericht, 'LG')* Hagen, judgment of 13.9.2017 – 23 O 30/17). In the meantime there were a few first decisions of Federal Court of Justice (Bundesgerichtshof, 'BGH') in 2021 that gave some answers to some of these questions, especially in terms of so-called 'tap tags' (Federal Court of Justice (Bundesgerichtshof, 'BGH'), decision of 09.09.2021 – I ZR 90/20 ('Influencer I') and decision of 09.09.2021 I ZR 125/20 ('Influencer II') with the result that there is no obligation to label a tap tag without consideration. At least there will be a change in the law against unfair competition (Gesetz gegen den unlauteren Wettbewerb, 'UWG') under Section 5a (4).

[1] Thereon hitherto Berlin Regional Court *(LG Berlin)*, MMR 2018, p. 543 ff. or Berlin Court of Appeal *(KG Berlin)*, judgment of 08.01.2019 (case ref. 5 U 83/18), which stands back from any general duty of identification and relies mainly on the sales promotion at the focal point of the posting to be evaluated.

3.4 Bots

A development that is difficult to categorise politically, socially and for media-law purposes is the generation of media contents by autonomous machines. In the social-media sector, these are **social bots** which can independently generate answers to questions or contribute to chats not only in social networks such as Twitter or Facebook but also in commentary columns of press platforms. Owing to its abstract control by algorithms, a bot has many different ways of 'behaving' in the course of specific communications. The recipient is not normally able to recognise bots as such because, as partially autonomous computer programs, they feign a human identity by communicating like humans on the Internet. The recipient of the communication remains unaware of any intention to manipulate and pretence of being a person (Steinbach 2017, p. 102).

Bots are used in a variety of very different ways. Whereas, when inquiring about the ingredients of a product, hardly any user would expect a personal reply, there may be a problem if social bots are used for manipulative political purposes, for example, by having machine-generated statements give the impression of originating from real persons infiltrating a discussion on political issues.

The use of social bots has not hitherto been regulated by statute. Under applicable law, the use of social bots is unlikely to be punishable (Volkmann 2018, p. 63). It concerns neither a data alteration under Section 303a (1) German Criminal Code *(Strafgesetzbuch, 'StGB')* nor computer sabotage under Section 303b (1) or (2) StGB. Also, the offence of disruption of the election or election fraud under Section 107 ff. StGB is unlikely to apply.

The prohibition of bots is being discussed, or at least a duty of identification. Specifically called for was an administrative offence for failure to comply with the duty of identification of opinion bots or a duty to use a real name as the result of 'protection of the free nature of public discourse' derived from Article 5 (1) Basic Law (Federal Constitutional Court decision of 4.11.2009 - 1 BvR 2150/08), in which case it may be asked whether this is compatible with the 'right to anonymity' guaranteed in Section 13 (6) TMG (Milker 2017, p. 216). Finally a **duty of identification** has been established in Section 18 (3) MStV.

As well as the idea of creating a criminal offence of 'digital trespass', there are also specific approaches to the regulation of bots. Apart from the technical problems, attempts to define bots in law have hitherto failed, which is due on the one hand to the diverse nature of the media involved; on the other hand, protection of this form of opinion dissemination by the programmer is likely to be affirmed as being within freedom of speech.

As also in a different context, this raises the question of whether the state can or should protect the population against 'fake' media content, because there is a great risk of misuse of such possibilities of influence.

3.5 *'Robot Journalists'*

It is necessary to distinguish between purely supportive acts in the production of media contents, for example by software, and the automatic generation of texts without the involvement of an editor. Until now, this facility is commonly used to announce sports results which are incorporated in a machine-generated text.

The requirements of the federal states' press statutes cannot be met if robot journalists are used; here especially, there can be no journalist's duty of care and there is no weighing of the public interest in freely available information against the privacy rights of the persons affected (Weberling 2018, p. 737). Nevertheless, this does not release editorial teams from liability when using such aids. Computer-based contents are not protected by German copyright law because they lack a personal intellectual achievement. Nevertheless, press publishers can invoke an ancillary copyright (Sections 87f to 87h German Act on Copyright and Related Rights *(Urheberrechtsgesetz, 'UrhG').* This conclusion based on ordinary law is the same as the conclusion in constitutional law. Media freedom takes effect for natural persons only and not for robots.

3.6 *Media Contents Based on Artificial Intelligence (AI)*

In future, it will be necessary to clarify the very fundamental question of whether artificial intelligence (AI), which is based on self-learning systems, requires state regulation. The underlying question is whether there is a need to provide for a statutorily regulated '**right to a human decision**' (von Graevenitz 2018, p. 241).

There is no single definition of AI consistently used by all players. Attempts at regulation to date refer to phenomena of 'weak' AI. In contrast to 'strong' AI, where the systems have the same intellectual skills as a human being or may even surpass the latter, 'weak' is based on the solution to specific application problems whereby the systems developed are capable of self-optimisation (cf. AI strategy of the German government). One special area of application involves attempts to make AI applicable to legal questions ('Legal Tech').[2]

Numerous legal issues arise in this case which can only be intimated at this point. It is to be clarified in very fundamental manner how the development potentials for economic activities can be promoted without impairing the rights of the persons affected. In particular, there is a call for the transparency, comprehensibility and verifiability of the AI systems to be guaranteed to allow effective protection against

[2]For further reference: Hartung/Bues/Halbleib: Legal Tech. Die Digitalisierung des Rechtsmarkts, 2017; Breidenbach/ Glatz Rechtshandbuch Legal Tech, 2018; Herberger: 'Künstliche Intelligenz' und Recht, in: Neue Juristische Wochenschrift (NJW) 2018, p. 2825 ff.; Hähnchen/Bommel: Legal Tech. Perspektiven der Digitalisierung des Rechtsdienstleistungsmarktes, in: Anwaltsblatt (AnwBl) 2018, p. 600 ff.

distortion, discrimination, manipulation or other abusive uses especially when algorithm-based forecasting and decision-making systems are deployed. Also to be clarified is the relationship between Big Data and the individual's right to self-determination of information.

3.7 State and Office-Bearers

A central issue of media law in the technical development of a media world 4.0 is the extent to which state entities and office-bearers are permitted to generate media content. The spectrum ranges from state warnings, for example, about health risks, as far as successful accounts of office-bearers in social networks.

Media content must generally **be kept free of state influence.** Only in this way can the population receive plural contents and independently form an opinion, which is absolutely vital for a liberal democracy. This principle does not apply absolutely as broadcasting law merely requires independence of the state which is related to the composition of the broadcasting boards *(Rundfunkräte),* and which was more narrowly defined by the Federal Constitutional Court (FCC) in 2014 ('*ZDF-Verwaltungsrat*', judgment of 25.3.2014 – 1 BvF 1/11, 1 BvF 4/11). Without doubt, 'state television' would be incompatible with the Basic Law as was already established in the FCC's first broadcasting decision (*'Deutschland-Fernsehen GmbH'*, BVerfG judgment of 28.2.1961 - 2 BvG 1 and 2/60). Whereas it was long disputed how far the **duty of impartiality of politicians** extended in social networks, the FCC clearly defined the boundaries in 2018. State information policy with an objectively informative or warning nature is, accordingly, also admissible in social networks. Nevertheless, members of the government, with a view to the equality of opportunity of political parties, should not use public media to support government parties or attack opposition parties. In particular, government members may not invoke the authority of a government office and thereby exploit associated resources which are not available to the political competitors (*'Wanka'*, FCC judgment of 27.2.2018 - 2 BvE 1/16). They may equally not block unwanted followers from their accounts (i.e. blacklisting).

4 Media Enterprises and Public-Law Media Providers

4.1 Institutional Framework of Media Creation

Traditionally at least, media creators required an institutional framework to be able to create media content. If this need has since become less pronounced in the age of digitalisation, the majority of media creators are nevertheless employed by press publishers, TV/radio broadcasters or by Internet companies and agencies. Press and broadcasting especially, even in the age of 'Media Law 4.0', are still characterised to a great extent by **conventional media structures** and, as things look, this will remain so in the near future.

4.2 Public-Law Broadcasting Organisations

4.2.1 State Guarantee of Diversity of Opinion

The main task of media law is, and will remain so, to guarantee **media diversity**. Only if diversity of opinion is guaranteed in the media, can the citizen make an informed decision during elections especially. Consequently, media that are diverse and independent of the state are a keystone of democratic development of informed opinion. The state can assure diversity of opinion by guaranteeing a large number of media providers who are independent of each other (**external pluralism**), as is the case in the press sector, or, if structures such as broadcasting do not allow the same, provide for **internal pluralism**, as is achieved by means of the broadcasting board in the public-law broadcasting organisations. Since it was formerly not possible, owing to the scarcity of transmission frequencies and the high costs incurred for broadcasting, to guarantee a plurality of providers, diversity of opinion was assured by the broadcasting board in the public-law broadcasting organisations by composing this board from members of highly diverse social groups. At the same time, the Federal Constitutional Court assigned public-law broadcasters the task of guaranteeing the basic supply to the population in that they should provide the latter not only with political information but should also be active over the whole spectrum of the classic broadcasting remit up to and including the broadcasting of entertainment shows. As far as the basic supply remit extends, which the court later also called 'functional remit', the state must make available the financial resources required for the same (*'Rundfunkgebühren II'*, FCC judgment of 11.9.2007 - 1 BvR 2270/05, 1 BvR 809/06 and 1 BvR 830/06). The co-existence of public-law broadcasting organisations and private broadcasters is referred to as a '**dual** broadcasting system'. Since the state may not exercise any influence on the programme contents, financing cannot be made out of tax revenues; instead, charges made must be levied directly on the users, for which reason the broadcasting fee exists. The latest decision of the Federal Court of Constitution is the judgement of 20. 7. 2021 - 1 BvR 2756/20, which the Federal Court of Constitution strengthened the role of public-law broadcasters and judged the increase in the broadcasting fee to be constitutional.

4.2.2 'Broadcasting Law 4.0'?

The media context has changed to a considerable extent since the broadcasting case law of the Federal Constitutional Court began more than 60 years ago. Above all, private broadcasting has established itself since then and offers a wide range of programmes so that one can talk, without more, of an external pluralism that is not dissimilar to that of the press context. During the discussion of the contents of 'Media Law 4.0', it can quite rightly be asked whether there is still justification for the considerable privileges of the public-law broadcasting organisations. This applies particularly to financing by means of mandatorily levied broadcasting fees and which is also demanded from those persons who have no wish to use public-law broadcasting. This question also applies to entertainment shows offered in similar manner by private broadcasters, but where private broadcasters, on the other hand, are reliant on advertising revenues. Argued in favour of public-law broadcasting is that the latter can guarantee '**quality journalism**' as well as children's programmes that are free of advertising and violence, and is only attractive if it also offers entertainment. It remains uncertain whether these arguments are persuasive, especially in view of the media use by younger recipients for whom public-law offerings tend to play a lesser role. Ultimately, such demands are merely legal policy considerations for the future development of the law since the Federal Constitutional Court, even in more recent decisions, had not even begun to depart from its previous model of the dual broadcasting system including the functional remit of public-law broadcasting. Since the Federal Constitutional Court infers this 'system' from Article 5 (1) Basic Law, this represents a constitutional law requirement that also cannot be rejected by the legislature but especially not by the federal state legislature.

As far as the broadcasting context is concerned, 'Media Law 4.0' initially more or less leaves things as they were. Some of the details are still being wrestled with, for example, the extent to which the public-law broadcasters may also provide press-like Internet offerings. The present system represents a compromise between the right to provide the basic supply also on the Internet and the interests, in particular, of daily newspapers in being able to cultivate this area without competition from public-law broadcasters. According to the compromise, public-law broadcasters are prohibited from making 'non-broadcast-related offerings similar to the press' (Cologne Higher Regional Court, *(Oberlandesgericht Köln)*, judgment of 30.9.2016 – 6 U 188/12). In future, public-law broadcasters are not to be allowed to have 'press-like' offerings available on their websites; instead, the websites must have their emphasis on moving images and sound. In contrast, there is also discussion as to whether the functional remit of public-law broadcasting should not be defined more strongly in favour of an information-oriented remit, which would not currently be compatible with the FCC interpretation of broadcasting freedom. The legislature is faced with the task of more precisely classifying the diverse forms of offerings because telemedia offerings have blurred the dividing lines between the classic opposites of radio/television and the press.

4.3 Private Broadcasters

4.3.1 Relationship to Public-Law Broadcasting

Since, as outlined above, 'Broadcasting Law 4.0' will also be dominated by public-law broadcasters, private broadcasting is organised differently from the press in legal terms. The existence of public-law broadcasting procures advantages to some extent for private broadcasting, at least according to the case law of the FCC. Given the basic supply remit of public-law broadcasting, the requirements made of private broadcasters in terms of diversity are lower than they would be if there were no public-law broadcasting (*'Niedersachsen'*, FCC, judgment of 4.11.1986 - 1 BvF 1/84). On the other hand, there is no mistaking the competitive advantage of public-law broadcasters with regard to the guarantee of their existence and the associated state-funding guarantee.

4.3.2 Streaming Services As Broadcasting?

If broadcasting is offered on the Internet, then the requirements of the State Treaty on Media *(Medienstaatsvertrag)* for broadcasting take effect, in particular the duty of the broadcasting provider to obtain a **licence,** which incurs costs. In view of the convergence of the media with some streams and channels (e.g. via live-streaming portals such as Twitch, YouTube or Lets-Play) becoming more similar to broadcasting offerings, especially if these streaming services offer live streams at regular intervals, some federal states media organisations are of the opinion that such offerings should be subject to the broadcasting regime (Bodensiek and Walker 2018, p. 137). In the State Treaty on Media, an exception was to be codified for licence-free broadcasting programmes (Section 54 (1) MStV). Excepted accordingly will be programmes which, inter alia, owing to their limited importance for the formation of individual and public opinion or which reach, or are forecast to reach, fewer than 20,000 users simultaneously on average over a period of six months. This should put an end to the dispute conducted over a longer period as to whether bloggers also require broadcasting approval.

4.3.3 Media Concentration Law

Nevertheless, there are various limits to the relaxed treatment of private broadcasting with regard to the diversity of opinion. Private broadcasters must also guarantee a minimum degree of **balanced** content of the programmes, **objectivity** and **mutual respect** (*'FRAG'*, BVerfG, judgment of 16.06.1981 - 1 BvL 89/78). Moreover, MStV makes requirements of the guarantee of **diversity of content** if a media enterprise appears capable of acquiring a dominant influence on opinion, as far as the duty to establish a programme committee which has functions similar to those of the broadcasting board of the broadcasting organisations.

The issue of diversity of opinion becomes important to private broadcasters primarily when they are the sole providers at local or regional level. In these cases, diversity of opinion does not come from a large number of offerings—which possibly originate from the same source—but from a multitude of providers. To prevent a dominant influence on opinion as far as possible, it may be necessary to prohibit mergers of competing broadcasters ('broadcasting centre model') whereas, on the other hand, spatial mergers in the sense of 'office-sharing' could be admissible (Fechner and Arnhold 2014, p. 288).

The legislature may make requirements of the financial resources of a broadcaster or of pluralism within a broadcaster; however, the legislator is generally prevented from making demands on the content of the broadcasts. Accordingly, any conditions imposed on a target age group to be reached with priority or to the distinction of the content from that of another specific radio broadcaster would be irreconcilable with the broadcasting freedom of Art. 5 (1) sentence 2, 2nd variant Basic Law. The duty of the legislator to exercise restraint on the content and programming applies with regard to public-law broadcasting organisations but even more so towards private broadcasters.

Not yet sufficiently explored are connections between different forms of media, i.e. 'cross-media offerings' which are particularly capable of undermining diversity of opinion. The ever-stronger growing role of the intermediary has hitherto been ignored.

4.4 Telemedia Providers

4.4.1 Telemedia Act *(Telemediengesetz)*

Telemedia are subject twofold to statutory requirements. First, they are covered by the Telemedia Act *(Telemediengesetz, 'TMG')*, a federal statute that governs the fundamental issues of all telemedia. Insofar as telemedia specifically provide journalistic-editorial offerings, regulations of the State Treaty on Media also apply.

TMG subjects telemedia to different **information duties** (Section 6 TMG) and governs their **liability** in individual cases. Section 7 ff. TMG primarily contains exclusions of liability because otherwise business models such as those of access-providers and service-providers would fail solely because of the general liability provisions of civil law and criminal law. In contrast, the provider remains liable for own contents to which third-party contents are nevertheless also attributed if the provider 'makes them its own'. There is a proliferation of case law on the details of liability which is not always consistent.

4.4.2 Comparison With Broadcasting Providers

On the Internet, alongside enterprises which merely offer access to the Internet or enable the exchange of contents, **content providers** are primarily of importance for media law. They are subject to much less regulation than broadcasting providers. Telemedia do not require a licence or even just registration (Section 4 TMG) whereas broadcasting providers—even if they disseminate their programme via the Internet—require a licence (Section 52 (1) first sentence MStV). The more liberal regulation of content providers on the Internet can lead, in view of the convergence of the media, to two contradictory legal policy requirements. Since the influence on the opinion of the population by Internet media will hardly be less than by broadcasting, it is demanded either that content providers be regulated like broadcasting or, however, that the organisation of broadcasting be made less stringent.

4.4.3 Ad-Blockers

Telemedia providers are normally reliant on advertising revenues. State regulation of advertising contents has, for this reason, a direct impact on the corresponding Internet websites. An interesting question concerns the extent to which state provisions take effect if companies block third-party advertising. This business model is based on software (e.g. 'Adblock Plus') that makes it possible to block advertising content from websites. While certain advertising contents are blocked (blacklisting), 'acceptable' advertising is approved (whitelisting)—if the companies placing the advertising have paid for the same. This issue at the interface of fair competition law and media law was the subject of a decision by the Federal Court of Justice *(Bundesgerichtshof, 'BGH')* in 2018. Therein the BGH affirms the application of the **competition law** codified in the Act against Unfair Competition *(Gesetz gegen den unlauteren Wettbewerb, 'UWG')*. The fact that the software is provided to users free of charge and that advertising is also partly unlocked by the whitelist function free of charge does not prevent its being legally classified as a 'commercial practice' *(geschäftliche Handlung)* as defined in Section 2 (1) No. 1 UWG. Also assumed is a concrete competitive relationship as defined in Section 2 (1) No. 3 UWG. Nevertheless, the BGH negated that the procedure was an act of unfair competition. It did not involve either deliberate obstruction (Section 4, No. 4 UWG) or a general market distortion as defined in Section 3 (1) UWG and also not the exercise of an aggressive influence under Section 4a (1) UWG. The Federal Constitutional Court must still clarify whether this decision gives adequate consideration to the media freedom of the media service-providers affected.

4.4.4 Media Platforms and User Interfaces

Another concern of the federal state legislator was to adapt the **regulation of platforms** to the consequences of convergence. For this reason, it is no longer the country in which the platform operator has its registered office which is decisive, but the principle of market location. Since providers of media platforms, media intermediaries and user interfaces frequently do not have their seat or registered office in Germany, the State Treaty on Media is declared to apply to services from foreign countries if they are intended for use in Germany. This is definitely the case if, in the overall picture, especially because of the language used, the contents offered or the marketing activities are aimed at users in Germany (Section 1 (8) sentence 2 MStV). This then remains the case if they earn 'an essential part of their refinancing in Germany'. This refers to services from other countries, but which are largely financed by advertising clients from Germany. The attempt to subject foreign providers to national regulations has met with European law reservations, inter alia, of the Commission, for which reason the need for a domestic representative was dispensed with; this would have considerably facilitated enforcement of the law. Regulated initially are platforms. Under Section 2 (2) No. 14 MStV, a media platform is every service which combines broadcasting or telemedia similar to broadcasting into one overall offering determined by one provider. The definition of the platform in MStV differs from how the same term is used in the Network Enforcement Act *(Netzwerkdurchsetzungsgesetz, 'NetzDG')* which primarily covers social networks (Section 1 (1) sentence 1 NetzDG). While this concerns a State Treaty of the federal states on the one hand, and a federal statute on the other, the use of the same term in different ways is at least confusing. Providers of media platforms, unless they can rely on an exception as smaller platforms, have certain duties defined in MStV. First, the operation of a media platform must be notified to the competent federal state media organisation at least one month before going into operation (Section 79 (2) MStV). In addition, media platforms are subject to various duties with regard to the selection and treatment of the offerings of third parties conveyed by them. The platform provider may not by itself modify offerings of third parties, or incorporate them in offer packages or market them elsewhere and, above all, it may not superimpose advertising or other contents on them (Section 80 (1) MStV). To safeguard diversity of opinion and choice, there must equality of opportunity of access to media platforms and freedom from discrimination. Broadcasters, broadcasting-type or press-like telemedia may not be unduly obstructed in their access to media platforms and may not be treated differently from similar media without objective reasons (Section 82 (2) MStV). These requirements are supplemented by a prohibition of discrimination on the findability of all offerings in Section 84 (2) sentence 3 MStV. The prohibition of discrimination is supplemented by a requirement of transparency. This obliges the media platform provider to ensure transparency especially of the criteria according to which sorting, putting in order and presenting is made (Section 85 MStV).

It is necessary to mention user interfaces in connection with the platforms. According to the description in Section 2, No. 15 MStV, these are the overview of offerings or contents of individual or multiple media platforms conveyed by means of text, image or sound which serve as a guide and directly allow the selection of offerings, contents or software-based applications. As cases of application, the State Treaty particularly mentions overviews of offerings or programmes of a media platform.

Notification of the operation of a user interface or of a media platform is compulsory (Section 79 (2) MStV). As regards the user interfaces, the findability of offerings is of particular importance. Similar types of offerings or contents may not be treated differently for the purposes of findability, in particular their sorting, order or presentation without objectively justified reason and their findability may not be unduly obstructed (Section 84 (2) MStV). In addition to this prohibition of discrimination, public-law programmes and offerings 'which contribute considerably towards diversity of opinion and choice within the federal territory' must be easy to find. The user must also be able to define the sorting and ordering of the offerings individually for his or her own purposes in favourites lists (Section 84 (6) MStV). Added to this is the duty of transparency of the provider under Section 85 MStV. This includes, inter alia, the criteria according to which contents are sorted, ordered and portrayed. The information thereon is to be made available to users in easily accessible form (Section 85 sentence 3 MStV).

4.4.5 Intermediaries and Providers of Social Networks

In the light of more recent technical developments, the legislators of the State Treaty on Media *(Medienstaatsvertrag)* focus on intermediaries. The State Treaty understands '**media intermediary**' as being every telemedium which aggregates, selects and makes generally accessible journalistic offerings of third parties without bundling these into a combined offering (Section 2 No. 16 MStV). This especially includes, inter alia, search engines, social networks, app portals, blogging portals, etc. Of great practical importance are digital language assistants ('Alexa', 'Siri', etc.) which have high numbers of users and which are, not least, relevant for the formation of opinion because they have to give an even more limited range of answers to questions than search engines. The regulations for media intermediaries are, therefore, to cover all offerings which also offer contents of relevance for opinion forming which are not already included in the term 'media platform'. Examples are search engines and social networks which are also described as open media. To be covered are search engines, social networks but also user-generated portals and blogging portals. Certain offerings are almost completely exempt from the duties for intermediaries. This is initially the case for smaller media intermediaries. According to the definition in the State Treaty, these are media intermediaries who reach fewer than one million users in Germany per month on average over six months. Two other exceptions exist for intermediaries who specialise in the presentation of goods or services, or which serve solely personal or family purposes (Section 91 (2) MStV). These providers are required to ensure transparency (Section 93 (1) No. 2 MStV) and to identify social bots (Section 18 (3) MStV) and there are prohibitions on discrimination (Section 94 MStV).

Social networks play a major role in the age of 'Media Law 4.0'. Owing to content generation by amateur journalists and the possibility of pseudonymisation and anonymisation, fake information and hate crimes are disseminated in social media in a manner inconceivable in 'classic media'. The state attempts, by means of the Network Enforcement Act *('NetzDG')*, to hold providers themselves responsible by obliging them to delete certain unlawful content. Expressly excluded are platforms with journalistic-editorial offerings for which service-providers are themselves responsible. Another important exception arises from Section 1 (2) NetzDG for social networks with fewer than two million registered users in Germany. **Network service-providers** must remove obviously unlawful contents within 24 h of receipt of the complaint or block access to them (Section 3 (2) No. 2 NetzDG). Other unlawful contents must be removed or access to them must be blocked without delay, normally within seven days of receipt of the complaint (Section 3 (2) No. 3 NetzDG). In addition the providers are obliged to inform state authorities about illegal contents (Section 3a NetzDG). The provisions of NetzDG are of great importance to network service-providers because Section 4 NetzDG threatens high fines in the event of contravention.

Various persons have expressed doubt about the lawfulness of NetzDG (Liesching 2018, p. 26ff.; Müller-Franken 2018, p. 1 ff.; Peifer 2018, p. 14 ff.; Fechner 2018, p. 157 ff.), for example, as regards the jurisdiction of the federal government because this involves the regulation of media contents which is a matter reserved for the federal states. Criticism of the content of the statute is primarily based on impairment of the freedom of speech of the users. Although it is not a matter of a direct state intervention, the legislature nevertheless defines a specifically structured procedure to which non-compliance can be sanctioned by high fines (up to 50 million euros, Section 4 (2) NetzDG in conjunction with Section 30 (2) sentence 3 OWiG). The threat of a fine combined with the short deadlines set for deletion makes it likely that 'overblocking' will occur, which entails impairment of the freedom of speech of Art. 5 (1) sentence 1, 2nd variant Basic Law; especially as there are no reasonable opportunities for legal protection against deletion. Thus, the statute appears not to have been properly thought out and is also likely to be an unconstitutional attempt by the legislature to fight situations which do—in fact—justify complaint. The statute can provide no protection against '**fake news**' either inside or outside social networks. Much more dangerous than statements that conform to the definitions of criminal offences listed in NetzDG are tendentious reports or reports which twist the facts but the which recipients believe have been properly journalistically researched, such as those disseminated, e.g. by agencies which are financed by foreign states (*Frankfurter Allgemeine Zeitung*, 'FAZ' of 23.1.2019, p. 3). It would be illusory to believe that the state can protect us against fake news, but it would also be dangerous to lose a distrust of media contents. The fact that the 'classic media' are also not spared this mistrust was demonstrated recently by the 'Riotous case' already mentioned. This makes it even more important to remain aware of the content risks in the age of 'Media Law 4.0'.

5 Recipients

5.1 Freedom of Information

Other players in media law are the **recipients** who are not only the target group for media activity but who also directly or indirectly finance the same. Particularly in a Media 4.0 context, the freedom of information of Art. 5 (1), sentence 1, first variant of the Basic Law *(Grundgesetz)* must stand the test. Freedom of information includes the right to inform oneself without hindrance from generally accessible sources. These also particularly include foreign newspapers and the various Internet sources. The state may not attempt to obstruct or ban access to such sources. Since media services in Media 4.0 are less and less one-sided, it seems that freedom of information must be seen in connection with the freedom of the media. Freedom of information and the freedom of media activity are then no longer separate fundamental rights but may be understood as the fundamental right to use the media.

5.2 General Public's Interest in Having Freely Available Information

In each case involving media law, the interests of the recipients must also be considered as part of the 'general public's interest in having freely available information'. This right, which is similar to a fundamental right, is derived by the FCC from the freedom of the press, freedom of broadcasting or freedom of speech. It supplements the freedoms of the media and can fight on behalf of the media. Media contents which serve the general public's interest in having freely available information tend to be more admissible than contributions which merely wish to exploit the public desire for sensationalism for their own purposes. This principle, which is to be observed when rights are weighed against each other, must also be respected with regard to social networks.

6 Persons Affected

6.1 Privacy Rights

Privacy rights are particularly under threat in the digital media age. The skimming of data (as affected numerous politicians on a grand scale at the beginning of 2019), identity theft and cyberbullying are as equally relevant to media law as reporting which breaches privacy rights. If privacy rights are affected, there is a need to weigh the gravity of the breach of privacy on the one hand against the importance of the uncensored message to the general public on the other also in social networks. The gravity of the breach of privacy in social media may be difficult to quantify. Whereas the degree of dissemination in the case of the 'classic media' can be approximately calculated, this is hardly the case for social networks owing to the viral form of the dissemination of contents. In case of doubt, the effect may not be underestimated. Peculiarities particularly arise when **reporting suspicions**. If a person is suspected of having committed a crime, such reporting does not require any absolute certainty about the offender; however, the presumption of innocence also applies in the media, for which reason high demands are made of the journalist's duty of care, if criminal-law or civil-law sanctions are not to be imposed. In social networks, the inhibition threshold for reporting suspicions is very low. Moreover, the persons expressing opinions are in many cases significantly less knowledgeable about the law than professional journalists. Also reporting about celebrities (Stegmann 2018, p. 377; Thalmann 2018, p. 476ff.) requires weighing of the interests of the persons affected in the individual case.

6.2 Rights to Own Picture

A specially developed privacy right is the ‘right to your own picture’. This right has become highly developed because of the more than 100-year-old **Art Copyright Act** *(Kunsturhebergesetz, ‘KUG’)* and the extensive case law based on the same. However, more recent developments of so-called ‘Media Law 4.0’ in this respect are also leading to new confusion and ambiguity.

The European legislature has created a strong instrument for the protection of personal data by means of the **General Data Protection Regulation** (GDPR). The instrument is so strong because, as a regulation, it takes precedence over national law and, therefore, also over KUG wherever the areas of application overlap. Taking photographs is classified in Art. 2, 1. GDPR as ‘automated processing of personal data’. Thus, the strict requirements of GDPR on the consent of the person photographed and of duties of information and erasure generally when taking photographs must be observed (whereas KUG merely governs the publication and dissemination of photos). However, there are various exceptions which restrict the applicability of GDPR primarily in those cases where data processing is ‘necessary for the purposes of the legitimate interests pursued by the controller or by a third party’ (Art. 6, 1. (f) GDPR). This provision is interpreted in such a way that press photographers can also invoke the same. Moreover, Art. 85 (1) GDPR allows Member States to create legal provisions which are intended to reconcile the protection of personal data with the right to freedom of expression and information. It is disputed whether this also covers the—longstanding—KUG. If the former Federal Ministry of the Interior—probably correctly—assumes so as well as in meantime some law-courts, then only the European Court of Justice can finally clarify the matter. Until then, unless pragmatic solutions are pursued, uncertainty for photographers will remain.

7 Outlook

Whereas no field of law entitled ‘Media Law 4.0’ hitherto existed, numerous media-law problems nevertheless arise from the technical developments and the new business models associated with a Media 4.0 context. Aspects of law which fully consider the technical developments are still in the process of emerging. As long as new technical possibilities are constantly arising, the development of these aspects of law cannot reach completion.

Quite obviously, it is not merely a question of clarifying the extent to which existing statutory provisions can be applied to the new phenomena. Rather the question arises as to how the state can create a new regulatory framework geared to the competitive situation between the media providers and adequately safeguard interests in protecting young persons, privacy rights and intellectual property. This cannot be done without regulation, however, state regulation of the media may neither impair diversity of opinion nor lead to state influence on media contents. Finally to be noted is that purely national solutions soon reach their limits and even Europe-wide regulations are not sufficient, so that there is a need to entrench a genuine 'Media Law 4.0' in public international law.

References

Beaujean D (2018) Gewinner und Verlierer der Medienkonvergenz, Multimedia und Recht (MMR), 3 ff

Bodensiek K, Walker M (2018) Livestreams von Gaming Video Content als Rundfunk? – Notwendigkeit einer verfassungskonformen Auslegung der Rundfunkzulassungspflicht, MMR, 136 ff

Fechner F (2018) Fake News und Hate Speech als Gefahr für die demokratische Willensbildung. Staatliche Gewährleistung kommunikativer Wahrheit? In: Uhle (ed) Information and Einflussnahme, Berlin 2018, 157 ff

Fechner F (2021) Medienrecht. Lehrbuch des gesamten Medienrechts unter besonderer Berücksichtigung von Presse, Rundfunk und Multimedia, 21th edn., Tübingen

Fechner F, Arnhold J (2014) Meinungsvielfalt, Anbietervielfalt und Angebotsvielfalt – Medienkonzentration im neuen Thüringer Landesmediengesetz auf dem Prüfstand der Rundfunkfreiheit, Thüringer Verwaltungsblätter (ThürVerwBl), 285 ff

Gounalakis G (2016) Geldentschädigung bei vorverurteilenden Äußerungen durch Medien oder Justiz, Neue Juristische Wochenschrift (NJW), 737 ff

Heins M (2018) Anwendbarkeit rundfunkrechtlicher (Werbe-)Vorschriften auf Online-Videos, MMR, 794 ff

Hofmann R, Fries P (2017) Der äußerungsrechtliche Geldentschädigungsanspruch im digitalen Zeitalter, NJW, 2369 ff

Leeb C, Maisch M (2019) Social-Media-Stars und -Sternchen im rechtsfreien Raum?, Zentrale für Unterrichtsmedien im Internet (ZUM), 29 ff

Liesching M (2018) Die Durchsetzung von Verfassungs- und Europarecht gegen das NetzDG – Überblick über die wesentlichen Kritikpunkte, MMR, 26 ff

Mafi-Gudarzi M (2018) #MeToo: Wieviel Wahrheit ist erlaubt, Neue Juristische Online-Zeitschrift (NJOZ), 521 ff

Milker J (2017) "Social-Bots" im Meinungskampf, ZUM, 216 ff

Müller-Franken S (2018) Netzwerkdurchsetzungsgesetz: Selbstbehauptung des Rechts oder erster Schritt in die selbstregulierte Vorzensur? – Verfassungsrechtliche Fragen, AfP – Zeitschrift für Medien- und Kommunikationsrecht (AfP), 1 ff

Peifer K (2018) Netzwerkdurchsetzungsgesetz. Selbstbehauptung des Rechts oder erster Schritt in die selbstregulierte Vorzensur? – Zivilrechtliche Aspekte, AfP, 14 ff

Rodenbeck J (2018) #MeToo vs. Persönlichkeitsrechte und Unschuldsvermutung, NJW, 1227 ff

Schütz R (2018) Regulierung in der digitalen Medienwelt, MMR, 36 ff

Stegmann O (2018) Der Einkauf des fürsorgenden Familienvaters im Supermarkt als Zeitgeschehen – plus commentary on FCJ, judgment of 6.2.2018, Kommunikation & Recht (K&R) 375 ff

Steinbach A (2017) Social Bots im Wahlkampf, Zeitschrift für Rechtspolitik (ZRP), 101 ff
Thalmann D (2018) Die Gemeinfreiheit der Prominenz – Werbung mit Abbildern von Personen des öffentlichen Interesses im Spannungsfeld von Individual- und Allgemeininteressen, Gewerblicher Rechtsschutz und Urheberrecht (GRUR), 476 ff
Ulbrich C, Frey B (2017) Mediale und gesellschaftliche Vorverurteilungen – Ein neuer Ansatz zum Umgang mit einem alten Phänomen, ZUM, 31 ff
Volkmann V (2018) Hate Speech durch Social Bots – Strafrechtliche Zuordnung von Volksverhetzung gem. § 130 Abs. 1 StGB, MMR, 58 ff
von Graevenitz A (2018) "Zwei mal Zwei ist Grün" – Mensch und KI im Vergleich, ZRP, 238 ff
Weberling J (2018) Medienrechtliche Bedingungen und Grenzen des Roboterjournalismus, NJW, 735 ff
Willems C (2018) Influencer als Unternehmer, MMR, 707 ff
Wimmer N, Nawrath K (2016) Der schöne 15. Juni 2011 – Anmerkungen zu BGH ZUM 2015, 989 – Tagesschau-App, ZUM, 126 ff

Data Protection 4.0

Axel von dem Bussche

Contents

A. von dem Bussche (✉)
Taylor Wessing Partnerschaftsgesellschaft mbB, Hamburg, Germany
e-mail: A.Bussche@taylorwessing.com

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_8

1 Introduction

1.1 The Term Data Protection 4.0

'Data Protection 4.0' is an abbreviated and perhaps somewhat misleading term. At first glance, it suggests the existence of clearly definable development stages of data protection law or milestones falling within successive periods. Where there is version number 4.0, there must surely have been versions 3.0, 2.0 and 1.0. Such events can of course be accounted for. The enactment of the world's first data protection law in 1970 in Hessen, Germany, the introduction of the Federal Data Protection Act (BDSG) in 1978, the census ruling of the Federal Constitutional Court in 1983 and most recently the applicability of the General Data Protection Regulation (GDPR) in May 2018. Without a doubt, these events also heralded radical changes in data protection law. However, their significance for the development of the legal field is more complicated than simply updating the version as the numbering suggests. Data Protection 4.0 can only be explained against the background of the concept of Industry 4.0. For a comprehensive view, the concept must be considered from two perspectives:

On the one hand, Industry 4.0 raises the question of how data protection law under current law relates to the developments outlined above. Although the term Industry 4.0 describes a strategic vision for the future of industry, the changes from the advent of digitalisation are already in place today and are becoming increasingly noticeable. If one measures the pace of data protection law reforms according to the speed of change in the past, it will be quite some time before the next major turning point. In the meantime, the usual challenge in the digital legal sphere is that the law must consider developments that may not have been sufficiently considered when such legislation was created. Furthermore, even if the law is designed to be sufficiently flexible and well prepared for the (near) future, the underlying social views may also change. What was previously frowned upon as invasive to privacy may become socially acceptable in an economically changing environment. In both cases, the expectations placed on the law by society and business diverge in terms of protection and appropriateness. This raises different questions for data protection law with regard to Industry 4.0: Can it keep up with the changes brought about by Industry 4.0? Are the current premises compatible with the expected changes? Where do areas of tension arise? In short: How does Industry 4.0 fit in with current data protection?

On the other hand, the question arises as to what data protection law should look like in the future in Industry 4.0. Although a reform of data protection law was recently on the agenda, the following applies: After the reform is before the reform. The frictions arising in the course of the first question between a competitive Industry 4.0 and data protection must be resolved appropriately in the interest of competitiveness. In doing so, however, the goals and protected interests of data protection must not be abandoned; in this respect, they function as crash barriers for modern data protection law. Building on the first question, it must therefore be clarified how the tensions between data protection and Industry 4.0 can be resolved. What problems does data protection law must deal with, what tools does it need to have at the ready? What are the goals of data protection law in Industry 4.0 and how can it achieve them? In other words: How can data protection fit in with Industry 4.0 in the future?

These two questions cannot be answered separately. Wherever the status quo in terms of data protection law does not do justice to the development that has already taken place, the next step must be to consider whether and how a remedy can be found. Therefore, data protection circumstances will be compared below with the issues arising in the context of Industry 4.0 and areas of tension will be identified together with concrete solutions.

1.2 Changes Due to Industry 4.0 and Relevance of Data Protection

The core of Industry 4.0 is the digitalisation of production. When the value chain is optimised through intelligent networking of production systems, this is also accompanied by the processing of a previously unrealised potential of data. This is made more favourable by cheaper procurement and the resulting greater use of sensor technology and storage media. As the existing amount of data rapidly grows, a wealth of data is emerging. Industry 4.0 therefore also gives rise to the **emergence of a data economy**. The application scenarios developed by Plattform Industrie 4.0 (hereinafter referred to as Plattform Industrie 4.0 2016a) provide an overview of the expectations associated with the term Industry 4.0. The following areas are particularly relevant in terms of data protection law:

Order-driven production, i.e. the networking of production capabilities beyond one's own factory limits, requires an exchange of data between the producing companies. With the aim of making production more flexible with regard to customer wishes, personal data may also be involved. The same applies to self-organising adaptive logistics, which should allocate products in a responsive manner. The area of value-based services envisages using process and status data from production and production use as raw material for the optimisation of products and the creation of new services. The idea of making already delivered products capable of transformation in the interest of product maintenance or customised services is based on findings from processing usage and status data. Another field is human-technology interaction in production. Both physical support through capability enhancers and digital assistance systems benefit particularly from personalisation to the respective user. Data is therefore at the heart of these application scenarios.

With the increasing importance of data for business models, the view of the accompanying data protection law is also changing. Apart from the erroneous classification as a mere compliance task, the analysis of data protection law is moving further to the beginning of the value chain and already finds a place in the selection and design of the business models (Moos 2015, p. 12). Since the human factor as an employee is (still) indispensable in Industry 4.0, changes in the working environment mean challenges for employee data protection. The great importance of this data for Industry 4.0 therefore means that data protection law is of **paramount importance**.

1.3 Overview of Data Protection Regulations

The purpose of data protection law is to protect natural persons when processing personal data, see Art. 8 (1) CFR. To achieve this purpose, there is a confusing number of special regulations. The provisions of the European General Data Protection Regulation (GDPR) and the BDSG, which supplement it, are authoritative. In addition, various sector-specific, more specific data protection regulations are relevant for Industry 4.0 applications, particularly Sections 11 et seq. Telemedia Act (TMG), Sections 88 et seq. Telecommunications Act (TKG) and finally Sections 49 et seq. of the Metering Point Operations Act (MsbG) (see Lüdemann et al. 2016, p. 125 et seq.).

The **GDPR**, a regulation from the European Union, is directly applicable law. After the first presentation of a draft by the Commission in 2012 and an extensive legislative process, it entered into force on 24 May 2016 and has been applicable since 25 May 2018. The **BDSG**, previously the central regulatory regime for data protection in Germany, was adapted to the GDPR accordingly. The BDSG applies where the GDPR grants the Member States room for manoeuvre (see further Kühling 2017, p. 1985) and is subsidiary to specialised area-specific data protection law.

The data protection regulations of **Sections 11 et seq. TMG** address service providers under Section 2(1) No. 1 TMG. It is planned to adopt an amended version of the underlying ePrivacy Directive (Directive 2002/58/EC) in the form of a regulation. This so-called ePrivacy Regulation is intended to specify the GDPR with regard to electronic communication data and would replace the data protection provisions of the TMG. However, the legislative process has been considerably delayed; entry into force is not expected before 2020 and applicability is not expected before 2022. Sections 11 et seq. TMG are no longer applicable in the meantime according to the prevailing view (Conference of the Independent Data Protection Supervisory Authorities of the Federal Government and the Federal States 2018, p. 2 et seq.; Gola, in: Gola 2018, Art. 95 (19) with further references), meaning that the GDPR applies.

Sections 88 et seq. TKG are basically directed at service providers under Section 3 No. 6 TKG. They are based on Directive 2002/58/EC. Due to the conflict of laws rule of Art. 95 of the GDPR, the obligations and permissions for telecommunications services are based on the implementation of the Directive in the TKG, not on the GDPR (Schmitz 2018, Section 11 TMG para. 21). Consequently, there is initially no change for these services until the applicability of the ePrivacy Regulation.

2 Scope of Application

2.1 The Facts

The provisions of data protection law are applicable to the (automated) processing of personal data (see Art. 2 (1) GDPR, Sec. 1(1) Sentence 2 BDSG, Sec. 91 TKG). If there is no personal reference, the scope of application of data protection law is not available. In the context of Industry 4.0, the decisive factor for opening up the scope of application is therefore the question of whether personal data is processed at all.

2.1.1 Reference to Persons

Personal data is defined in the GDPR as any information relating to an **identified or identifiable natural person** (Art. 4 No. 1 GDPR). A person is identifiable if he or she can be identified directly or indirectly, particularly by means of an association with an identifier. Examples of such an identifier are the name, an identification number, location data, an online identifier or other special characteristics that are an expression of the identity of the person. Identification does not mean naming; it is sufficient that a person can be directly identified from the information itself (see decision of CJEU dated 19.10.2016 - C-582/14, ECLI: EU: C: 2016: 779, para. 38 - Breyer) a person can be individualised, i.e. recognised (Karg, in: Simitis 2019, Art. 4 No. 1 para. 49).

The question of **identifiability** is more difficult to assess. What is meant is whether a link could be established between the information and the person. Under the previous law, it was disputed whether it was sufficient that any third party could establish a link (objective/absolute theory), or whether it was precisely the body responsible in the specific individual case that had the means and would make the effort to establish a link (subjective/relative theory) (see Karg, in: Simitis 2019, Art. 4 No. 1 para 49 with further references). The GDPR also requires the means of third parties, not only those of the controller, to be considered (see Recital 26 to the GDPR). However, the use of these means must also be probable according to general discretion and not merely hypothetically possible. In particular, the costs, the time required, the technology available at the time of processing and its development must be considered (see Recital 26). Therefore, even purely technical data can be personal data. It must always be examined on a case-by-case basis whether and with what probability a reference to a person can be established.

Accordingly, there is a personal reference if, for example, in the context of order-controlled production or adaptive logistics, products are provided with an identifier of the customer in the manufacturing/shipping process. If, in the area of human-technology interaction, **wearables**, for example, generate information about their wearers, personal data is often involved; in any case, the probability of identifiability increases with the extent of use (Kopp and Sokoll 2015, p. 1352). In the course of collecting data about delivered products, there is also often a personal reference. For example, data collected when a **vehicle is used** is personal if it is linked to the vehicle identification number or registration number (Conference of the Independent Data Protection Authorities of the Federation and the Federal States and the German Association of the Automotive Industry 2016, p. 1). With regard to **IP addresses**, the CJEU supports the classification as personal data insofar as it is possible for the controller to establish a connection to the user in a legally permissible and technically reasonable manner (decision of CJEU dated 19.10.2016 - C-582/14, ECLI: EU: C: 2016: 779, para. 49 - Breyer). Consequently, wherever there are points of contact with employees or end customers in Industry 4.0 processes, a thorough examination of any personal reference is recommended.

2.1.2 Anonymisation and Pseudonymisation

The opening up of the scope of application means that the possibilities of processing and consequently the economic value creation of the data are limited (see Sects. 3 and 4 below). It may therefore make economic sense to eliminate a personal reference from the data to avoid the applicability of data protection law. Not least of all, this approach is less invasive for the rights of data subjects.

Excluded from the scope of data protection law is all non-personal data, which is recognised to be **anonymous and anonymised data**. This is information that from the outset does not relate to an identified or identifiable person and personal data that has been anonymised in such a way that identification can no longer take place (Recital 26 sentence 5). The process of anonymisation itself is not defined in the GDPR. The content of a data set remains unaffected, only the reference is removed (Ernst, in: Paal and Pauly 2018, Art. 4 para. 49). In contrast to personal data, it must be sufficient that identification is unlikely according to general discretion (Roßnagel 2018, p. 244). Data protection law places high demands on this, which are often not met in practice. Criteria for assessing the effectiveness of individual anonymisation techniques are, in particular, whether after anonymisation it is still possible to pick out a person from the data set, to link data sets relating to a person or to infer information about a person (see Art. 29 Data Protection Working Party 2014, p. 3). If this tracing to a natural person is technically possible or if only the transmission of the data is anonymised, the requirements for anonymisation are not met.

Data does not have to have a personal reference for every purpose. When evaluating process and status data from production and product use, it does not necessarily depend on the individual user. Operating parameters and queries can also be evaluated anonymously. For Industry 4.0 scenarios, such as the development of value-based services or the optimisation of versatile products, anonymisation is therefore a helpful means of simplifying the legal situation.

Pseudonymisation also plays an important role in the GDPR. Pseudonymous data is data that cannot be attributed to a person without adding additional information (Recital 26). As a rule, identifiers are replaced by pseudonyms for this purpose. The additional information required for re-identification (so-called attribution rule) must be kept separate from the actual data and be subject to technical and organisational measures to ensure that no attribution takes place (see Art. 4 No. 5 GDPR, Recital 26 sentence 2).

In contrast to anonymisation, the scope of data protection law is in principle applicable in the case of pseudonymisation, because pseudonymised data is nevertheless to be regarded as personal data (see Recital 26, sentence 2). This applies in any event to cases in which the attribution rule remains with the controller or is passed on to a third party, but it is not guaranteed that the controller will not gain access to it. If, on the other hand, it is ensured that the controller does not receive the attribution rule, the controller can no longer attribute the data to individuals. Recital 26 is to be understood in such a way that for this data controller, the data is not personal, but anonymous (see Roßnagel 2018, p. 244).

Pseudonymisation of personal data can also help to comply with data protection obligations (see Recital 28, see Sect. 3.1 below on obligations). For example, pseudonymisation can influence the balancing of interests within the framework of Art. 6(1)(f) of the GDPR in favour of the controller, facilitate further processing for a different purpose (Art. 6(4)(e) of the GDPR), as well as serve the realisation of privacy by design under Art. 25(1) of the GDPR (see Sect. 3.1.2) and the implementation of technical and organisational measures under Art. 32(1)(a) of the GDPR. In this respect, it is advisable for Industry 4.0 processes, insofar as complete anonymisation is not feasible, to pseudonymise data if possible to facilitate fulfilment of obligations under data protection law.

2.1.3 Big Data

A significant phenomenon for Industry 4.0 is big data, i.e. the processing of large amounts of data at a high speed to generate economic benefits (Arning 2015, p. 7). The increasing availability of data (see Sect. 1.2 above) makes it attractive for companies to generate value and efficiency gains from their analysis. The amount of data processed is not in itself relevant for opening up the scope of application. Once again, the sole decisive point is the personal reference.

Big data applications that accumulate data for pattern recognition, such as predictive models, simulations or intelligent networks (Schulz, in: Gola 2018, Art. 6 para. 256), do not regularly rely on personal references. In the interest of processing that is as minimally invasive as possible for data subjects and legally secure for data controllers, it makes sense to use anonymous data or to anonymise existing data with a personal reference. The scope of application of data protection is not triggered (see above Sect. 2.1.2), and the controller is free in processing. However, it should not be overlooked that the concept of personal reference is dynamic. The question of identifiability, i.e. which means are likely to be used to identify a person, must particularly consider the technology and technological developments available at the time of the processing (Recital 26). This means that an initially anonymous or anonymised data can become personalised through technological progress. It is also possible that, over time, a company gains information that makes it likely that it can be identified. The decisive factor for the permissibility of such processing is therefore the **probability of re-identification** (Schulz, in: Gola 2018, Art. 6 para. 256). On the one hand, this risk can be reduced by the anonymisation technique providing a kind of ‘protection reserve’ that prevents identification (see Schaar 2016, p. 225). On the other hand, it is necessary to regularly check whether linkable information obtained in the meantime is likely to make the data identifiable—a single risk analysis is not sufficient for this (Laue et al. 2016, section 1 para. 22), but must be repeated at regular intervals.

2.2 Territorial Cope

The GDPR has a broad territorial scope. First, it applies if the processing of personal data is carried out in the context of activities of an establishment in the European Union (so-called **establishment principle**, see Art. 3(1) GDPR). Establishment presupposes the effective and actual exercise of an activity by a fixed establishment—irrespective of its legal form (see Recital 22, sentences 2 and 3). Under the old law, the CJEU required a combination of personal and material means, as well as a certain degree of consistency (see decision of CJEU dated 1.10.2015 - C-230/14, ECLI: EU: C: 2015: 639, para. 29 - Weltimmo). The independent operation of a server should not be sufficient for this (Art. 29 Data Protection Working Party 2010, p. 15; also Hornung, in: Simitis 2019, section 3 para. 23). The decisive factor is that it is solely the place of establishment that matters, not the place of actual processing (Art. 3(1) GDPR). Outsourcing operations to third countries, e.g. in the context of cloud computing applications, therefore has no influence on the obligations under the GDPR (cf. Hornung, in: Simitis 2019, Art. 3 para. 27). For Industry 4.0 applications of companies based in Germany, the geographical applicability is therefore already regularly determined by the principle of establishment.

If the controller is not established in the Union, the GDPR is furthermore applicable if goods or services are offered to persons located in the European Union and the data processing is related to this or is intended to monitor their behaviour (so-called **place of market principle**, see Art. 3(2) GDPR). For the offer of goods or services, it is irrelevant whether a payment is to be made (Art. 3 (2)(a) GDPR), which means that free Internet services financed by advertising is also covered. Behaviour means all measurable physical activities (Hornung, in: Simitis 2019, Art. 3 para. 27). This includes, for example, the monitoring of user behaviour in the EU from a website operated in a non-EU country. The scope of application therefore goes far beyond the EU borders, which is a declared goal of the GDPR.

2.3 Other Data

The concept of personal reference that characterises the scope of application does not recognise any gradual differences, but is binary (see Karg, in: Simitis 2019, Art. 4 No. 1 para. 14). **Non-personal data** is therefore not subject to data protection regulation in its entirety. On the one hand, this is data for which no identification or identifiability is possible, namely anonymous data. Since the GDPR does not apply to the processing of personal data of legal persons (see Recital 14), it also covers data that is personal but does not relate to a natural person. A reference to a person only exists in exceptional cases if the processing relates to the natural persons behind it (Art. 29 Working Party 2007, p. 27).

This means that, for example, **machine data** generated in Industry 4.0 is free to be processed, but is not protected by data protection law. However, especially for sensitive company-related data, there is an increased need for protection in Industry 4.0 due to comprehensive integration in network structures. Apart from any contractual provisions, the data is protected under constitutional and EU law through freedom of enterprise and freedom of property (Frenz 2016, p. 122); protection under simple law may be provided by the right of the database producer under sections 87a et seq. of the Copyright Act (see Wiebe 2017, p. 338), or the provisions for the protection of company secrets. Also under discussion is protection under property law (see e.g. Wiebe 2016, p. 877 et seq.) or a platform regulatory approach (Spindler, in: Hornung 2018, 151 et seq.). It remains to be seen whether and to what extent the (European) legislator will take action.

For Industry 4.0 application scenarios, free handling of this data is essential. The European Commission has recognised this need and addressed the free movement of non-personal data with Regulation (EU) 2018/1807, which became valid on 20 May 2019. Under Art. 2(1) and Art. 3 No. 1 of the Regulation, it constitutes the legal framework for the processing of electronic data which is not personal data under Art. 4 No. 1 of the GDPR and is therefore the counterpart to the GDPR (see Recital 10 of the Regulation). The Regulation covers for example aggregated and anonymised data sets for big data analysis, data related to precision agriculture and data on maintenance needs of industrial machinery (Recital 9 to the Regulation). If personal and non-personal data is inextricably linked in a data set, the application of the GDPR remains unaffected (Art. 2(2) of the Regulation). In the absence of a personal reference, the Regulation does not adopt the protection perspective familiar from data protection laws, but attempts to improve the free movement of data. To this end, the Regulation prohibits the Member States from imposing data localisation requirements other than for reasons of public security (Art. 4 para. 1 of the Regulation). In return, the availability of data for competent authorities is ensured (Article 5 of the Regulation). Finally, the development of rules of conduct for the transfer of data is promoted through self-regulation of the service providers (Art. 6 of the Regulation). The Regulation explicitly envisages the transfer between cloud providers or to one's own IT system (see recitals 5, 6, 31 to the Regulation).

3 Rights and Obligations

3.1 Responsibility

3.1.1 Definition

The determination of responsibility is of particular importance; the controller is responsible for compliance with the principles for the processing of personal data (Art. 5(2) GDPR), the addressee of the obligations under data protection law and the point of contact for data subjects. '**Controller**' within the meaning of Art. 4 No. 7 of the GDPR is the body which alone or jointly with others determines the purposes and means of the processing. It is significant that it does not matter whether the body itself processes data, but whether it exercises the decision-making power over it. In the case of connected cars, for example, the controller is whoever receives data from the car, which is usually the manufacturers and, if applicable, third-party service providers (Conference of the Independent Data Protection Authorities of the Federal Government and the Federal States and the German Association of the Automotive Industry 2016, p. 2). With regard to smart TVs, device manufacturers are considered data controllers if they process personal data in the context of software updates, app store/portal operators, app providers and operators of personalisation services (Düsseldorfer Kreis 2015, p. 22). A cloud user is a controller, but a cloud provider can also be a controller if it offers additional services, such as appointment or contact synchronisation (Art. 29 Data Protection Working Party 2010, p. 27).

In the context of Industry 4.0, the networking of production structures gives rise to the problem of assigning the multitude of players with new types of responsibilities to the roles that have evolved under data protection law. This complexity is partly controlled by Art. 26 of the GDPR, according to which several persons who jointly decide on the processing are **jointly responsible** (see most recently instructive decision of the CJEU dated 5.6.2018 - C-210/16, ECLI: EU: C: 2018: 388, para. 27 et seq. - ULD v. Wirtschaftsakademie Schleswig-Holstein). In external relations, both processors are jointly responsible; in internal relations, they agree on their responsibility (Martini, in: Paal and Pauly 2018, Art. 26 para. 4). The challenge for the jointly responsible parties is therefore to contractually structure their relationship accordingly (Hornung and Hofmann, in: Hornung 2018, p. 46).

3.1.2 Privacy by Design and Privacy by Default

The concepts of data protection under 'privacy by design' and 'privacy by default' laid down in Art. 25 of the GDPR are highly relevant for Industry 4.0 processes. The obligation to **privacy by design** means that the controller is also already obliged at the time of the development of products or processes to design the processing in a data protection-friendly manner. Industry 4.0 processes should therefore be designed from the outset to be data protection-friendly so that processing is, as it were, automated and legally compliant (Richter 2012, p. 576). **Privacy by default** also focuses on reducing processing to the extent necessary for the respective purpose through the default settings of a system. For the concrete meaning in individual cases, guidelines and recommendations of the European Data Protection Board are to be observed (Albrecht and Jotzo 2017, Part 5, para. 5).

Art. 25 of the GDPR primarily addresses only the controllers of personal data. However, the controller using a system for processing and the producer of this system will often diverge. However, a controller is already obliged to carry out a risk analysis when deciding on the selection of appropriate systems. Producers who are not data protection controllers are therefore exposed to competitive pressure to offer GDPR-compliant systems (Mantz 2018, Art. 25 para. 79 f.). There is therefore an incentive to comprehensively consider the requirements of Art. 25 of the GDPR when designing systems; **producers are indirectly obliged in this respect** (see Recital 78 sentence 4).

Implementation includes, on the one hand, the definition of internal strategies, i.e. guidelines that specifically describe how the requirements of the GDPR are implemented (such as design and documentation of processing, selection of measures, involvement of auditors) (Hansen, in: Simitis 2019, Art. 25 para. 60). Second, concrete measures must be taken, such as minimising processing, pseudonymisation as quickly as possible, establishing transparency with regard to the functions and processing of the system, enabling the data subject to monitor the processing and improve security functions (see Recital 78). It should be emphasised that this is an obligation of the controller that can be sanctioned. It is therefore essential to document the implementation efforts, as also provided under Arts. 5(2) and 24(1) of the GDPR.

3.1.3 Information Requirements

The processing is accompanied by information obligations for the controller under Arts. 12–14 of the GDPR. Compared to the old law, their scope has been **considerably expanded** (see Laue et al. 2016, section 3 para. 3). Following the catalogue of Arts. 13(1) and 14(1) of the GDPR, the controller must, in addition to the information already provided under the old BDSG, provide the contact details of the data protection officer (if any), in the case of processing under Art. 6(1) (f) of the GDPR, make transparent the balancing exercise carried out, inform about the intended legal basis, name the recipients or categories of recipients and inform about the intention to transfer data to a third country, as well as whether an adequacy decision or appropriate safeguards exist. Under Arts. 13(2), 14(2) of the GDPR, the following information must also be provided in deviation from the old legal situation: Storage period of the data, information on data subject rights, reference to the right of withdrawal in the case of consent, existence of a contractual obligation to collect data or a data collection that is necessary for the conclusion of a contract, existence and, if applicable, disclosure of the logic and scope of automated decision-making. The time of issue is **when the data is collected** from the data subject. If the data is not collected from the data subject, the information must be provided no later than one month after collection (Art. 14(3) GDPR). If the data subject already has the information, the information is not required (Art. 13(4), 14(5)(a) GDPR).

Industry 4.0 processes with direct contact with data subjects must therefore be structured in such a way that the data subject is informed before the first processing operation. If it is not a direct collection, the corresponding effort must be made to provide the information within one month at the latest. If the provision of information requires a disproportionate effort, it may exceptionally be omitted. The higher the effort and the lower the interest of the data subject in the information or the risk of the processing, the more likely this is the case (see Bäcker, in: Kühling and Buchner 2018, Art. 14 para. 55). The exception that also exists for statistical purposes is not relevant for big data applications that serve economic purposes such as customer behaviour analysis and whose results are to be used for decisions vis-à-vis individuals (Caspar, in: Simitis 2019, Art. 89 para. 24; Roßnagel et al. 2016, p. 159). In the case of smart systems, the information can be provided to the data subject not only in written text, but also, for example, via QR codes, videos, SMS/email or audio instructions (see Art. 29 Data Protection Working Party 2018, p. 21).

3.2 Order Processing

Order processing is the instrument under data protection law for the controller to have processing operations carried out by other entities and is therefore also highly relevant current law in Industry 4.0 processes, which often provide for processing by different entities. Under Art. 4 No. 8 of the GDPR, a **processor** is any entity that processes data on behalf of the controller. In distinction to the so-called transfer of functions between two controllers and joint responsibility, this refers to bodies that do not decide on the processing themselves, but are bound by the mandate of the controller and carry out the processing in its interest (Petri, in: Simitis 2019, Art. 28 para. 21). **No further authorisation standard is required** for the involvement of the processor (Albrecht and Jotzo 2017, Part 5 para. 22; Petri, in: Simitis 2019, Art. 28, para. 33 with further references). Article 28(3) of the GDPR sets out specific requirements for the contract between the controller and the processor, which can be concluded in the text form of Section 126b of the German Civil Code (BGB) (cf. Hartung, in: Kühling and Buchner 2018, Art. 28 para. 96 with further references). In particular, it must be provided that data is **only processed on the instructions** of the controller. In contrast to joint responsibility (see Sect. 3.1.1), the processor is therefore subordinate to the controller and has no discretion or decision-making leeway of its own (Martini, in: Paal and Pauly 2018, Art. 28 para. 2). In a departure from the old law, the processor may be directly exposed to claims for damages by data subjects under Art. 82(1) and (2) of the GDPR, and may be jointly and severally liable under Art. 82(4). In addition, he is liable to the controller for any subcontractors he uses under Art. 28(4) sentence 2 of the GDPR. Use cases of commissioned processing are, for example, cloud computing services, insofar as no access to the content of the operator's data is required, the outsourcing of backups, remote maintenance applications or external support (Bavarian State Office for Data Protection Supervision 2018, p. 1).

The ever-growing amount of data requires **intermediaries** such as platforms or data aggregators to facilitate the exchange. Under existing law, these are often classified as processors. In the course of the networking of production structures and the associated multilateral exchange of data, it is questionable whether the image of a subordination relationship underlying order processing will do justice to the practice in the same way in the future (Plattform Industrie 4.0 2016b, p. 14). The value chains associated with Industry 4.0 do not run in a linear fashion and therefore do not always provide for the processing companies to be dependent on instructions (see Plattform Industrie 4.0 2016b, p. 14). In addition, unlike the prototype of order processing (see Petri, in: Simitis 2019, Art. 28 para. 3), the recipients of data will regularly have a vested interest in handling the data. Apart from focusing on the mere possibility of the original controller to exert an influence, it is more expedient to clearly and transparently assign the data protection obligations to the new players, e.g. through a certification system (Plattform Industrie 4.0 2016b, p. 14). For Industry 4.0 applications, the institute of commissioned processing is nevertheless worthy of attention for the time being.

3.3 Rights of the Data Subject

Under Art. 4 No. 1 of the GDPR, data subjects are identified or identifiable natural persons to whom information relates. Arts. 15–22 of the GDPR give data subjects various rights through which they can influence the processing. First, a data subject has a **right of access** to whether and how (see the catalogue of Art. 15(1)(a-h)) data relating to him or her is processed, and the controller must provide a copy of the data (see Art. 15(1), (3) GDPR). They can demand the **rectification** and **completion** of incorrect data without delay (Art. 16 GDPR). In addition, under certain conditions, there is the right to **erasure** (so-called '**right to be forgotten**', Art. 17 GDPR) and the right to **restriction of processing** (Art. 18 GDPR). Under Art. 19 of the GDPR, the controller must also notify all recipients of this data of such changes. Under Art. 21 of the GDPR, the data subject may **object** to the processing of data for direct marketing or insofar as the processing is carried out under Art. 6(1)(f) of the GDPR. In the case of lit. f, the data subject must assert that he or she is in a special situation and there must be no compelling reasons worthy of protection on the part of the controller, such as ensuring data/IT security (Kamlah 2018, Art. 21 para. 5 f.).

Finally, Art. 20 of the GDPR contains something completely new with the **right to data portability**. It is intended to prevent so-called 'lock-in' effects, which occur when users do not switch providers because the effort involved is too high (Hornung 2012, p. 103). To this end, the data subject is given the right to receive the data concerning him or her in a structured, commonly used and machine-readable format and to transfer this data to another controller if the original processing was carried out with the aid of automated procedures and based on consent or a contract. The data must have been provided by the data subject, which may limit the scope of application against the background of Industry 4.0. In any case, data created by the data controller itself, such as an assessment of a person's state of health created by analysing raw data, should no longer be understood as 'provided' in this sense (Art.-29-Data Protection Working Party 2017, p. 9 f.). The borderline case is data that was not knowingly and willingly provided by the data subject, but was generated through the use of a service, such as traffic and location data or data that is generated when using smart devices. In particular, the purpose of the GDPR argues in favour of considering such data to be covered by Art. 20 of the GDPR, since its transfer precisely makes it more difficult to change providers (in detail Wrobel 2018, p. 247). For Industry 4.0 applications such as platforms, where the change of customers to competing offers is not excluded, it is advisable to establish a workflow for dealing with requests for data portability, e.g. by setting up an online tool.

3.4 Infringements

Infringements of the provisions of the GDPR can have different consequences. First, under Art. 82 of the GDPR, any person who has suffered material or immaterial damage has a claim for **compensation** against the controller or processor. They may be jointly and severally liable (Art. 82(4) GDPR). Under Art. 82(3), **fault is presumed**, so that the controller or processor must exculpate himself. Under Art. 79(2), the courts of the Member State in which the liable party has an establishment or the habitual residence of the data subject have jurisdiction. Whether **competitors** can also issue a warning notice for violations of data protection law depends on whether the provisions of the GDPR—as still advocated for the BDSG (KG, judgment of 22.9.2017 - 5 U 155/14 with further references)—are market conduct rules within the meaning of Section 3a UWG. This is not the case if the system of sanctions of the GDPR would have to be regarded as conclusive (see OLG Hamburg, decision dated 25.10.2018, 3 U 66/17, Köhler 2018, p. 1269 et seq.; a.A. Laoutoumai u. Hoppe 2018, p. 533 et seq.). Clarification by the Federal Court of Justice is to be expected here shortly.

Under Art. 83 of the GDPR, supervisory authorities may impose **administrative fines**. Considering the catalogue of Art. 83(2)(a-k) of the GDPR, a fine of up to €10 million or, in the case of companies (see the definition in Art. 4 No. 18 GDPR), 2% of the total worldwide annual turnover of the previous financial year may be imposed for violations of the provisions of para. 4; for violations of the provisions of para. 5, the fine may amount to up to €20 million or 4% of the annual turnover; the same applies to non-compliance with an instruction of the supervisory authority. The possible amount of fines has thus **increased drastically** compared to the old legal situation. While the supervisory authorities were initially cautious during the transition period to the new law, there are increasing reports of (higher) fines being imposed (e.g. recently €50 million fine for Google for non-transparent information imposed by the French data protection authority CNIL and €80,000 imposed by the State Commissioner for Data Protection and Freedom of Information of Baden-Württemberg due to insufficiently secured health data). Legal protection by the courts arises within the framework of the remedies granted by Section 41 (2) BDSG in application of the Administrative Offences Act (Gola, in: Gola 2018, Art. 83 para. 34). Finally, **criminal sanctions** under Section 43 BDSG, Sections 201a, 202a, 206 StGB may come into consideration.

4 Processing

Under Art. 6(1) of the GDPR, the processing of personal data is only lawful if it can be based on at least one of the legal bases of sub-sections a-f. The concept of processing is conceivably broad and includes any operation that is somehow related to personal data (Roßnagel, in: Simitis 2019, Art. 4 No. 2 para. 11). Consequently, data protection law can become relevant in the entire value chain; possible processing activities must be provided with a legal basis.

4.1 Contract Performance and Pre-Contractual Measures

Under Art. 6(1)(b) of the GDPR, processing is lawful if it is **necessary for the performance of a contract with the data subject**. This is the case if the processing has a direct factual connection to the contractual relationship (Wolff 2017, para. 540) and is necessary for the performance of obligations arising from the contractual relationship (Schulz, in: Gola 2018, Art. 6 para. 38). On the one hand, the processing must not be merely useful, but on the other hand, it does not have to be indispensable either (Buchner and Petri, in: Kühling and Buchner 2018, Art. 6 para. 42 ff.). Typically, this includes data such as name, address and payment data. The scope of the permissibility of data processing can thus be determined to a certain extent by the design of the content of the contract (see Schulz, in: Gola 2018, Art. 6 para. 27; on the distinction from consent see Wolff 2017, para. 543 et seq.).

4.2 Legitimate Interest

Alongside consent (see Sect. 4.3 below), Art. 6(1)(f) of the GDPR is the most relevant legal basis in practice. According to this, **processing necessary to protect the legitimate interests of the controller** is lawful unless the interests or fundamental rights of the data subject prevail. All legitimate interests, meaning legal, economic and non-material interests approved by the legal system must be considered (see Schulz, in: Gola 2018, Art. 6 para. 57). The processing must be necessary to protect such an interest. These interests must be balanced against the interests or fundamental rights of the data subjects. The **weighing up of individual cases** is based on factors such as the type and scope of the data, the context of the collection and the consequences of the processing (Schantz, in: Simitis 2019, Art. 6 para. 1 para. 105 et seq.). Generally, the greater the informative value of a piece of data about the data subject, the more serious the interference with his or her rights (cf. Schantz, in: Simitis 2019, Art. 6(1) para. 106).

The so-called reasonable expectations of the data subject must be explicitly considered in the weighing up process (see Recital 47, sentences 1 and 3). For example, networked devices equipped with sensors can collect data through their integration into the actions of data subjects which, when analysed, provide information about sensitive characteristics. This collection often occurs in the context of everyday actions that are irrelevant to the person and possibly in the imagined privacy of one's own home (cf. Schantz, in: Simitis 2019, Art. 6 (1) para. 121). These expectations then influence the consideration to the disadvantage of the controller. If the processing serves the maintenance of the product, such as in the case of versatile products or the determination of the demand for value-based services, the characteristics of the product are more important than the specific user. The use of pseudonymised data or data that is anonymised immediately after collection for such big data analysis protects the interests of the data subjects. This means that the data controller can use compensatory measures, such as an unconditional right of objection or additional information (cf. Schantz, in: Simitis 2019, Art. 6(1) para. 114), to influence the balance in his or her interest.

The flexibility of sub-section (f) as a legal basis offers the temptation to make use of it as extensively as possible. Although it is advantageous for the controller not to be dependent on the participation of the data subjects, the weighing up to be carried out entails a high degree of legal uncertainty. In any case, the controller must actually carry out and document the weighing up exercise, see Art. 5(1)(a) and (2) of the GDPR.

4.3 Consent

Apart from the legal justification, the consent of the data subject can also be used. In this way, circumstances that do not fit within the narrow legal parameters can be provided with a legal basis. Effective consent must be given voluntarily, for a specific case, in an informed manner and unmistakably in the form of an unambiguous confirmatory act (Art. 4 No. 11, Art. 7 GDPR). **Voluntariness** may be lacking especially if there is a particular imbalance between the data subject and the controller (Recital 43). However, it is not sufficient for a processing company to have a unique selling proposition on the market, for example through an innovative business idea (Schulz, in: Gola 2018, Art. 7 para. 22). The decisive factor is whether (reasonable) alternatives exist for the data subject and whether, if he or she refuses consent, he or she suffers a disadvantage that precludes a free decision (Wolff 2017, para. 503 et seq.).

The characteristics of **level of information and specific purpose** are problematic with regard to Industry 4.0 scenarios. Under Art. 4 No. 11, consent must be given in an informed manner and for the specific case (Art. 6 (1) lit a GDPR). The data subject should be able to understand the scope of the data processing to which he or she consents (Stemmer 2018, Art. 7 para. 74). Therefore, at the time of giving consent, i.e. before the processing, the controller, the data to be processed and the type of processing as well as the purposes of the processing (cf. Art. 6(1)(a) GDPR) must be known (Stemmer 2018, Art. 7 para. 75) and the data subject must be informed about these facts (Recital 42). Particularly in the case of big data applications, the specific purposes for which the data is to be used are not necessarily known when the data is collected. Those responsible are then faced with the task of drafting consent declarations that are as broad as possible, but at the same time sufficiently specific, which, depending on how they are structured in individual cases, cross the line, and then become non-specific (cf. Arning 2015, 10 f).

If effective consent was obtained, but for a different purpose, the permissibility of the **change of purpose** is based on Art. 6(4) of the GDPR. For this to be the case, the purposes must be compatible with each other. The connection between the purposes, the context of the collection, the type of personal data, the consequences of the intended further processing and the existence of appropriate safeguards must be considered, whereby the latter may explicitly include pseudonymisation. Subject to the risk in individual cases, compatibility may exist, for example, if self-learning systems or assistance systems process usage data to adapt to the preferences of the user (Roßnagel, in: Simitis 2019, Art. 6(4) para. 37).

In other application scenarios, such as in the area of value-based services or for the evaluation of usage-related data of delivered products, on the other hand, sufficiently specific and informed consent can usually be obtained for a determined purpose. In the area of order-driven production, it must be considered that consent must cover all processing operations throughout the entire production chain (BMWi 2016, p. 108). However, it should be noted that under Art. 7(3) sentence 1 of the GDPR, the data subject may **withdraw** consent at any time with effect for the future. Subsequently, the erasure of the data processed based on consent can be demanded, unless the controller can base the processing on another legal ground (see Art. 17(1)(b) GDPR). To avoid this risk and legal uncertainty, it is advisable to use legal facts as far as possible and to reduce consent to cases where no legal permission comes into question.

4.4 Employee Data Protection

The upheavals of digitalisation are also particularly affecting the world of work. The buzzword of the **smart factory**, a comprehensively networked factory, which is common in the context of Industry 4.0, concerns not only communication between machines, but also human-machine interaction (Krause 2017, p. 14). In addition to the use of computers and smartphones, which is already common today, future application scenarios include the use of wearables that data file operating processes, such as trackers, data glasses or intelligent gloves (Krause 2017, p. 14), or the support of employees through assistance systems such as exoskeletons. One advantage of these systems is that their use can be analysed and optimised through data collection. For example, an employee can be alerted if they make a mistake in a standardised process (see the example scenario by Hofmann 2016, p. 13). To be operated sensibly and securely, assistance systems require access control in the form of identity management, usually via personalised user accounts (Plattform Industrie 4.0 2015, p. 82). As a result, the data can regularly be assigned to individual employees and consequently fall under the scope of data protection law.

For the protection of employee data, Art. 88 of the GDPR gives the Member States room for manoeuvre, which the German legislator has developed with **Section 26 of the BDSG**. According to this, the processing of personal data of employees is only permissible if it is necessary for the decision on the establishment, implementation or termination of an employment relationship, Section 26(1) sentence 1 BDSG. To determine the necessity, the employer's interest in the processing and the employee's right to privacy must be weighed up, in a similar way to Art. 6(1)(f) of the GDPR (BT-Drs. 18/11325, p. 97). The principles of case law developed under the old law can be transferred (Nebel 2018, p. 523). In addition, processing can now—in deviation from the old legal situation—also take place based on **consent**. Under Section 26(2) sentence 3 BDSG, consent must be given in writing as a rule. The question of whether consent was given voluntarily must particularly consider the dependency existing in the employment relationship and the circumstances in which it was given (Section 26(2) sentence 1 BDSG), which is why consent will remain the exception in employee data protection. Only in cases where the employee gains an advantage or the interests of both parties are aligned, does voluntariness suggest itself, see Section 26 (2) sentence 2 BDSG. In all other cases, the type of data, the proximity to the employment relationship and the existence of an actual pressure situation must be considered.

Primarily, Industry 4.0 applications aim to optimise operational processes and are not intended to monitor the behaviour or performance of employees, which must be considered when determining necessity (Krause 2017, p. 33), but does not play a role in the fundamental applicability of data protection law. Due to the ubiquitous collection of data, this enables extensive monitoring of employees. In any case, a 'total surveillance' of employees is inadmissible (Kort 2018, p. 25); work behaviour may only be subject to selective data collection (Krause 2017, p. 34). Apart from this, the necessity depends on the design in the individual case (for individual constellations of Industry 4.0, see Kort 2018, p. 24; according data category Hofmann 2016, p. 12). The company's ability to function does not have to be affected; in the interest of technical innovations, significant efficiency gains are sufficient (Krause 2017, p. 34). For example, for the protection of employees in collaboration with robots, it is not necessary for these employees to be identifiable by the robot; for the efficient use of employees, it is not necessary for them to be located permanently, but only when necessary (Krause 2017, p. 34).

In addition to data protection law, the works council's right of co-determination under Section 87(1) No. 6 of the Works Council Constitution Act (BetrVG) is also involved as soon as a technical device is objectively suitable for monitoring performance or behaviour in a specific case (see BAG, AP No. 1 on Section 87 BetrVG 1972 monitoring). Accordingly, the introduction of Industry 4.0 systems will regularly be accompanied by a works agreement under Section 77 BetrVG (Hofmann 2016, p. 14). However, under Art. 88(1) of the GDPR and Section 26(4) sentence 1 of the BDSG, such a **collective agreement** can also be the legal basis for data processing if it does not fall below the level of protection of the GDPR (see Gräber and Nolden, in: Paal and Pauly 2018, Section 26 marginal no. 5). Collective agreements can therefore function as a means of shaping company-appropriate regulations, especially in areas where the legal regulations remain abstract (see Nebel 2018, p. 524). In the practice of larger companies, collective agreements are—precisely because of the very limited use of consent in employee data protection—the standard case for achieving permissibility under data protection law.

4.5 Special Categories of Personal Data

In principle, the GDPR treats all data in the same way. However, differentiation is made in the processing of so-called special categories of personal data (Art. 9 GDPR) and data on criminal convictions or offences (Art. 10 GDPR), which may only be processed under official supervision. The former special categories particularly include data revealing political opinions or religious or ideological beliefs, as well as genetic and biometric data and health data. Whenever the processing of these data is involved in an Industry 4.0 application, special caution is required, as their processing is generally prohibited. The exceptions of Art. 9(2) of the GDPR may then have to be examined.

4.6 Transfers to Third Countries

The networking associated with the digital transformation does not stop at national borders—a secure legal framework for third country transfers is necessary to prevent businesses located in the area covered by the GDPR from being cut off from global supply chains (Plattform Industrie 4.0 2016b, p. 14). The permissible transfer of personal data to third countries is governed by Art. 44 et seq. of the GDPR. What is required is an adequacy decision by the European Commission with the third country in question (Art. 45 GDPR), as most recently reciprocated with Japan on 23.1.2019, or the existence of appropriate safeguards under Art. 46 of the GDPR, such as binding corporate rules (see Art. 47) or **standard contractual clauses**

('SCCs'). The use of SCCs has proven itself in practice, as they can be used in an uncomplicated manner. They can be adopted unchanged and do not require any additional approval by or notification to supervisory authorities (Schantz, in: Simitis 2019, Art. 46 para. 31). Under Art. 46(5) sentence 2 of the GDPR, SCCs concluded prior to the applicability of the GDPR continue to apply (there are two versions for transfers to third countries to controllers and one version for transfers to third countries to processors).

4.7 Codes of Conduct

Art. 40 of the GDPR allows business or industry associations representing several controllers to develop codes of conduct. In this way, the application of the GDPR can be specified for certain areas, for example with regard to legitimate interests of controllers, pseudonymisation of personal data, information of data subjects and the exercise of data subjects' rights (Art. 40(2) GDPR). Even if comparable options have only rarely been used in the past (twice in Germany, see Roßnagel, in: Simitis 2019, Art. 40 para. 4), this offers the opportunity for individual sectors with Industry 4.0 applications to concretise the very abstract regulations of the GDPR in the sector-specific interest, to achieve a higher degree of legal certainty and to reduce the likelihood of legal disputes (Bitkom 2016, p. 10; Paal, in: Paal and Pauly 2018, Art. 40 para. 3). The rules of conduct are submitted to the supervisory authority and enforced by own supervisory bodies under Art. 41 of the GDPR (so-called regulated self-regulation). Here, the associations are called upon to take action.

4.8 Purpose Limitation and Data Minimisation

The principle of **purpose limitation** guaranteed by Art. 8(2) sentence 1 of the CFR and laid down in Art. 5(1)(b) of the GDPR is particularly relevant for Industry 4.0 applications. According to this principle, data must be collected for specified, unambiguous purposes and may not be further processed in a way that is incompatible with these purposes. The data subject therefore has the certainty that data collected from him or her will only be used for the purpose that was originally specified. This principle is somewhat at odds with Industry 4.0. Application scenarios, such as the creation of value-based services or transparent and adaptable products, require the flexible processing of extensive data sets (Plattform Industrie 4.0 2016b, p. 12). Even if processing can be justified by consent for several purposes, the principle of purpose limitation requires that the respective processing purposes are already known and defined at the time of collection. This is difficult to achieve in the face of constantly new demands and requirements. For example, in big data analysis, data that was originally processed for another purpose is torn out of its existing context (see Helbing 2015, p. 150) and processed in a new context for a new purpose. Such secondary processing is only possible under narrow conditions (see Sect. 4.3). In any case, the purposes of processing data must already be considered as far ahead as possible when developing business models.

A similar tension exists with the principle of **data minimisation** from Art. 5(1)(c) of the GDPR. It provides that data may only be processed to the extent necessary to achieve the purpose (Roßnagel, in: Simitis 2019, Art. 5 para. 116). In contrast, Industry 4.0 processes are based on an increasing wealth of data, in that more data can be collected through sensor technology and stored longer and to a greater extent through more favourable storage options. The larger the treasure trove of data, the more extensive big data analysis can be carried out and the more meaningful the results. There is therefore an incentive for responsible parties to collect as much data as possible, which conflicts with the principle of data minimisation. Although the conflict can be circumvented by anonymising the data, not all business models can manage without personal data, and there is also an increasing risk of re-identifiability (see Sect. 2.1.3).

These tensions can therefore only be resolved to a limited extent. On purpose limitation, the European legislator has listed the criteria by which the compatibility of purposes is measured (see Art. 6(4)), but there is no clear basis for the final assessment (see Roßnagel et al. 2016, p. 159). Responsibility lies with the individual data processor; however, the existing legal uncertainty prevents already collected data from being used for purposes arising from new business models. In the interest of innovation, it is necessary to wait for the European Data Protection Board to specify how this is to be handled.

References

Albrecht JP, Jotzo F (2017) Das neue Datenschutzrecht der EU, 1. Aufl., Baden-Baden

Arning MA (2015) Datenpools – Big Data datenschutzkonform umsetzen, K&R Beilage Nr. 3, 7 ff

Artikel-29-Datenschutzgruppe (2007) Stellungnahme 4/2007 zum Begriff "personenbezogene Daten", 01248/07/DE, WP 136, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_de.pdf

Artikel-29-Datenschutzgruppe (2010) Stellungnahme 8/2010 zum anwendbaren Recht, 0836-02/10/DE, WP 179, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2010/wp179_de.pdf

Artikel-29-Datenschutzgruppe (2014) Stellungnahme 5/2014 zu Anonymisierungstechniken, 0829/14/DE, WP 216, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

Artikel-29-Datenschutzgruppe (2017) Leitlinien zum Recht auf Datenübertragbarkeit, 16/DE, WP 242 rev.01, http://ec.europa.eu/newsroom/document.cfm?doc_id=44099

Artikel-29-Datenschutzgruppe (2018) Leitlinien für Transparenz gemäß der Verordnung 2016/679, 17/DE, WP 260 rev.01, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=51025

Bayerisches Landesamt für Datenschutzaufsicht (2018) FAQ zur DSGVO, https://www.lda.bayern.de/media/FAQ_Abgrenzung_Auftragsverarbeitung.pdf

Bitkom e.V. (2016) Rechtliche Rahmenbedingungen von Industrie 4.0 – Stellungnahme, http://www.bitkom.org/sites/default/files/file/import/Bitkom-Position-Rechtliche-Rahmenbedingungen-Industrie40.pdf
Bundesministerium für Wirtschaft und Energie (Hrsg.) IT-Sicherheit für die Industrie 4.0 – Produktion, Produkte, Dienste von morgen im Zeichen globalisierter Wertschöpfungsketten, https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/it-sicherheit-fuer-industrie-4-0.pdf?__blob=publicationFile&v=4 (zit.: BMWi 2016)
Düsseldorfer Kreis – Aufsichtsbehörden für den Datenschutz im nicht-öffentlichen Bereich (2015) Orientierungshilfe zu den Datenschutzanforderungen an Smart-TV-Dienste, https://www.datenschutzkonferenz-online.de/media/oh/20150917_oh_smart_tv.pdf
Frenz W (2016) Industrie 4.0 und Datenschutz im fairen Wettbewerb, EuZW, 121 f
Gola P (Hrsg.) (2018) Datenschutz-Grundverordnung, 2. Aufl., München (zit.: Bearbeiter, in: Gola 2018)
Helbing T (2015) Big Data und der datenschutzrechtliche Grundsatz der Zweckbindung, K&R, 145 ff
Hofmann K (2016) Smart Factory – Arbeitnehmerdatenschutz in der Industrie 4.0 – Datenschutzrechtliche Besonderheiten und Herausforderungen, ZD, 12 ff
Hornung G (2012) Eine Datenschutz-Grundverordnung für Europa? – Licht und Schatten im Kommissionsentwurf vom 25.1.2012, ZD, 99 ff
Hornung G (Hrsg.) (2018) Rechtsfragen der Industrie 4.0, Datenhoheit – Verantwortlichkeit – rechtliche Grenzen der Vernetzung, Baden-Baden (zit. als Autor, in: Hornung 2018)
Kamlah W (2018) In: Plath K-U (Hrsg.) DSGVO/BDSG, 3. Aufl., Köln
Köhler H (2018) Durchsetzung der DS-GVO mittels UWG und UklaG?, WRP, 1269 ff
Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder (2018) Zur Anwendbarkeit des TMG für nicht-öffentliche Stellen ab dem 25. Mai 2018, https://www.bfdi.bund.de/SharedDocs/Publikationen/Entschliessungssammlung/DSBundLaender/95DSK_Positionsbestimmung_TMG.pdf?__blob=publicationFile&v=4
Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder und Verband der Automobilindustrie (2016) Datenschutzrechtliche Aspekte bei der Nutzung vernetzter und nicht vernetzter Kraftfahrzeuge, https://www.vda.de/dam/vda/Medien/DE/Themen/Innovation-und-Technik/Vernetzung/Gemeinsame-Erkl-rung-VDA-und-Datenschutzbeh-rden-2016/Gemeinsame-Erklaerung-VDA-und-Datenschutzbehoerden-2016.pdf
Kopp R, Sokoll K (2015) Wearables am Arbeitsplatz – Einfallstore für Alltagsüberwachung?, NZA, 1352 ff
Kort M (2018) Neuer Beschäftigtendatenschutz und Industrie 4.0 – Grenzen einer "Rundumüberwachung" angesichts der Rechtsprechung, der DSGVO und des BDSG nF, RdA, 24 ff
Krause R (2017) Expertise – Digitalisierung und Beschäftigtendatenschutz, https://www.bmas.de/SharedDocs/Downloads/DE/PDF-Publikationen/Forschungsberichte/fb482-digitalisierung-und-beschaeftigtendatenschutz.pdf?__blob=publicationFile&v=1
Kühling J (2017) Neues Bundesdatenschutzgesetz – Anpassungsbedarf bei Unternehmen, NJW, 1985 ff
Kühling J, Buchner B (Hrsg.) (2018) Datenschutz-Grundverordnung/BDSG, 2. Aufl., München (zit.: Bearbeiter, in: Kühling/Buchner 2018)
Laoutoumai S, Hoppe A (2018) Setzt die DSGVO das UWG Schachmatt?, K&R, 533 ff
Laue P, Nink J, Kremer S (2016) Das neue Datenschutzrecht in der betrieblichen Praxis, 1. Aufl., Baden-Baden
Lüdemann V, Ortmann MC, Pokrant P (2016) Datenschutz beim Smart Metering – Das geplante Messstellenbetriebsgesetz (MsbG) auf dem Prüfstand, RDV, 125 ff
Mantz R (2018) In: Sydow G (Hrsg.) Europäische Datenschutzgrundverordnung, 2. Aufl., Baden-Baden
Moos F (2015) "Geht nicht" gibt es nicht: Datennutzung als rechtliche Gestaltungsaufgabe, K&R Beilage, 12 ff

Nebel M (2018) Big Data und Datenschutz in der Arbeitswelt – Risiken der Digitalisierung und Abhilfemöglichkeiten, ZD, 520 ff
Paal BP, Pauly DA (Hrsg.) (2018) Datenschutz-Grundverordnung – Bundesdatenschutzgesetz, 2. Aufl., München (zit.: Bearbeiter, in: Paal/Pauly 2018)
Plattform Industrie 4.0 (2015) Umsetzungsstrategie Industrie 4.0 – Ergebnisbericht der Plattform Industrie 4.0, https://www.bitkom.org/Bitkom/Publikationen/Publikation_5575.html
Plattform Industrie 4.0 (2016a) In: Bundesministerium für Wirtschaft und Energie (Hrsg.) Fortschreibung der Anwendungsszenarien der Plattform Industrie 4.0, https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/fortschreibung-anwendungsszenarien.pdf?__blob=publicationFile&v=7
Plattform Industrie 4.0 (2016b) In: Bundesministerium für Wirtschaft und Energie (Hrsg.) Ergebnispapier Industrie 4.0 – wie das Recht Schritt hält, https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/i40-wie-das-recht-schritt-haelt.pdf;jsessionid=B8394BB485838445387890601A799761?__blob=publicationFile&v=8
Richter P (2012) Datenschutz durch Technik und die Grundverordnung der EU-Kommission, DuD, 576 ff
Roßnagel A (2018) Pseudonymisierung personenbezogener Daten – Ein zentrales Instrument im Datenschutz nach der DS-GVO, ZD, 243 ff
Roßnagel A, Geminn C, Jandt S, Richter P (2016) Datenschutzrecht 2016 – "Smart" genug für die Zukunft? Ubiquitous Computing und Big Data als Herausforderungen des Datenschutzrechts, Kassel, http://www.uni-kassel.de/upress/online/OpenAccess/978-3-7376-0154-2.OpenAccess.pdf
Schaar K (2016) DS-GVO: Geänderte Vorgaben für die Wissenschaft – Was sind die neuen Rahmenbedingungen und welche Fragen bleiben offen?, ZD, 224 ff
Schmitz P (2018) In: Spindler G, Schmitz P, Liesching M (Hrsg.) Telemediengesetz mit Netzwerkdurchsetzungsgesetz, 2. Aufl., München
Simitis S, Hornung G, Spiecker gen. Döhmann I (Hrsg.) (2019) Datenschutzrecht – DSGVO mit BDSG, 1. Aufl., Baden-Baden (zit.: Bearbeiter, in: Simitis 2019)
Stemmer B (2018) In: Wolff HA, Brink S (Hrsg.) BeckOK Datenschutzrecht, 26. Edition, München
Wiebe A (2016) Protection of industrial data – a new property right for the digital economy?, GRUR Int., 877 ff
Wiebe A (2017) Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken, GRUR, 338 ff
Wolff HA (2017) In: Schantz P, Wolff HA, Das neue Datenschutzrecht – Datenschutz-Grundverordnung und Bundesdatenschutzgesetz in der Praxis, 1. Aufl., München
Wrobel M (2018) Datenportabilität – Demokratisierung der digitalen Wirtschaft?, DSRITB, 247 ff

Challenges in Data Protection in Business Transactions: An Overview Over Civil Law Approaches to Data Protection and Data Ownership

Indra Spiecker gen. Döhmann and Alexander Benecke

Contents

1 Introduction

Within the context of Industry 4.0, data is commonly referred to as “Oil of the digital economy” (Riehm 2018, p. 73). This metaphor stresses the idea that data—just like oil—serves as both a commodity and a “lubricant” for business cooperation (Riehm 2018, p. 73; Kling-Straub and Straub 2018, p. 3201). As compelling as this comparison may sound at first glance, it does not consider that data—unlike oil—is not a scarce good (Badmann 2019, p. 578 ff.; Drexl 2017, p. 339 f.; different Wiebe 2017, p. 887). In addition, this view disguises that both oil and data contain

I. Spiecker gen. Döhmann (✉) · A. Benecke
Faculty of Law, University Frankfurt, Frankfurt, Germany
e-mail: spiecker@jur.uni-frankfurt.de; benecke@jur.uni-frankfurt.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_9

heavy risks—despite their undeniable benefits (Hirsch 2014, p. 374 f.; on the meaning of Danger, Risk, and Uncertainty, Spiecker gen. Döhmann 2021).

Data processing as such goes along with risks for both individuals and businesses. Within the context of Industry 4.0, those risks typically contain possible violations of, on the one hand, each individuals' general right to protection of personality, and, on the other hand, unwanted publicity of industrial knowledge and business secrets. Clearly, industrial knowledge not only gains but also keeps its value by remaining a secret to potential competitors and other outsiders. Therefore, businesses prefer strict technical protection to prevent unwanted access from third parties. However, technical protection *alone* proves to be insufficient for two reasons. First, once technical barriers are overcome, there is no second line of protection. If exclusivity is broken and information becomes free flowing, an unlimited and uncontrollable amount of competitors and other outsiders can take both notice and advantage of this information. There is no factual "getting it back" and although there is a "right to be forgotten" to some extent, e.g., Art. 17 GDPR, this right can hardly be enforced effectively when it comes to information that almost solely holds its value by remaining in the dark (Riehm 2018, p. 74). Second, technical protection does not offer any solution to the problem of voluntarily disclosing data for a specific purpose, which the processor then uses for other purposes—and neither for situations where the state forces disclosure and then third parties require access of this information from the state.

These risks are neither revolutionary new nor the sole challenges of Industry 4.0 applications. However, data processing in networks that is characteristic for Industry 4.0 pushes these risks further due to the increased amount of generated data, to new ways of evaluating, and to generating specific information from them (concerning the difference between data and information Deißler 2018, p. 34 ff.). Ideally, legal provisions can protect against this inherent vulnerability by subsequently regulating both how to handle illegally processed data and providing companies with legal tools of asserting specific counterclaims.

Both the character of Industry 4.0 legal provisions as a fluid subject matter and the variety of—partly even interlaced and sometimes counteracting—parties involved in Industry 4.0 push legal limitations. Businesses, competitors, workers, unions, intermediaries, and customers each have their own interests regarding collecting, withholding, and using "their" data (Specht 2017, p. 1040, 1041). Hitherto, legal regulation can only catch up with these conflicts to a rather insufficient extent as there is no coherent legal framework for Industry 4.0 but rather a fragmented legal ragbag. Legal protection of personal and business data is a cross-sectional framework comprised of criminal, civil, and public law. Additionally, Industry 4.0 applications have their own legal implications and dynamics. Rapidly paced innovations within Industry 4.0 constantly expose loopholes within the rather ponderous legal framework. Questions about who owns data within B2C and B2B-relations and concerns about liability and responsibility for autonomous systems—only to name a few—remain unanswered by the law (concerning the problem of systemic digitalization Spiecker gen. Döhmann 2016a, p. 698 ff.). This chapter addresses some of the most relevant and common issues within Industry 4.0 when it comes to data protection regulation, namely the questions of how to incorporate data protection in business transactions, and—based on that—whether businesses can in fact have a "right to data."

2 Industry 4.0: Terminology and Challenges

Industry 4.0 is a German term (Müllmann 2018, p. 1177, 1178) and inspired by the common differentiation between certain phases of the era of industrial revolution (Mertens et al. 2017, p. 46; Staffler 2018, p. 269 f.; Mühlich 2014, p. 381). While all previous phases were identified as such in hindsight, Industry 4.0 was proclaimed anticipatory (implicit Staffler 2018, p. 269). Although there is some sort of common sense that Industry 4.0 does in fact describe somewhat of a revolution (Bräutigam and Klindt 2015, p. 1137; Mühlich 2014, p. 381; Leimster 2015, p. 176), there is neither an intra- nor interdisciplinary understanding one can consider to be commonly shared (Mertens et al. 2017, p. 46; Müllmann 2018, p. 1178). Hence, some refer to it as a "pseudo-scientific" term (Wilhelm 2015, p. 187; similar Bräutigam and Klindt 2015, p. 137) and some refer to it as a "hype cycle" (Hirsch-Kreinsen 2016, p. 24 f.).

Leaving the many attempts of finding a comprehensive definition aside, this chapter does not intend to make another approach. We will rather concentrate on different technical procedures and structures that challenge data protection in latest data processing situations between companies, such as the use of cyber-physical systems (CPS), autonomously acting artificial intelligence (AI), embedded data usage, and horizontal as well as vertical networks within the value chain which on their very own allow for decentralized production (Staffler 2018, p. 269, 270). These factors accumulate to an ideal picture of a "smart factory" (Hofmann 2016, p. 12 f.; Thalhofer 2017, p. 225).

Many of these rather new terms seem to imply small steps of a real revolution. However, we should not mistake new terms for actually new technical procedures and/or the cause of yet unknown problems. Instead, some manufacturing structures used within the smart factory as such have been used and known for quite a while (Mertens et al. 2017, p. 47; Müllmann 2018, p. 1177, 1178; Mühlich 2014, p. 381). Hence, legal challenges arise much more from network-based intensification of automation, autonomization, customer-focus, and personalization. Services generated by user data ("Smart Services"), for example, can only exist through usage of "Smart Products" (such as Internet-connected refrigerators or production machines). These services help with collecting and processing personal data to provide specific services and are sometimes even a prerequisite for the service to exist.

3 Principles: Data Protection Regulation of Private Relationships

Data Protection Law is one of the most important regulatory tools concerning data-driven applications within Industry 4.0. Main provisions and general principles are laid out in the relatively new (25.05.2018) General Data Protection Regulation (GDPR). Although the GDPR is a directive in terms of Article 288 TFEU, it does contain a multitude of "opening clauses" (Benecke and Wagner 2016, p. 600 ff.), which allow for Member States to pass each their own new legal frameworks and regulations or keep their already existing ones at least to a certain extent (Simitis et al. 2019, Introduction No. 206).

Data Protection Law aims to regulate risks and benefits of automatized data processing. One of its main purposes is to prevent power asymmetries and misinterpretations that arise from one-sided informational preponderance and that can lead to economic, social, personal or political discrimination of individuals (Simitis et al. 2019, Introduction No. 10). Due to Data Protection Law having its historical basis in the General Right to Protection of Personality (concerning German legal tradition BVerfGE 65, p. 141), which both the EU Data Protection Directive of 1995 and the EU General Data Protection Regulation of 2018 are also based on, it nevertheless evolved factually as a broad legal framework of how to deal with information on national nor on European level in general, although it was originally a specific regulatory regime that concentrated on mediating negative impacts of automatized data processing for individuals. In that sense, Data Protection Law is technology law and therefore bound to implement preventive measures.[1] Hence, one of the main principles of Data Protection Law is that there must not be processing of data without a legal basis allowing for that specific processing—"principle of prohibition with reservation of permission" (Simitis et al. 2019, Introduction No. 236). However, evolving from this regulatory approach, Data Protection Law has become another set piece of Economic Regulatory Law (Simitis et al. 2019, Introduction No. 47). Data Protection Law is now also a subset of Consumer Protection Law (Weichert 2001, p. 264), which becomes apparent when looking at the representation of data subjects in Art. 80 GDPR and the Right to data portability in Article 20 GDPR (Simitis et al. 2019, Article 20 Rn. 1) which are both not original concerns of personality protection.

The main goal of Data Protection Law is to ensure protection of *personal* data. However, significant problems arise on the extent of that protection. Despite a common sense (Simitis et al. 2019, Article 4 Nr. 1 No. 3 and 66), which includes the European Court of Justice (ECJ C-582/14, NVwZ 2017, p. 213—Breyer), of the term Personal Data as purely factual information, a huge portion of information especially in the business sector remains uncovered by Data Protection Law. On the

[1] *Simitis/Hornung/Spiecker gen. Döhmann*, in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, 2019, Einl. Rn. 17; *Hornung/Spiecker gen. Döhmann*, in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, 2019, Art. 1 Rn. 8.

other hand, purely factual information could well be included in the regulatory regime if mixed with personal information. The core requirements for Data Protection Law concerning information are laid out in Article 4 (1) GDPR (personal data), whereby it does not matter whether the processor cannot distinguish between personal and non-personal data or whether the collected data was gathered randomly or unintentionally (ECJ C-131/12, NJW 2014, p. 2257—Google Spain). Hence, even if Data Protection Law may not apply to certain Industry 4.0 applications at first glance, in most cases it actually will be applicable because there will be a connection of relevant information to specific individuals such as employees or customers. This should make Industry 4.0 applications more aware of the content of the GDPR.

The GDPR governs processing of data exclusively; the provisions are final to the extent that no other provision will be applicable to this processing (the exception being Article 6 (1) (c) and (e) GDPR). Thus, the legal situation concerning processing of personal data is mostly uniform all over Europe, with—inter alia— the in the context of Industry 4.0 notable exception being employee data protection, which Member States can regulate themselves due to the opening clause in Article 88 GDPR.

While at the very outset, one of the main goals of Data Protection Law was to countervail power asymmetries between the state and the people, the making of the GDPR was also influenced by the risks and dangers that arise from data processing between private players and the informational asymmetry possible there.[2] Due to it being a cross-sectional matter, Data Protection Law covers a variety of data processing in different situations typical for Industry 4.0. The following chapters will cover some of the most challenging ones.

4 Challenges for the Protection of Data and Information Between Companies

4.1 Business Transactions

Challenges for Industry 4.0 companies concerning data protection typically are a result of customary business transactions. These challenges foremost arise when it comes to handling personal data of customers, e.g., during due diligence[3] and the execution of specific acquisitions. As the GDPR imposes severe penalties and

[2] Concerning the third-party effect of Artt. 7, 8 CFR *Marsch*, Das europäische Datenschutzgrundrecht, 2018, p. 247 ff.; concerning the genesis *Albrecht*, in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, 2019, Einl. Rn. 185; *Hornung/Spiecker gen. Döhmann*, in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, 2019, Einl. Rn. 208.

[3] The actual execution of the transaction is generally preceded by a due diligence review, which is intended to enable potential buyers to make a "careful" purchase decision. For this purpose, the seller provides the acquirer with information about the company to be sold. *Plath*, in: Bussche v.d./ Voigt, Konzerndatenschutz, 2014, Teil. 6, Rn. 5.

considering how easy it is to violate them during complex transactions (Göpfert and Meyer 2011, p. 486), complying with the GDPR's legal provisions is of utmost importance (Grünwald and Hackl 2017, p. 556 ff.), also in building trust between partners and in relation to customers and employees.

Business transactions can take many different shapes, from mergers and acquisitions to sales and divestitures. Concerning their practical relevance, however, there are typically *two* forms of acquisitions that lead to problems with Data Protection Law.

Share Deal: Share deal means sale of the shares of a company, so a legal purchase of preventive nature and a transfer of shares by way of assignment. The share deal leads to a buyer's universal succession, i.e., a complete transition of existing rights and obligations, particularly fiscal nature (Beyer and Beyer 2016, p. 241).

Asset Deal: An asset deal occurs when a buyer is interested in purchasing the operating assets of a business instead of stock shares. The buyer completes the transaction by providing the selling company consideration for some or all of the assets they own (Schröder 2017, Rn. 36).

4.1.1 Transfer of data Within Share Deals

Usually, a share deal will not result in a change of the legal person in charge and legally responsible, the so-called controller (Article 4 (7) GDPR). The company itself remains unchanged and, therefore, from a legal perspective, it is the same controller processing the relevant personal data (Baranowski and Glaßl 2017, p. 199, 201). Hence, neither do the legal responsibilities laid out by Data Protection Law change (Simitis et al. 2019, Article 6 (1) No. 126; Platz 2014, No. 64 f.; Göpfert and Meyer 2011, p. 486, 490). As there is no transfer of personal data to a third party (Schröder 2017, No. 36; Härting 2017, p. 724, 725), there is no "operation" by definition of Article 4 (7) GDPR. Thus, the legal requirements for the lawfulness of processing laid out by Article 6 GDPR are inapplicable for any transfer of data in the course of this action (Nebel 2016, p. 417, 418; Härting 2017, p. 724, 725).

When there is a corporate group involved, however, problems arise during the process of incorporation. The GDPR does not contain an intra group exemption (Simitis et al. 2019, Article 6 (1) No. 116). From the perspective of Data Protection Law, businesses within a corporate group remain independent and thus remain independently responsible controllers (Gola 2018, Article 6 No. 170). As a consequence, any transfer of customer data from holding companies to sister companies requires its own legal basis (Simitis et al. 2019, Article 6 (1) No. 116; Schröder 2017, No. 36) and needs to follow the procedural requirements of the GDPR. A fitting legal basis for intra-group transfer of data is Article 6 (1) (f) GDPR according to which processing is lawful when it is necessary for the purposes of the *legitimate* interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data. Concerning the legitimacy of interests, recital 48 GDPR states that "controllers that are part of a group of undertakings or institutions affiliated to a central body may have a legitimate interest in transmitting personal data within the group of undertakings for internal administrative purposes, including the processing of clients' or employees' personal data." Hence, the transfer of employees' data and customers' data within a corporate group will be legitimate *by default* (Gola 2018, Article 6 No. 170 f.; conservative Schröder 2017, No. 38). However, this is only a refutable legal presumption, which requires case-by-case decision. In particular, the processor must consider the data subjects' expectations at the time of data-collection since the new group context can result in different purposes of data processing unforeseeable for the data subject at the time of data collection (Simitis et al. 2019, Article 6 (1) No. 126).

One important asset of companies are their customer data sets. Disclosure of any customer data (Plath 2015, No. 45 ff.) during contract initiation and due diligence requires a legal basis, since that time of transaction the acquiring company is a "third party" in the sense of Article 4 (10) GDPR and disclosure is "processing" by "making available otherwise" as defined by Article 4 (2) GDPR. The legal basis for this proceeding is Article 6 (1) (f) GDPR. On the question about who is to be attributed as the "processor" within a share deal, the following rules apply: The company holding shares in the company to be sold is the seller. However, it is the company to be sold itself that generated the data and disclosed it during due diligence. Hence, the target company is "processor" by means of Article 4 (7) GDPR and, thus, the question rises whether the target company protects *their own* "legitimate interests" by disclosing the data (Plath 2014, No. 26 ff.; Göpfert and Meyer 2011, p. 486, 490). Although both target company and mother company will have corresponding interests, under Article 6 (1) (f) GDPR, processing is also lawful when it is necessary for the purposes of the legitimate interests pursued by a *third party*, which, in this case, could also be the holding company (Göpfert and Meyer 2011, p. 486, 490). Eventually, from a legal standpoint, there are relatively few legal obstacles when it comes to processing of data within a share deal on the grounds of legitimation. However, all parties involved should take care of additional safeguards and the procedural requirements, such as duties to inform (Art. 12 et. Seq. GDPR), potential risk assessment and also protection against fiduciary breaches.

4.1.2 Transfer of Data Within Asset Deals

On asset deals, one has to differentiate. There are rather unproblematic cases, e.g., when the acquiring company intends to (at least to some extent) continue the seller's business. In this case, the acquiring company typically takes over the rights and obligations in relation to the customer through a three-sided contract between seller, customer, and acquiring company. Although the fact that the customer consents to the transfer of contract does not automatically mean he or she also consents to the transfer of "his" data,[4] the parties involved can in fact make such an agreement within the same contract (Baranowski and Glaßl 2017, p. 199, 201). In that case, Article 6 (1) (a) GDPR covers the transfer of data. If the data subject does not consent, Article 6 (1) (b) GDPR is the legal basis for the transfer of data (Härting 2017, p. 724, 727), since the processing "is necessary for the performance" of the contract.[5]

On the other hand, there are much more complicated asset deals (especially when the seller is about to become insolvent) that involve isolated transfer of customer data or of old data from contracts that already expired (concerning the differentiation Schröder 2017, No. 38; concerning transfer of data from ongoing contracts Nebel 2016, p. 417, 420). Article 6 (1) (f) GDPR covers both cases if the transfer of data is necessary for the purposes of the legitimate interests pursued by the seller and such interests are not overridden by the interests or fundamental rights and freedoms of the data subject.

If customer data is sold in an isolated way via asset deal, the seller's sole interest lies within the fulfillment of his contractual obligations to transfer the relevant data to the acquiring company (Simitis et al. 2019, Article 6 (1) No. 127). However, a customer will not have any obvious interest in a third party processing her data from now on. From the customer's perspective, there is an exchange of the responsible entity by another controller processing his or her data without resulting in any relevant benefit for the customer per se (Schröder 2017, No. 38). It is even worse for the customer since – at least according to the principles of the GDPR—the customer will regularly have a substantial interest in the person of the processor (Simitis et al. 2019, Article 6 (1) No. 127; differing view Nebel 2016, p. 417, 421 f.) particularly if there has been an ongoing relationship in which trust has been established. This does not automatically apply to the successor and even less so if the successor's reputation concerning data protection is doubtful. It goes without saying that lawfulness of an isolated transfer of data is a matter of individual consideration in the sense of Article 6 (1) (f) GDPR. However, after all, the sellers' interests will only outweigh the customers' interests if he can claim other interests than the contractual obligation of data transfer as such.

Concerning old data from expired contracts, the legal basis for their transfer can only be Article 6 (1) (f) GDPR. Again, this requires a weighing of interests between the sellers' and the customers' interests in the transfer of data. In this case, however, the consideration will turn out in favor of the controller. If there is no consumer data involved, a major argument in favor of the controller is his or her interest in creating a basis for the buyer to continue the business as unconditionally as possible

[4]Differently *Schröder*, in: Forgó/Helfrich/Schneider, Betrieblicher Datenschutz, 2. ed. 2017, ch. 4 Rn. 37 and *Nebel*, CR 2016, 417 (418), who assume that the consumer's wish to transfer the contractual relationship indicates regularly (at least implicitly) consent to the transfer of data or that it is not rejected.

[5]*Plath*, in: Bussche v.d./Voigt, Konzerndatenschutz, 2014, part 6, Rn. 68; in result also *Thode*, PinG 2016, 26 (28); *Härting*, CR 2017, 724 (725 f.); *Nebel*, CR 2016, 417; *Schröder*, in: Forgó/ Helfrich/ Schneider, Betrieblicher Datenschutz, 2. Aufl. 2017, ch. 4 Rn. 37.

(Schröder 2017, No. 38). If the asset deal involves consumer data, common sense requires the consumer to have reasonable expectations about the goods and services offered, e.g., the buyer knowing previous purchases made by the customer just so he can give him specific advice or offer specific products that meet his requirements and expectations (Schröder 2017, No. 38). Alternatively, the seller can grant the consumer a temporary and unconditional right to object. If the consumer does not object in time this very fact—together with further indicators—could speak in favor of the sellers' interests.[6]

As already discussed above in conjunction with the Share Deal, Article 6 (1) (f) GDPR can also serve as a legal basis for transferring of personal data during due diligence within an asset deal. In this case, however, the seller pursues his very own interests, e.g., downsizing or repositioning his business actions. It should be noted that particularly in the preparatory steps of a business transfer such as an asset deal, the seller need not make available all data to the seller but rather restrict access to selected files. Additional protective tools for due diligence are pseudonymization and anonymization of the relevant data (Simitis et al. 2019, Article 6 para. 1 No. 127).

4.1.3 Summary and Relevance

Just like under the former Data Protection Directive and its national laws, there are typically two main legal challenges for business transactions. While share deals are rather unproblematic, problems arise when it comes to asset deals. Fortunately, the GDPR has not changed much concerning business within the Industry 4.0. Article 6 (1) (f) GDPR now contains a *central* provision for the lawfulness of data processing. Considering the complexity of business transactions, the overwhelming amount of data involved and drastic sanctions for violations of provisions laid out by the GDPR, companies should be cautious about each transfer of / transferring personal data. They should make sure they follow the rules of the GDPR including procedural, organizational, and technological safeguards, particularly the steps before the finalization of the deal, e.g., in due diligence settings.

4.2 Whom Does Data "Belong To"? Status Quo and Perspectives (German Law)

Most authors refer to "data protection" as defensive and protective law in the sense of individual powers of control over data, e.g., claims of data subjects for

[6] *BayLDA*, 7. TB 2015/2016, 74 f.; *Baranowski/Glaßl*, BB 2017, 199 (202); *Schröder*, in: Forgó/Helfrich/Schneider, Betrieblicher Datenschutz, 2. Aufl. 2017, ch. 4 Rn. 38; *Schantz*, in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, 2019, Art. 6 Abs. 1 Rn. 127.

information, omission, and deletion (Article 15, 17 GDPR). The for a long time sufficient, predominantly non-commercial access to the assessment of data processing within the law (except tort law) has its roots in the historic idea of public law and regulation as much as in the close connection between Data Protection Law and the right to personality. Concerning the increased monetary value of data within the knowledge society and E-Commerce, however, the regulatory approach shows flaws. Rather, a modern legal system also requires protection from the economic exploitation of data in the form of clear provisions on scope and ownership of rights of disposal, access, profit-oriented use, claims for compensation in the event of unlawful use, and deletion of data including fair compensation for the added value deriving from individual data for large data sets (Spiecker gen. Döhmann 2016b). Other areas of information law, such as intellectual property law, patent law or information law, show that this is, in fact, possible. The current law can hardly meet any of these requirements, if any of them at all, and Data Protection Law so far has not concentrated on these aspects. One reason for this is that corresponding regulations are scattered across many different areas of law. In addition, German constitutional law and European primary law impose additional requirements on the processing of personal data, thus creating a complex normative multi-level system that forces a distinction between personal data on the one hand, and, on the other, factual data, e.g., machine data.

4.2.1 Where Law Protects Data As Assets Already

To date, German law does not have any civil law property classification of data per se; in terms of property law, there is no exclusive assignment to a specific person (Specht 2016, p. 288, 289). While data is often used as a payment component of pseudo-synallagmatic contracts (Specht 2016, p. 288, 289; skeptical Hornung and Goeble 2015, p. 265, 270 f.) and subject of various legal regulations or at least governed by them by reflex, there is yet no coherent set of regulations in this respect. The patchwork of regulations ranges from general civil law (such as protection of ownership of the physic data medium) to copyright law (database protection) and competition law (access obligations) to Data Protection Law (rights of data subjects).

4.2.1.1 Protection by Contracts

Civil law serves economical protection by contractual agreements—which are, undoubtedly, particularly relevant in Industry 4.0. Among other tools, sui generis contracts allow for allocation of data by provisions on access to and utilization of data as well as contractual penalties and profit absorption clauses. Concerning contractual agreements, the provisions of the General Terms and Conditions of Business (Section 305 ff. German Civil Code) apply as much as other mandatory law (e.g., provisions of Data Protection Law). In the event of violation of contractual obligations, contractual legal consequences take precedence; otherwise, mandatory law applies, e.g., in the form of claims for damages and torts under Section 280 (1), 241 (2) German Civil Code (Riehm 2018, 74 f.).

The allocation of rights to use and exploit data by means of a contract has the significant advantage of leaving room for negotiation when compared to mandatory law (Hornung and Goeble 2015, p. 265, 271). However, it is exactly this freedom in combination with the special qualities of information (Spiecker gen. Döhmann 2020) which in turn, as a result of information asymmetries, leads to monopolies and economic imbalances between the parties involved, often enough turning into disparity within negotiations and, above all, is detrimental not only to consumers. The essential problem of contracts, moreover, follows from their mere inter-partes effect; they can at best bring about relative, exclusivity-like ties (Markendorf 2018, p. 409, 410), but cannot constitute (hitherto non-existent) property rights to non-physically bound data (Specht 2016, p. 299, 289 f.). Also, the contractual assignment of rights cannot cover the many effects of exploitation of data on third parties, e.g., in personalization. Thus, what often is referred to as data purchase, is ultimately nothing more than the mere granting of (exclusive) rights of use under the law of obligations by one contracting party to the other (Markendorf 2018, p. 409, 410; Paal and Hennemann 2017, p. 1697, 1698), maybe to the detriment of third parties or the common good. In addition, the relativity of contractual obligations results in a decisive weakness compared to absolute rights, particularly on the exchange of data between companies (Stender-Vorwachs and Steege 2018, p. 1361, 1363). Non-contractual third parties frequently involved in Industry 4.0 are not part of the contract. This, for example, applies to the field of intelligent vehicles as well as the problem of a contract between the vehicle owner and the company as a suitable legal basis for the processing of personal data about third parties travelling with the vehicle (Hornung and Goeble 2015, p. 265, 272). Drafting contracts in the case of corporate cooperation is comparably problematic if third parties not involved gain de facto access to data, manipulate, delete, or spy on it (Riehm 2018, p. 74, 75). Contracts reach a further limitation when it comes to the use of data in foreclosure and insolvency proceedings (Riehm 2018, p. 74, 75; Berberich and Kanschik 2017, p. 1ff.). From the perspective of Industry 4.0 companies, however, the advantages of private agreements apparently outweigh the disadvantages; in practice, they have been the means of choice to date. The stakeholders involved should be heard with their assessment of the suitability of contracts in business transactions and their considerations concerning practicability should certainly be considered in reformatory efforts to introduce a strict statutory "data right."[7]

[7] This is valid at least as far as no exclusive right of personal data is concerned—there the situation is completely different due to the involvement of structurally inferior consumers. However, data protection law with its characteristic as a defensive law should be a better solution for the posed problems than an economic over forming of the data protection law.

4.2.1.2 Protection by Tort Law

Tort law offers a different starting point for protection of data as an economic good. It is not only because of the widespread debate about the buzzword "data ownership" that the tort variant of ownership (Section 823 (1) Var. 5 German Civil Code) comes to the fore. Under Section 903 German Civil Code, ownership can only be acquired in objects, i.e., physical objects (Section 90 German Civil Code) (BGH NJW 1972, p. 43; NJW 2009, p. 1947; NJW-RR 2016, p. 982; MüKo BGB 2017, Section 854 No. 2). Therefore, tort claims for infringement of property arise concerning their storage on physical data carriers only.[8] However, this is subject to several—decisive—restrictions. On the one hand, property law does not cover the data itself but only the data carrier; the data stored on it is only protected by reflex.[9] On the other hand, claims for damages arise in the event of deletion or damage to the data only, not in the event of any other type of impact. Finally, only the owner of the affected data carrier is entitled to a tort claim. However, since the vast majority of data is now processed within the cloud, where ownership and possession of the data carrier typically diverge, protection is porous at best. This is even more the case because the person designated as the beneficial owner of the data in private law contracts is usually not the owner of the data carrier (Riehm 2018, p. 74, 78; Specht 2016, p. 288, 289). This shall not diminish the protection of ownership of the data carrier in tort; in some cases, it continues to be of decisive, independent importance (OLG Oldenburg BeckRS 2011, 28832; OLG Karlsruhe NJW 1996, p.200). However, this makeshift construct should not be considered an appropriate solution for the vast majority of legal cases any longer.

The idea of so-called data ownership does not help much either as much as proponents of this concept may claim it. Some assume that one cannot *own* data, but rather only *possess* it within the meaning of Section 854 of the German Civil Code (e.g., Hoeren 2019, p. 5, 7 f.). The inevitable consequences of this construct would not only be the application of the possession protection regulations in favor of the data owner, but also the assumption of data possession as other right in the sense of Section 823 para. 1 Var. 6 German Civil Code. As far as possession of the data medium and its extension to the data stored on are concerned, doubts raised about the concept of data ownership apply. Some, however, argue in favor of ownership of data *itself*, stating that the lack of material quality of data would be no valid argument against this construct and the need of a broad interpretation of the concept of material quality in this respect (Hoeren 2019, p. 5, 7 f.). Moreover, the protection of ownership would convey a relatively weak legal position, so that one could only consider other rights and general access to data when assigning data to ownership

---

[8] *Markendorf*, ZD 2018, 409 (410); *Wandtke*, MMR 2017, 6 (11); *Dorner*, CR 2014, 617 (626); *Determann*, MMR 2018, 277 ff.; differently *Welp*, iur 1988, 443 (448); *Hoeren*, MMR 2013, 486, who plead for an application of § 903 BGB by analogy.

[9] BGH, NJW 2007, 2394; BGH NJW 2016, 1094; OLG Karlsruhe NJW 1996, 200; *Zech*, CR 2015, 137 (142); data is moreover not to be seen as fruit of the thing data carrier, *Specht*, CR 2016, 288 (292); *Zech*, CR 2015, 137 (142).

(in favor of this, most recently Hoeren 2019, p. 5, 7 f.; concerning the comparability of the legal asset of information with possession Redeker 2011, p. 634, 638 f.; against this, Markendorf 2018, p. 409, 410). Regardless of its already questionable practicability, the wording of Section 854 (1) of the German Civil Code speaks against the construct of data ownership. It clearly refers to the "actual power over the thing" and, thus—like ownership—to Section 90 German Civil Code. The immaterial good "data," however, lacks the constituent element of physicality for classification as a "thing". There is also no room for a corresponding application of the ownership provisions due to the lack of comparability of the interests involved.[10] Finally, it should be noted that data ownership on personal data is difficult to conceptualize as most personal data arises in a social context where a multitude of competing ownerships would arise—a situation which serves more to obscure the legal status than to clarify who has access and who can grant access under which conditions.

The scope of the general right to personality as a "framework right" within Section 823 (1) of the German Civil Code is similarly limited. On the one hand, it only covers the handling of personal data, and, on the other hand, case law has been restrictive in the past concerning claims in the event of infringements of personality rights.[11] The general right to privacy grants everyone the right to respect for private sphere (BGH NJW 2012, p. 3645) and, thus, according to German constitutional law, also to the right to informational self-determination, which is relevant to data protection, as an outgrowth of the right to privacy (Schertz 2013, p. 721, 722 ff.). At the same time, the right to personality is an "open provision" without absolute validity. Therefore, to determine whether a violation has occurred or not, there must always be a balance between the goods and interests of the person affected and those of the person acting (Jauernig 2018, Section 823, No. 67), which in a communicative society can often be in favor of the person acting based on the freedom of opinion guaranteed in Article 5 (1) German Constitutional Law. In addition to Section 823 (1) German Civil Code, Section 823 (2) German Civil Code is also capable of protecting data under tort law. In particular, the criminal law provisions of sections 202a-c and 303a German Criminal Code, which protect both exclusive access to data and its integrity, can serve as necessary protective laws (MüKo BGB 2017, Section 823 No. 296; OLG Dresden NJW-RR 2013, p. 27; Zech 2015a, p. 137, 143). These standards are protective laws for individual legal assets, which is why they are applicable within the framework of Section 823 (2) German

[10] *Denga*, NJW 2018, 1371 (1372); furthermore, there might be lacking an unplanned regulatory gap: the legislator recognized that due to its lack of physicalness data is not to be subsumed under the term of § 90 BGB, cf. *Arbeitsgruppe „Digitaler Neustart" der Konferenz der Justizministerinnen und Justizminister der Länder*, Bericht v. 15.5.2017, p. 34, available under: https://jm.rlp.de/fileadmin/mjv/Jumiko/Fruehjahrskonfe-renz_neu/Bericht_der_AG_Digitaler_Neustart_vom_15._Mai_2017.pdf.

[11] E.g., the rating of a teacher by name on the Internet platform spickmich.de was not classified as an infringement of his personal rights, BGH, NJW 2009, 2888; see also *Wybitul/Haß/Albrecht*, NJW 2018, 113 (115); *Kühling*, NJW 2017, 1985 (1990); *Schertz*, NJW 2013, 721 (724 ff.).

Civil Code. Criminal law literature has solved the problem in favor of the person who records the data by way of a "scriptural act," so, to conform to the principle of a uniform legal order, one can argue in favor of transferring this concept to civil law (Hoeren 2017, p. 1587, 1592; Zech 2015a, p. 137, 143). According to this, the person who has created the data by entering or using a program is the owner of the data (Hoeren 2013, p. 486; skeptical Peschel and Rockstroh 2014, p. 309, 312).

4.2.1.3 Assignment of Rights of Disposal by Data Protection Law

The idea of deriving exclusive rights to personal data from the very fact that individuals are affected by them under Data Protection Law (Specht 2017, p. 1040, 1042) is linked to a fundamental debate about how the right to personality—and, thus, also its category of the right to informational self-determination (BVerfGE 65, p. 1 ff.; Schertz 2013, p. 721, 722 ff.)—is structured as a property-like right (Specht 2016, p. 288, 291 f.). It is true that in the relevant census ruling, the German Constitutional Court stated that the individual has no right "in the sense of absolute, unrestricted control" over his or her data (BVerfGE 65, p. 1, 43 f.). Nonetheless, this does not *a priori* prevent the simple legal attribution of individual data exploitation powers under narrower conditions, particularly those flanked by rules on limits (Specht 2016, p. 288, 292; 2017, p. 1040, 1041 f.). Admittedly, this is a road not taken (yet) by the legislature. This idea is, at least for the time being, therefore, nothing more than an intellectual game. It has been pushed forward, however, particularly in the international debate, also as a counter-argument to the protective measures of the GDPR. In any case, since the GDPR came into force, one must consider the overarching effect of EU law and its application over national Data Protection Law. In this respect, EU law neither contains a constitutional-like construct of informational self-determination nor does the GDPR, as a regulatory expression of the European fundamental right to data protection (Art. 8 CFR), provide indications for a (re)orientation of Data Protection Law as property law (Härting 2016, p. 63). It is rather clear that EU law adheres to a personality-rights-based approach, as does the ECHR.

Some authors object to Art. 20 GDPR (Data Portability), which they refer to as the initial legal-dogmatic door-opener for market-based data trade (Jülicher et al. 2016, p. 358, 361). However, this interpretation is misleading since the norm is not linked to personal data, but merely establishes a (legal) obligation on the part of the data controller, which cannot be waived, for example by agreement.[12] Even if one neglects the fact that the GDPR pursues elements of competition policy and consumer protection goals in its approach to prevent against informational power

[12] *Von Lewinski,* in: BeckOK Datenschutzrecht, Wolff/Brink, 26. Ed., Art. 20 Rn. 8; *Kamann/Braun*, in: Ehmann/Selmayr, DSGVO, Art. 20 Rn. 4; critically probably also *Dix*, in: Simitis/Hornung/ Spiecker gen. Döhmann, Datenschutzrecht, Art. 20 Rn. 1; to the question, which data Art. 20 GDPR covers *Spiecker gen. Döhmann* p. 194.

asymmetry, Article 20 GDPR fits into the canon of defensive rights of data subjects and complements their concern to give data subjects better control over their data (Kühling and Buchner 2018, Art. 20 No. 4; Ehmann and Selmayr 2018, Art. 20 No. 1; Gola 2018, Art. 20 No. 1; Simitis/Hornung/Spiecker gen. Döhmann 2019, Art. 20 No. 1). In this respect, data portability is more similar to the right to information under Data Protection Law, albeit without the primary aim of informing the data subjects—e.g., as a precondition for exercising other rights (Wolff/Bring 2018, Art. 20 No. 7; Kühling and Buchner 2018, Art. 20 No. 2). An interpretation of data portability as a trading right would be counterproductive. Art. 20 GDPR wants to disable natural monopolies due to the network-related economies of scale and scope of information. At present, therefore, Data Protection Law does not constitute any individual exclusive rights to personal data.[13] Rather, Data Protection Law, as it is currently structured, would be an instrument of downstream limitation of any exclusive rights to data that might be introduced (Denga 2018, p. 1371, 1373; Specht 2017, p. 1040, 1042). The holder of such a right must always observe the mandatory requirements of Data Protection Law when making dispositions and obtain the consent of the data subject or base his disposition on a statutory element of permission (Specht 2016, p. 288, 294).

4.2.2 Concepts for a Future Allocation of Exclusive Rights to Data

Uniformly, legal literature has sought to address the consequences of the economic valorization of data and the concomitant need to create an economic framework. The approaches, however, differ.

Some authors argue in favor of keeping the legal status quo. They find the law to be already sufficiently capable of balancing the different interests of the parties involved (Dorner 2014, p. 617, 626). Strict exclusive rights to data would not only be a disincentive to economic investment and of little value to consumers, but also collide with constitutional requirements. In line with the popular saying, "thoughts are free," some authors fear the danger of overregulation and advocate for data to oscillate as free from legal boundaries as possible (Determann 2018, p. 503, 506 ff.), often in sync with the early idea of the Internet as one of the core drivers of data-driven economy being a free sphere.

On the other hand, there are reform ideas, some of which build on existing instruments, and, where appropriate, provide for selective modifications to the law that the legislature can implement without extensive intervention or that the courts

[13] *Paal/Hennemann*, NJW 2017, 1697 (1698 with further references.); implicitly *Denga*, NJW 2018, 1371 (1373); critically to possible tendencies *Dix*; ZEuP 2017, 1 ff.; questionable wording at *Stender-Vorwachs/Steege,* NJOZ 2018, 1361 (1362, "verfügen" ("dispose"). In result against a positive power of disposal *Bundesministerium für Verkehr und digitale Infrastruktur,* "Eigentumsordnung" für Mobilitätsdaten? – Eine Studie aus technischer, ökonomischer und rechtlicher Perspektive", August 2017, p. 50; available under: http://www.bmvi.de/SharedDocs/DE/Publikationen/DG/eigentumsordnung-mobilitaetsdaten.pdf?blob=publicationFile.

can implement through interpretation. In the shadow of the data ownership debate—which is considered to have failed, at least for Europe—there is, for example, the approach of broadly interpreting the provisions on data bank producer rights under Sections 87a et seq. of the German Copyright Act (Wiebe 2017, p. 338 ff.; 2018, p. 97 ff.). Some also consider whether an independent right to data should serve as a miscellaneous right within the meaning of Section 823 (1) German Civil Code (Riehm 2018, p. 74, 84 ff.). Another concept with limited legislative intervention could be the creation of platforms and platform regulation. With minor changes to content control (Section 307 ff. German Civil Code) and the extension of antitrust exemption, some consider them as an efficient alternative to the introduction of exclusive rights because they allow flexible agreements for the use of data (Spindler 2018, p. 151, 169 f.; Metzger 2019, p. 129 ff.). On personal data alone, there are also considerations in the direction of a data use right linked to existing copyright law (Wandtke 2017, p. 6 ff.; Zech 2015b, p. 1151, 1154) and a sui generis intellectual property right in behavior-generated data (Fezer 2017a, p. 3ff.; 2017b p. 99 ff.). Finally, administrative agencies and courts have actively opened competition law for remedies under this legal regime for violation of data protection (OLG Düsseldorf, Beschluss vom 24.03.2021—Kart 2/19 (V)) and thus enlarged the field of compliance in companies on Data Protection Law.

Recently, there have been various reform proposals for a uniform European regulatory framework, particularly for data markets and thus stressing the commercial aspects of data.[14] At the national German level, two major studies were published. First, the working group "Digitaler Neustart" (Digital Restart) of the Conference of Ministers of Justice of the German Federal States dealt with the preliminary work of the literature on property rights. However, the Conference assumed that the civil law "patchwork" of data protection regulations forms a "sufficiently closed system of protection."[15] Another study on ownership of mobility data by the federal ministry for justice and consumer protection identifies weaknesses in all existing legal instruments for data exploitation protection. A property-like right to data to be developed could adopt elements of existing approaches and should ultimately be based on two pillars: The idea of investment protection and the scriptural act.[16] Finally, the German Ministry of Justice and Consumer Protection

[14] EU Commission, Communication "Building a European Data Economy", COM (2017) 9 final; EU-Commission, Communication "Towards a common European data space", COM (2018) 232 final; to this instead of many *Peitz/Schweitzer*, NJW 2018, 275 ff. with further references; *Schweitzer*, ZEuP 2019, 1 ff.

[15] *Arbeitsgruppe "Digitaler Neustart" der Konferenz der Justizministerinnen und Justizminister der Länder*, Bericht v. 15.5.2017, p. 98; available under: https://jm.rlp.de/fileadmin/mjv/Jumiko/Fruehjahrskonferenz_neu/Bericht_der_AG_Digitaler_Neustart_vom_15._Mai_2017.pdf.

[16] *Bundesministerium für Verkehr und digitale Infrastruktur*, "Eigentumsordnung" für Mobilitätsdaten? – Eine Studie aus technischer, ökonomischer und rechtlicher Perspektive", August 2017, p. 104; available under: http://www.bmvi.de/SharedDocs/DE/Publikationen/DG/eigentumsordnung-mobilitaetsdaten.pdf?blob=publicationFile. Critically to this Determann, ZD 2018, 503 (506).

has pushed for a scientific debate on the rights of access to data where several potentially contradictory legal regimes were presented. The conclusion remained that economic value of data is not all that should govern the transfer of data and access to it, especially on data protection and the necessity to provide data as a backbone for innovation for all.[17]

However, each of these approaches will measure against the comprehensive, albeit not always precise, legal framework of the GDPR. The problem remains that the GDPR does not contain a convincing protection of operational as well as business data.

5 Conclusion

The protection of data in Industry 4.0 under Data Protection Law is no independent form of protection in disregard of other legal regimes touching upon data, but rather a conglomerate of existing regulatory approaches. As a result, the major problem of allocating data generated in Industry 4.0 ventures remains largely unresolved. There is no fundamental resolution about who has access to the data generated using individualized manufacturing processes and under which conditions.

Data Protection Law is of considerable help to the extent that it specifically links the protection of personal data to the data subject and assigns the central role to the data subject. It can also be considered to be the most comprehensive legal framework on data, so far (Spiecker gen. Döhmann, in Peukert (Ed.), Global Digitality, forthcoming 2022). The legal basis of Article 6 (1) (f) GDPR can also legitimize data processing within the framework of a balancing of interests that is independent of the ideas of the data subject. In the area of data transactions, it is easy to show that there are clear assessments, which enable the companies involved to make a legally secure assessment. However, this does not imply a right to data or a data sovereignty that is equivalent to monetary terms, from which further implications for non-personal data could derive.

Approaches to a regulation already exist; however, they can only insufficiently specify and translate the diverse references in Industry 4.0. This is even more true as they are predominantly based on existing concepts and are, thus, hardly able to integrate the specifics of processing in automated and, at the same time, individualized manufacturing processes.

[17] *Spiecker gen. Döhmann* in: German Federal Ministry of Justice and Consumer Protection/Max Planck Institute for Innovation and Competition, Data Access, Consumer Interests and Public Welfare, 1. ed. 2021 p. 206 f.

References

Badmann M (2019) Grundlagen der Allgemeinen Betriebswirtschaftslehre, 3rd edn. Springer Gabler, Wiesbaden

Baranowski A, Glaßl R (2017) M&A im Internet. Transaktionen von Daten und Content. BB 199–207

Benecke A, Wagner J (2016) Öffnungsklauseln in der Datenschutzgrundverordnung und das deutsche BDSG – Grenzen und Gestaltungsspielräume für ein nationales Datenschutzrecht. DVBl:600–608

Berberich M, Kanschik J (2017) Daten in der Insolvenz. NZI:1–10

Beyer A, Beyer A (2016) Verkauf von Kundendaten in der Insolvenz – Verstoß gegen datenschutzrechtliche Bestimmungen? NZI:241–245

Bräutigam P, Klindt T (2015) Industrie 4.0, das Internet der Dinge und das Recht. NJW 1137–1142

Deißler L-S (2018) Gewährleistung von Informationsqualität in europäischen Informationssystemen. Nomos, Baden-Baden

Denga M (2018) Gemengelage privaten Datenrechts. NJW:1371–1376

Determann L (2018) Gegen Eigentumsrechte an Daten. ZD:503–508

Dix A (2017) Daten als Bezahlung – Zum Verhältnis zwischen Zivilrecht und Datenschutzrecht. ZEuP:1–5

Dorner M (2014) Big Data und "Dateneigentum", Grundfragen des modernen Daten- und Informationshandels. CR 617–628

Drexl J (2017) Neue Regeln für die Europäische Datenwirtschaft? NZKart 339–344

Ehmann E, Selmayr M (2018) Datenschutz-Grundverordnung. C.H. Beck, München

Fezer K-H (2017a) Dateneigentum Theorie des immaterialgüterrechtlichen Eigentums an verhaltensgenerierten Personendaten der Nutzer als Datenproduzenten. MMR:3–5

Fezer K-H (2017b) Dateneigentum der Bürger. ZD:99–105

Gola P (2018) Datenschutz-Grundverordnung. C.H. Beck, München

Göpfert B, Meyer S (2011) Datenschutz bei Unternehmenskauf: Due Diligence und Betriebsübergang. NZA:486–493

Grünwald A, Hackl J (2017) Das neue umsatzbezogene Sanktionsregime der DS-GVO: Bußgeldbemessung nach kartellrechtlichen Maßstäben? ZD:556–560

Härting N (2016) Datenschutz-Grundverordnung – Das neue Datenschutzrecht in der betrieblichen Praxis. Otto-Schmidt, Köln

Härting N (2017) Kundendaten beim Unternehmenskauf nach DSGVO. CR:724–727

Hirsch D (2014) The glass house effect: big data, the new oil, and the power of analogy. Maine Law Rev 66:374–395

Hirsch-Kreinsen H (2016) Industrie 4.0 als Technologieversprechen. Soziologisches Arbeitspapier 46/2016, TU Dortmund

Hoeren T (2013) Dateneigentum – Versuch einer Anwendung von § 303a StGB im Zivilrecht. MMR:486–491

Hoeren T (2017) IT- und Internetrecht – Kein Neuland für die NJW. NJW 1587–1592

Hoeren T (2019) Datenbesitz statt Dateneigentum. MMR:5–8

Hofmann K (2016) Smart Factory – Arbeitnehmerdatenschutz in der Industrie 4.0 – Datenschutzrechtliche Besonderheiten und Herausforderungen. ZD 12–17

Hornung G, Goeble T (2015) "Data Ownership" im vernetzten Automobil. CR 265–273

Jauernig O (2018) Bürgerliches Gesetzbuch Kommentar, 17. ed. C.H. Beck, München

Jülicher T, Röttgen C, von Schönfeld M (2016) Das Recht auf Datenübertragbarkeit – Ein datenschutzrechtliches Novum. ZD:358–362

Kling-Straub J, Straub T (2018) Nächste Ausfahrt DS-GVO – Datenschutzrechtliche Herausforderungen beim automatisierten Fahren. NJW:3201–3206

Kühling J, Buchner B (2018) Datenschutz-Grundverordnung/Bundesdatenschutzgesetz. C.H. Beck, München

Leimster JM (2015) Einführung in die Wirtschaftsinformatik, 12th edn. Springer Gabler, Wiesbaden
Markendorf M (2018) Recht an Daten in der deutschen Rechtsordnung. Blockchain als Lösungsansatz für eine rechtliche Zuordnung? ZD 409–413
Marsch N (2018) Das europäische Datenschutzgrundrecht. Mohr Siebeck, Tübingen
Mertens P, Barbian D, Baier S (2017) Digitalisierung und Industrie 4.0 – eine Relativierung. Springer Vieweg, Wiesbaden
Metzger A (2019) Digitale Mobilität – Verträge über Nutzerdaten. GRUR:129–136
Mühlich R (2014) Datenschutz und Industrie 4.0 – eine sensible Annäherung ist notwendig. ZD 381–382
Müllmann D (2018) Auswirkungen der Industrie 4.0 auf den Schutz von Betriebs- und Geschäftsgeheimnissen. WRP 1177–1182
Nebel J (2016) Die Zulässigkeit der Übermittlung personenbezogener Kundenstammdaten zum Vollzug eines Asset Deals. CR:417–424
Paal BP, Hennemann M (2017) Big Data im Recht – Wettbewerbs- und daten(schutz)rechtliche Herausforderungen. NJW:1697–1701
Peschel C, Rockstroh S (2014) Chancen und Risiken neuer Daten-basierter Dienste für die Industrie, Tagungsband DSRI Herbstakademie 2014. OLWIR, Edewecht, p 377–390
Plath (2014) Teil. 6 Datenschutz bei M&A – Transaktionen. In: von dem Bussche A, Voigt P (eds) Konzerndatenschutz. C.H. Beck, München Herausforderungen im Datenschutz zwischen Unternehmen
Riehm T (2018) Dateneigentum – Schutz nach Zivilrecht. In: Hornung G (Hrsg) Rechtsfragen der Industrie 4.0. Nomos, Baden-Baden, p 74–96
Schertz C (2013) Der Schutz des Individuums in der modernen Mediengesellschaft. NJW:721–728
Schröder C (2017) Ch. 4 Datenschutz in der Unternehmenstransaktion. In: Forgó N, Helfrich M, Schneider J (Hrsg) Betrieblicher Datenschutz. C.H. Beck, München
Simitis S, Hornung G, Spiecker gen. Döhmann I (2019) Datenschutzrecht. Nomos, Baden-Baden
Specht L (2016) Ausschließlichkeitsrechte an Daten – Notwendigkeit, Schutzumfang, Alternativen. CR 288–296
Specht L (2017) Das Verhältnis möglicher Datenrechte zum Datenschutzrecht. GRUR Int:1040–1047
Spiecker gen. Döhmann I (2016a) Zur Zukunft systemischer Digitalisierung – Erste Gedanken zur Haftungs- und Verantwortungszuschreibung bei informationstechnischen Systemen. CR 698–704
Spiecker gen. Döhmann I (2016b) Schlusswort: Jedem Anfang wohnt ein Zauber inne. In: Dreier T, Fischer V, van Raay A, Spiecker gen. Döhmann I Informationen der Öffentlichen Hand – Zugang und Nutzung. Nomos, Baden-Baden Studien zur Informationsfreiheit, p 585–598
Spiecker gen. Döhmann I (2020) § 33: Information Management. In: Cane P, Hofmann H, Ip E, Lindseth P/ (Ed.), The Oxford Handbook of Comparative Administrative Law. Oxford University Press, Oxford, p 677–694
Spiecker gen. Döhmann I (2021) The legal framework for access to data from a data protection viewpoint – especially under the GDPR. In: Data Access, Consumer Interests an Public Welfare, Edited by German Federal Ministry of Justice and Consumer Protection, Max Planck Institute for Innovation an Competition. Nomos, Baden-Baden, p 175–208
Spindler G (2018) Plattformen und Plattformregulierungen als Alternative zu Dateneigentumsrechten. In: Hornung G (ed) Rechtsfragen der Industrie 4.0. Nomos, Baden-Baden, pp 151–174
Staffler L (2018) Industrie 4.0 und wirtschaftlicher Geheimnisschutz. NZWiSt 269–274
Stender-Vorwachs J, Steege H (2018) Wem gehören unsere Daten? Zivilrechtliche Analyse zur Notwendigkeit eines dinglichen Eigentums an Daten, der Datenzuordnung und des Datenzugangs. NJOZ:1361–1367
Thalhofer T (2017) Recht an Daten in der Smart Factory. GRUR-Prax:225–227

Thode J-C (2016) Kundendaten beim Unternehmenskauf – tatsächlich ein Datenschutzproblem? PinG:26–29
Wandtke A-A (2017) Ökonomischer Wert von persönlichen Daten. MMR:6–12
Weichert T (2001) Datenschutz als Verbraucherschutz. DuD:264–270
Welp J (1988) Datenveränderung (§ 303a StGB) – Teil 1. iur 443–449
Wiebe A (2017) Schutz von Maschinendaten durch das sui-generis Schutzrecht für Datenbanken. GRUR:338–345
Wiebe A (2018) Schutz von Daten im Immaterialgüterrecht. In: Hornung G (ed) Rechtsfragen der Industrie 4.0. Nomos, Baden-Baden, pp 97–118
Wilhelm B (2015) Anmerkungen zur Produktion von morgen. zfo 187
Zech H (2015a) Daten als Wirtschaftsgut – Überlegungen zu einem "Recht des Datenerzeugers". CR:137–146
Zech H (2015b) "Industrie 4.0" – Rechtsrahmen für eine Datenwirtschaft im digitalen Binnenmarkt. GRUR 1151–1160

Information Security Law 4.0

Paul Voigt

Contents

1 Introduction

Information technology forms the basis of basically all company procedures. Although it offers the advantage of high efficiency, it also harbours a great potential for risk. Many companies have been hit by a hacking attack recently (Bitkom 2018, p. 19). At the same time, the increasingly rapid pace of digitisation in the private

P. Voigt (✉)
Attorney at Law, Taylor Wessing, Berlin, Germany
e-mail: p.voigt@taylorwessing.com

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_10

sector is leading to a growing dependence on the functionality of information technology systems (Schober 2018, p. 12). Increasing household appliances are linked to the Internet. Consequently, there will probably be 797.6 million networked devices in Germany by 2022; on average, this would mean that for every citizen there would be around ten smart devices (Cisco 2018, p. 3) that could be exposed to potential threats in the form of malware or targeted cyberattacks. Even the simplest household and consumer goods—such as light bulbs (see Steigerwald 2018), vacuum cleaners or refrigerators—are becoming a threat, as they can enable hackers to access the entire local network.

In connection with the **Internet of Things** (IoT), two hazardous situations are conceivable. On the one hand, an IoT device could be compromised to harm the user, for example by manipulating or spying on data or sabotaging the device (BSI 2018, p. 21). On the other hand, compromising the device can also aim at harming third parties by setting up a botnet and using it for DDoS attacks or by using the compromised device as a proxy to disguise identity in further attacks (BSI 2018, p. 21).

In view of the increasing risk potential, it is in the self-interest of every company to ensure that its information technology equipment is protected against dangers and damages of any kind. However, there is often also a legal obligation to take protective measures against IT threats. The most important of these obligations under information security law, as well as possible threat scenarios, are briefly described below.

2 Information Security As a General Obligation Under Commercial Law

Obligations on information security may arise from the applicable commercial law regulations. In the case of corporations, the management bodies are responsible for the essential obligations of managing the company and complying with legal requirements. This management task also includes **ensuring information security**, which includes risk prevention and IT compliance.

2.1 Obligation to Identify Existential Risks at an Early Stage

Under § 91 Sec. 2 of the German Stock Corporation Act (AktG), the Management Board of a stock corporation must take appropriate measures to ensure **early detection of risks that could jeopardise the continued existence of the company.** In view of the omnipresent use of information technology in companies, both the financial damage directly associated with a security incident—such as production or operational failures—and any loss of reputation due to high-profile cyberattacks can have consequences that could jeopardise the company's continued existence. Therefore, the achievement and maintenance of information security is a **task of the Executive Board** under § 91 Sec. 2 AktG, whereby in practice often an individual Executive Board member is assigned with the duty to establish information security.

Possible security gaps in the IT infrastructure must be detected early enough to ensure that it is still possible to prevent the high-risk development (Spindler 2014, marginal 27). Under § 91 Sec. 2 AktG, this particularly requires the **establishment of an early detection and monitoring system**, which regularly forms a partial component of the general risk management system (for details on the implementation of such a system: Voigt 2022, marginals 44–48). In addition, under § 91 Sec. 3 AktG, the Executive Board may have to establish an appropriate and effective internal control and risk management system.

The principle of risk distribution contained in § 91 Sec. 2 AktG can have a **spillover effect on other forms of companies**—particularly the German Ltd. 'GmbH'—because of the similar duties carried out by the managers (Spindler 2014, marginal 87). The selection and resolution of the management on risk monitoring is then based on the requirements of § 91 Sec. 2 AktG (Spindler 2014, marginal 87). However, this only applies to companies above a certain size and complexity, as the establishment of a comprehensive monitoring system cannot be imposed on small companies to the same extent as on large corporations (Spindler 2014, marginal 87).

2.2 IT Compliance

Furthermore, § 76 Sec. 1 AktG imposes the general obligation to act in accordance with the applicable legal system—and consequently also in **accordance with the legal requirements relating to the lawful use of information technology.** Within the scope of this 'IT compliance', both preventive measures, such as regular internal training of personnel and organisational security measures, must be met to contain identified risks. This can and should counteract threat scenarios such as infection with malware, the exploitation of software vulnerabilities or phishing attacks.

3 Information Security As a Contractual Obligation

However, information security can become relevant not only in the context of internal company legal obligations, but also as a contractual obligation of the company towards third parties. In principle, information security can both represent a main contractual obligation or a secondary obligation.

3.1 General Information

The fact that information security (also) represents a **material contractual duty** must be regularly assumed, for example, in the case of non-disclosure agreements, data processing agreements, contracts for the maintenance and support of hardware or software, cloud computing contracts or outsourcing contracts (for details see Voigt 2022, marginals 97–102).

Information security as a **secondary contractual obligation** within the meaning of § 241 Sec. 2 of the German Civil Code (BGB) must be assumed, for example, in online banking, which is becoming increasingly important for the processing of everyday transactions. Since security incidents can easily result in damage to the customer, the maintenance of information security is a secondary obligation of the bank, the breach of which can result in claims for damages by the contractual partner or trigger rights of termination or damage claims (Voigt 2022, marginals 134–137 and marginal 143).

However, with respect to **IoT**, contracts for the sale of IT products—such as **sales contracts for smart home products**—and the associated obligations of the seller or manufacturer are of particular importance (see Sect. 3.2).

3.2 Purchase Contracts for IT Products

If the object of purchase has a 'defect', the purchaser is entitled to certain warranty rights for defects. In this respect, it appeared questionable whether a **lack of IT security** in the form of a security gap is to be regarded **as a defect.** However, under the new purchase law in force since the beginning of the year 2022, this issue has now been clarified, as the relevant provisions (§§ 434 Sec. 3 sent. (2), 327e Sec. 3 No. 2 BGB) expressly mention the safety of the product as an objective requirement. Decisive is the condition of the IT product at the **time of transfer of the risk (handover or provision)**. If the IT product does not have a security gap at the time of transfer of the risk, the purchaser's warranty claims for defects in this respect shall not be considered.

Provided that the threat level for **IT security** subsequently changes—for example, due to new types of malware or targeted cyberattacks that overcome the security standards previously regarded as sufficient—the question arises as to whether and to what extent the provider or seller is obliged to **provide security updates.** At least for the B2C space, this question has been answered by the new purchase law applicable from January 2022: For digital products and so-called goods with digital elements (e.g. smart watches) that are sold to consumers, the respective sellers are obliged to provide (IT security) updates for a certain period of time, cf. § 327f Sec. 1 sent. 1, 2 and § 475b Sec. 4 No. 2 BGB. In B2B scenarios, such update obligation is generally not stipulated in the new purchase law.

Differing from the legal situation in place until 2022, an IT security gap may constitute a defect, irrespective of wheter a certain quality of the IT product has been agreed upon or whether the completeness of IT security is a **customary quality** of the same products in the general legal perception.

If the purchaser suffers any **consequential damage** due to the IT security gap, a claim for damages against the seller under § 280 I BGB as well as against the producer under § 823 I BGB (producer liability, for details see Spindler 2007, p. 48 ff.) may be possible. Such a claim would be conceivable, for example, if unauthorised persons gained access to the account data of the buyer due to an IT security gap and consequently carried out disadvantageous transactions.

The factual problem is that it is **usually difficult for consumers to know** whether the software is up to date and whether and to what extent updates are made available. The sellers often provide insufficient information in this respect; sometimes it can even be observed that products with known IT security gaps and no possibility of a security update are sold as new (BSI 2018, p. 45 on the 'Stagefright' security gap of the Android operating system).

4 Information Security to Protect Business Secrets

Business secrets are regularly an essential part of a company's assets; their disclosure can have existential consequences that threaten the company's existence (see Brammsen 2016, p. 2193 f.). The extent of the risks emanating from industrial espionage became clear in October 2018, when it was reported that harmful computer chips were allegedly installed directly in hardware components used worldwide by a Chinese supplier (Scherschel 2018).

The protection of trade secrets should therefore be a very high priority for companies. From a legal point of view, § 2 Sec. 1 of the German Act on the Protection of Trade Secrets (GeschGehG) stipulates that information can only be classified as **'trade secrets'** if it is subject to **appropriate secrecy measures**. It is therefore largely up to the company to determine in each individual case, by taking appropriate precautions, which information is subject to secrecy protection. Since it is not necessary to take the best possible, but only 'appropriate' protective measures, it should first be critically examined whether and to what extent the economic and personal implementation effort of a specific measure is in adequate relation to the importance of the information to be protected and the connected imminent risks (Voigt et al. 2019, p. 144). Conceivable **secrecy measures** include, for example, the establishment of password protection, the use of two-factor authentication, the encryption of data and connections as well as state-of-the-art virus and malware protection (Voigt et al. 2019, p. 144).

5 Information Security Obligations for Critical Infrastructures and Companies of special public Interest

The Act on the Federal Office for Information Security (BSIG) imposes high information security requirements on companies in critical supply sectors (**KRITIS operators**), as e.g. hacker attacks in this area can have considerable consequences for a large number of consumers. The progressing interconnectedness of technical devices also increases the vulnerability of supply service providers across multiple channels. One threat scenario is that attackers gain access to a system via intelligent electricity meters (**smart meters**) and trigger incorrect control actions via the communication infrastructure, which can result in overloads and system failures (Meyer 2015).

Under § 2 Sec. 10 (1) BSIG, critical infrastructures are facilities, installations or parts thereof which (1) belong to the **energy, information technology and telecommunications, transport and traffic, health, water, food or finance, insurance and municipal waste sectors** and (2) are of high importance for the functioning of the community since their failure or impairment would result in significant supply shortages or threats to public safety. The term is defined in greater detail by the Regulation on Critical Infrastructures (BSI-KritisV), which was issued based on § 10 Sec. 1 BSIG.

KRITIS operators are obliged under § 8a Sec. 1 (1) BSIG to take appropriate organisational and technical **precautions to avoid disruptions to** the availability, integrity, authenticity and confidentiality of their information technology systems, components or processes that are essential for the functioning of the critical infrastructures operated by them. To determine the specific scope of obligations, a proportionality test is carried out, taking into account the state of the art, § 8a Sec. 1 sent. (2) and (3) BSIG. In this context, the 'state of the art' is a dynamic reference value, which continuously adapts to changing possibilities. Consequently, the obligated company must also regularly update its security measures (Voigt 2022, marginal 269).

Furthermore, under § 8a Sec. 3 sent. (1) BSIG, KRITIS operators must provide appropriate **evidence of** compliance with the above requirements at least **every two years**. As possible forms of proof, § 8a Sec. 3 sent. (2) BSIG mentions safety audits, tests or certifications. In addition, under § 8b Sec. 4 BSIG, KRITIS operators are **obliged to report significant incidents** to the BSI (for details on the reporting obligation: Voigt 2022, marginals 277–292).

If an operator of KRITIS intentionally or negligently violates its obligations under the BSIG, a fine of up to 20,000,000 euros can be imposed under § 14 Sec. 5 BSIG, § 30 OWiG.

The BSIG also affects manufacturers of IoT devices that are not KRITIS operators. One point of criticism of the Federal Ministry of the Interior concerning IoT in the explanatory statement of the draft bill for an 'IT Security Law 2.0' dated 9 December 2020 was that IoT devices are not developed under IT security aspects, which could obviously affect consumer protection in the view of the Ministry.

Therefore, under § 3 Sec. 1 sent. (2) No. 14a BSIG the BSI's tasks now include the promotion of consumer protection and consumer information in the field of security of information technology, particularly by **advising, informing and warning consumers**. The BSI shall thereby focus and expand its tasks and powers on consumer-related products and services, for example IoT in addition to operating systems (such as Windows 10 and Android), hardware consumer products (smartphones or SmartTVs), Alexa, GoogleHome and SmartHome devices. Another innovation is that so-called companies of special public interest (§ 3 Sec. 14 BSIG), which can be considered "KRITIS light operators", e.g. because they belong to the biggest companies in Germany, are now also subject to certain—somewhat softened in comparison with KRITIS operators—information security obligations under the BSIG.

In addition, the BSI is able to issue, at the request of the respective company, a voluntary IT security label for various product categories (cf. § 9c BSIG) which is defined in more detail by the BSI IT Security Label Regulation (BSI-ITSiKV). This IT security label contains a declaration of the manufacturer about the existence of certain IT security features ('Herstellererklärung') as well as information from the BSI about security gaps or other security-relevant IT features ('Sicherheitsinformation'). After the application for the use of the IT security label has been filed, the BSI checks the submitted documents for plausibility and then either allows the use of the label or refuses its use. The BSI can check at regular intervals or on an ad hoc basis after the IT security label has been issued whether the requirements of the IT security label are still met.

In this form, the IT security label could promote market transparency, but could also serve manufacturers of such devices as a suitable marketing instrument.

6 Obligations for Providers of Telemedial/Digital Services

For providers of digital or telemedial services, information security obligations may arise from § 19 Sec. 4 sent. (1) German Telecommunications and Telemedia Data Protection Act (TTDSG) on the one hand and § 8c Sec. 1 sent. (1) BSIG on the other. Risk scenarios consist primarily of hackers gaining access to customer or user data on a large scale.

6.1 Telemedia Service Providers (§ 19 Sec. 4 Sent. (1) TTDSG)

§ 19 Sec. 4 sent. (1) TTDSG obliges telemedia service providers to ensure, by means of technical and organisational precautions and within the scope of their respective responsibility for commercially offered telemedia that (1) there is no possibility of **unauthorised access** to the technical equipment used for their telemedia offers and

(2) that this equipment is secured against disruptions, including those caused by external attacks.

Under the definition of § 2 Sec. 2 No. 1 TTDSG, the **telemedia service provider**, thus the addressee of the provision, is any natural or legal person who provides his own or third-party telemedia, participates in the provision of such telemedia or provides access to the use of its own or third-party telemedia. The term includes services, which are not exclusively broadcasting or telecommunications services (German Bundestag 2006, p. 13), **for example, online search engines, shops, auction houses or social networks** (Voigt 2022, margin no. 365 w.f.r.). The scope of application of § 19 Sec. 4 sent. (1) TTDSG is also limited by the requirement of business-like conduct, which assumes scheduled and permanent activity (German Bundestag 2015, p. 34).

The precautions to be taken **depend on the individual case** and must be technically possible and economically reasonable. Concrete measures may include access and admission controls (Spindler and Schmitz 2018, margin no. 89) or the use of an encryption procedure recognised as secure, § 19 Sec. 4 sent. (3) TTDSG. If the service provider does not implement the necessary precautions, a fine of up to 10,000 euros can be imposed, § 28 Sec. 1 No. 10, Sec. 2 TTDSG.

6.2 Providers of Digital Services (§ 8c Sec. 1 Sent. 1 BSIG)

Under § 2 Sec. 11, Sec. 12 BSIG, providers of digital services are legal entities which offer **online marketplaces, online search engines or cloud computing services.** Under § 8c Sec. 1 sent. (1) BSIG, providers of digital services must take suitable and **adequate technical and organisational measures** to manage risks to the security of the network and information systems which they use to provide digital services within the European Union. Under § 8c Sec. 2 sent. (1) BSIG, the measures must warrant for an appropriate security level of the network and information systems corresponding to the existing risk, taking into account the state of the art. This mostly corresponds to the standard applicable to KRITIS operators in § 8a Sec. 1 BSIG. Furthermore, the measures must be suitable to prevent security incidents or to minimise the effects of a security incident, § 8c Sec. 1 sent. (2) BSIG. Under § 8c Sec. 2 sent. (2) BSIG, the following factors must be considered when determining the required IT security standard:

- the security of the systems and facilities
- the detection, analysis and containment of security incidents
- business continuity management
- regular monitoring, verification and testing
- compliance with international regulations

A more detailed definition of the measures is provided by the implementing acts of the Commission under Art. 16 (8) of Directive (EU) 2016/1148, § 8c Sec. 2 (3) BSIG. On 30 January 2018, a corresponding ‘Implementing Regulation laying down rules for the application of Directive (EU) 2016/1148 of the European Parliament and of the Council regarding further specification of elements to be taken into account by digital service providers for managing the risks posed to the security of network and information systems and of the parameters for determining whether an incident has a substantial impact' was adopted, which contains a list of security elements in its Article 2.

If a significant security incident occurs, providers of digital services are obliged to **report it immediately** to the BSI, § 8c Sec. 3 sent. (1) BSIG. In determining whether the materiality threshold triggering the obligation to report has been reached, the following must particularly be considered under § 8c Sec. 3 sent. (2) BSIG:

– the number of users affected by the security incident, particularly those users requiring the service to provide their own services
– the duration of the security incident
– the geographic area affected by the security incident
– the extent of interruption of the provision of the service
– the extent of the effects on economic and social activities

Under Art. 4 (1) of the Implementing Regulation, significant effects are deemed to be present if:

– the service provided by a digital service provider was unavailable for more than five million user hours
– the security incident has led to a loss of integrity, authenticity or confidentiality (. . .) affecting more than 100,000 users in the Union
– the security incident has created a public danger, a risk to public security or has caused loss of life
– the security incident has resulted in material damage of more than one million Euro for at least one user in the Union

If a provider of digital services intentionally or negligently fails to take the measures referred to in § 8c Sec. 1 sent. (1) BSIG or if they fail to report, to report properly, not completely or not in due time under § 8c Sec. 3 sent. (1) BSIG, they are acting in breach of the regulations under § 14 Sec. 2 No. 7 or No. 8 BSIG and are liable to a fine of up to 500,000 euros, § 14 Sec. 5 sent. (2) BSIG.

6.3 Overlapping Scope of Application

If a **provider of digital services** within the meaning of § 2 Sec. 11 and Sec. 12 BSIG is **also a telemedia service provider** within the meaning of § 2 Sec. 2 No. 1 TTDSG, **§ 19 Sec. 4 TTDSG** is **generally superseded,** due to the aim of attaining full harmonisation by the European Directive concerning measures for a high common level of security of network and information systems across the Union (NIS Directive, EU No. L 194 of 19.7.2016, pp. 1–30), cf. Art. 16 Sec. (10) NIS Directive. An **exception** applies to companies with fewer than 50 employees and an annual turnover of less than ten million euros (Art. 2 Sec. (2) of the Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises, ABl EU No. L 124 of 20.5.2013, pp. 36–41): Under § 8d Sec. 4 sent. (1) BSIG, these are not subject to the IT security obligations for providers of digital services, so § 19 Sec. 4 TTDSG applies (for details on the distinction between BSIG and TTDSG: Voigt 2022, marginals 377–379).

7 Information Security Obligations Under Data Protection Law

The information security obligations which have probably been the most frequently discussed in the recent past, originate in (general) data protection law—namely the GDPR—, which sets out various requirements for internal company IT systems and procedures in order to protect personal data and imposes, in some cases, drastic fines for failure to comply with these requirements. For IoT, the GDPR is particularly relevant because **smart home providers** process **user data** on a large scale, and such personal user data must be protected by various measures in accordance with the regulations listed below. However, trust in the data protection compliance of providers is shaken repeatedly. An incident caught great attention when many voice recordings of a voice assistant were sent to another user (see the article by: Bleich 2019). In another case, a smart home alarm system contained a microphone that users were not informed about at first (see the article by: Herbig 2019).

7.1 Technical and Organisational Measures

Under Art. 32 Sec. 1 GDPR, controllers and processors are obliged to take **appropriate technical and organisational measures** to ensure a level of security appropriate to the risk involved. Under the legal definition of Art. 4 No. 7 GDPR, the controller is the person who determines the purposes and means of processing personal data. In contrast, a processor only processes personal data on behalf of the controller, Art. 4 No. 8 GDPR.

The specific **measures to be taken** must be determined in each individual case based on an **objective risk assessment**, see recital 76 sent. (2) GDPR. Under Art. 32 Sec. 1 GDPR, the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons must be considered (for details on the individual factors to be considered: Voigt 2019, marginals 17–24).

However, Art. 32 Sec. 1 lit. a-d GDPR list some exemplary **minimum requirements**. The technical and organisational measures may accordingly include the following (for details see Voigt 2019, marginals 4–14):

- the pseudonymisation and encryption of personal data
- the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services
- the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident
- a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing

An obligation to take technical and organisational measures prior to the actual data processing results from Art. 25 Sec. 1 GDPR ('**Data protection by design**') as well as from Art. 25 Sec. 2 sent. (1) GDPR ('**Data protection by default**'). These preventive protection approaches aim to minimise the risks associated with any security gaps already during system or product development, for example by ensuring that IT systems are designed to collect as little data as possible in the development stage and to directly pseudonymise or render the data anonymous at the time of their collection.

If measures required under Art. 32 Sec. 1, 25 Sec. 1, Sec. 2 sent. (1) GDPR are not implemented, a fine of up to ten million euros or—in the case of a company—of up to 2% of the total worldwide annual turnover of the previous financial year, considering whichever is the highest, can be imposed under Art. 83 Sec. 4 lit. a GDPR.

7.2 Distinction from BSIG requirements

If the controller is also a KRITIS operator, a company of special public interest or a 'provider of digital services' and thus subject to obligations under the BSIG, the scope of application of the GDPR provisions overlaps with that of the respective BSIG provisions. The relationship between the obligations to implement security measures arising from the GDPR and the BSIG has not yet been clarified. However, as the respective purposes of the different regulations—the protection of personal data on the one hand and the security of network and information systems on the other—differ, the provisions may well **apply alongside each other**.

7.3 Reporting and Documentation Obligations

In the event of a personal data breach, Art. 33 Sec. 1 sent. (1) GDPR stipulates an **obligation to immediately notify** the personal data breach to the competent supervisory authority (in principle within a maximum of 72 h), unless the personal data breach is unlikely to result in a risk to the rights and freedoms of data subjects. Irrespective of this risk prognosis, however, such incidents must **always at least be documented** under Art. 33 Sec. 5 GDPR. If the personal data breach is likely to result in a high risk to the rights and freedoms of any data subject, Art. 34 Sec. 1 GDPR also requires the personal data breach to be communicated to the data subject without delay (for details on the information obligations in the event of data protection violations: von dem Bussche 2019).

The implementation of the obligation to notify without delay can only be effective—and above all timely—in practice if there are appropriate standard operating procedures on handling data breaches in place that define for all employees how they should behave in the face of a (suspected) data protection breach (for a list of issues to be addressed in a policy on handling data protection breaches: Voigt 2022, marginal 231).

8 Further Sector-Specific Information Security Law

Moreover, there are a number of area-specific information security regulations—for example, in the Telecommunications Act (TKG), the Social Security Code Volume V (SGB V), the Insurance Supervision Act (VAG) or the Banking Industry Act (KWG)—whose requirements are similar, at least in part, to those of the BSIG. In addition, the Energy Industry Act (EnWG) and the Act on the Peaceful Uses of Nuclear Energy and Protection against its Hazards (AtG) also contain information security requirements. However, these industry specific regulations are only of marginal interest to IoT and will therefore not be discussed here (for more details see Voigt 2022, marginals 383–507).

References

Bitkom (2018) Spionage, Sabotage und Datendiebstahl – Wirtschaftsschutz in der Industrie, Studyreport 2018. https://www.bitkom.org/Bitkom/Publikationen/Wirtschaftsschutzstudie-2018.html

Bleich H (2019) Alexa, wer hat meine Daten? Amazon gibt intime Sprachdateien preis, c't – Magazin für Computertechnik, issue 01/2019, pp 16–18

Brammsen J (2016) "Durchlöcherter" Bestandsschutz – Wirtschaftsgeheimnisse im 21. Jahrhundert, ZIP 2016, 2193

Bundesamt für Sicherheit in der Informationstechnik (BSI; 2018) Die Lage der IT-Sicherheit in Deutschland 2018. https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/bsi-lagebericht-2018.pdf?__blob=publicationFile&v=3

Cisco (2018) Visual Networking Index: Forecast and Trends, 2017–2022, Germany - 2022 Forecast Highlights. https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-highlights/pdf/Germany_2022_Forecast_Highlights.pdf

German Bundestag (2006) printed matter 16/3078 of 23 October 2006

German Bundestag (2015) printed matter 18/4096 of 25 February 2015

Herbig D (2019) Googles Alarmanlage hat ein Mikrofon, von dem niemand wusste. https://www.heise.de/newsticker/meldung/Googles-Alarmanlage-hat-ein-Mikrofon-von-dem-niemand-wusste-4313669.html

Meyer R (2015) Intelligente Stromzähler: Einfallstor für Hacker. https://politik-digital.de/news/smart-meter-einfallstor-hacker-147655/

Scherschel F (2018) Bericht: Winzige Chips spionierten in Cloud-Servern von Apple und Amazon. https://www.heise.de/security/meldung/Bericht-Winzige-Chips-spionierten-in-Cloud-Servern-von-Apple-und-Amazon-4181461.html

Schober M (2018) Regulierung für Digitale Dienste, BSI-Magazin 2018/01, p 12 f

Spindler G (2007) Verantwortlichkeiten von IT-Herstellern, Nutzern und Intermediären (study commissioned by the BSI). https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Studien/ITSicherheitUndRecht/Gutachten_pdf.pdf?__blob=publicationFile&v=2

Spindler G (2014) In: Münchener Kommentar zum Aktiengesetz, § 91 AktG, 4th editon, Munich 2014

Spindler G, Schmitz P (2018) Telemediengesetz, § 13 TMG, 2nd editon, Munich 2018

Steigerwald M (2018) Smart Home - Smart Hack: Wie der Weg ins digitale Zuhause zum Spaziergang wird, lecture available at: https://media.ccc.de/v/35c3-9723-smart_home_-_smart_hack

Voigt P (2022) IT-Sicherheitsrecht, 2nd edition, Köln 2022

Voigt P (2019) Technischer Datenschutz, in: von dem Bussche/Voigt, Konzerndatenschutz, 2nd editon, Munich 2019, Part 5 Chapter 3

Voigt P, Herrmann V, Grabenschröer J (2019) Das neue Geschäftsgeheimnisgesetz – praktische Hinweise zu Umsetzungsmaßnahmen für Unternehmen, Betriebs-Berater 4/2019, pp 142–146

Von dem Bussche A (2019) Informationspflichten bei Datenschutzverletzungen, in: von dem Bussche/Voigt, Konzerndatenschutz, 2nd edition, Munich 2019, Part 2 Chapter 7

Contract Law 4.0

Torsten Körber and Carsten König

Contents

1 Introduction

Legislators typically respond to technical innovations rather reluctantly. This also applies to the driving forces behind Industry 4.0, such as digitisation, automation and interconnection, cloud computing, the development of cyber-physical systems (CPS), and big data. The lack of relevant regulations results in both opportunities and risks for **contract design**. On the one hand, companies may enjoy a wide scope for developing creative and flexible solutions that truly reflect their business interests. **Freedom of contract** is often hardly restricted as specific regulatory requirements are not yet in force and the general framework does not properly relate to the

The authors thank Hannah Brand for her support in translating and updating this chapter. All weblinks were last accessed on 30 December 2021.

T. Körber (✉) · C. König
Chair for Civil Law, Competition Law, Regulatory Law, Law of the Digital Economy, University of Cologne, Cologne, Germany
e-mail: ls-koerber@uni-koeln.de; koenig@ls-koerber.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_11

novel circumstances. On the other hand, the legal toolkit provided by **statutory provisions** (e.g. the types of contracts described in the German Civil Code—Bürgerliches Gesetzbuch, BGB) often does not adequately reflect the situation of the parties. This can make the contract design complex and costly, especially since the burden of finding fair and reliable solutions for their contractual relations lies solely with the parties.

Against this background, this chapter first explores typical subjects of contract design in Industry 4.0 (see Sect. 2), before pointing out important restrictions on the freedom of contract through mandatory law (Sect. 3). Finally, the perspective is turned around and it will be examined how Industry 4.0 might affect the future development of contract law (Sect. 4). The focus of our discussion in this chapter is on Industry 4.0 in the narrower sense, i.e. on intelligent production environments in which manufacturing plants and logistics systems largely operate themselves without human intervention (so-called Smart Factories).[1]

2 Contract Design in Industry 4.0

The contracts to be concluded in Industry 4.0 are as diverse as industry itself. In many cases, the classic types of contracts defined in the German Civil Code will at best provide a rough basis for highly complex, largely **customised contracts**, which can often only be classified as mixed-type contracts or contracts *sui generis* (Heuer-James et al. 2018, p. 2822 f.). Thus, it is neither possible nor appropriate to discuss in this chapter the details of specific contract types or the items that should be dealt with when contracts are drafted. Rather, we highlight several topics that deserve special attention in the context of contract design.

a) ***Distribution of tasks***. The increasing automation of industrial processes can lead to companies cooperating more closely with each other. For example, a manufacturer of machine tools may no longer just sell the equipment as such, but may also offer their customers IT and data analysis services that can be used to optimise the production process (e.g. the production of metal parts).[2] The provision of these services may require the involvement of other companies (e.g. software developers or providers of cloud services). This may result in the creation of entire **value networks** whose members are typically linked to each other by **long-term contracts** (e.g. service contracts). Some fear that with increasing automation and interconnection, companies' respective areas of responsibility could become blurred (Kuß 2017, para. 2; Forschungsunion Wirtschaft - Wissenschaft und acatech 2013, p. 63), so that it would no longer be possible to clearly identify breaches of duty and assign liability. Where this

[1] Wikipedia, "Smart Factory" (http://t1p.de/feqq).

[2] An example is the platform TruConnect of the Trumpf company (https://t1p.de/3z8p).

danger is imminent, the specific distribution of tasks (e.g. responsibility for software updates, IT security, etc.) should be regulated by contract.

b) ***Contractual obligations.*** The contracts that are concluded in the Industry 4.0 often do not correspond to **contract types** of the German Civil Code. Examples include complex software development or software adaptation agreements, contracts for the use of online platforms or cloud computing contracts. In such cases, it is important that the parties specify their respective rights and obligations in the contract (Kuß 2017, para. 15), because it may be difficult, if not impossible, to develop appropriate solutions from the applicable statutory provisions. Particularly in the case of **software agreements**, it is important to describe the performance owed by the software developer as specifically as possible, e.g. by listing all the relevant requirements in a so-called software requirements specification. If such a performance specification is missing or incomplete, courts in Germany typically rely on the 'medium standard of performance' which is applicable to ordinary contracts for work and services (e.g. German Federal Court of Justice – Bundesgerichtshof, BGH, Judgment of 16 December 2003 – X ZR 129/01, NJW-RR 2004, 782, 783 with further references). Even with the help of experts it is, however, very difficult to precisely determine this standard with regard to software. Thus, to dispense with a performance specification in the aforementioned sense entails considerable legal uncertainties for the parties (Conrad and Witzel 2016, para. 26). This risk can be largely prevented by a careful drafting of the contract in line with the relevant interests of the parties.

c) ***Warranty rights***. Since it is often difficult to classify contracts for IT services according to the contract types defined in the German Civil Code (Conrad and Schneider 2016, para. 46 ff.), the parties should not rely on certain statutory provisions to apply in the event of non-performance. Instead, they should contractually define breach of duty and remedies available to the other party (Kuß 2017, para. 2). This particularly applies if the supplier or service provider is supposed to also remedy cases of non-performance that occur after the transfer of the product and thereby the risk. For example, a software developer could be obligated to provide software updates for certain period after the delivery of the software (see also below, under Sect. 4.3). Under German law, defects that occur after the transfer of risk are usually not covered by the statutory warranty in the case of sales contracts (cf. § 434 (1) (1) BGB) or contracts for work and services (cf. §§ 633 (2) (1), 640 (1) (1) BGB). This is different only for rental agreements (cf. § 536 (1) (1) BGB). A contractual solution also seems advisable if the parties wish to extend the warranty beyond the statutory periods (typically two or three years, cf. § 438 (1) no. 3, § 634a (1) nos. 1, 3 BGB). It is also important that the parties specify exactly what constitutes non-performance and what does not. For example, they should stipulate in their contract whether a certain tolerance is to be applied (e.g. with regard to failure frequency, reaction times, maintenance intervals, etc.). Quality requirements can be defined via Service Level Agreements (SLA).

d) ***Liability***. Of course, conditions and limitations of liability are also important elements of contract design (see also the chapter 'Liability Law 4.0' in this book). If the contractual obligations are clearly defined (above, under b), breaches of duty can usually be detected and identified easily. With respect to responsibility or fault, the creditor can in Germany rely on the presumption of § 280 (1) (2) BGB (the debtor, however, has the right to prove exoneration). Nevertheless, it is expected that difficulties of proof could arise in highly automated and interconnected production environments (Grapentin 2018, p. 213 ff.; Horner and Kaulartz 2015, p. 512 ff.) The parties might also prefer liability rules that deviate from statutory law, e.g. to simplify the handling of claims. It is conceivable, for example, that some parties might want to lower the requirements for the proof of damage or even allow for lump-sum compensation to make complex damage assessments unnecessary. In return, the other party might benefit from liability limitations, e.g. maximum amounts for compensation. Thereby it could also be ensured that the typical damage is insurable. However, in this context, the parties must generally observe the requirements of the law on standard business, which is mandatory in Germany also for contracts between businesses (see also below, under Sect. 3.1). An important topic for contract design can also be recourse in connection with product liability claims (see chapter 'Liability Law 4.0', under 3.2.2). Finally, it seems advisable that the parties agree whether and, if so, how and by whom the relevant processes in the Smart Factory are to be logged (typically with the aid of software) and who is to have access to the log files under what conditions (Horner and Kaulartz 2015, p. 514 f. p., see also the chapter 'Liability Law 4.0', under 3.2.4).

e) ***Standards and interfaces***. The automation and interconnection of industrial production plants in cyber-physical systems (CPS), based on network technologies and software, require that all relevant components of the production environment work together seamlessly. For this reason, technical specifications such as standards and interfaces (e.g. file formats, network protocols) should be determined by contract (Kuß 2017, para. 2). This particularly applies to programming and communication interfaces, which make it possible to integrate individual computer programs into software systems. These interfaces also ensure that information can be exchanged between machines or software from different manufacturers, which is of central importance for the desired universal interconnection that characterises intelligent production environments. It should be noted, however, that the parties may have very different interests in this respect. For example, the operator of a Smart Factory may tend towards open standards because she seeks to integrate machines from different suppliers into the CPS or because she wants to be able to customise the software used. In contrast, a manufacturer of machines may prefer proprietary solutions, e.g. because he may want to offer a complete package consisting of machines, software and IT services. Obviously, any lack of interoperability can lead to dependencies (Henseler-Unger 2017, paras. 74 f.), which may also pose problems under the competition laws (see also below, under Sect. 3.2).

f) ***M2M communication***. A central element of intelligent production environments is IT-based, automated communication between equipment, e.g. between individual stations of the production plant, between these stations and the workpieces going through the plant or between the plant and the outside world (machine-to-machine, M2M). Where communication takes place between facilities or objects owned or operated by different companies, laying down the basic rules of M2M communication in a **framework agreement** is recommended (see also Groß 2018, p. 6 f.; Baum et al. 2017, p. 1890 f.). It is widely accepted that machine declarations can be considered declarations of intent *(Willenserklärungen)* in terms of §§ 116 ff. BGB and are regularly attributable to the machine user. This is because the user decides to use the machine and therefore initiates the submission of the declaration at least indirectly (see below, under Sect. 4.2). Regulating this topic by contract could still eliminate any remaining uncertainties and create legal certainty. With the appropriate contract design, even highly automated or autonomous declarations issued by **software agents** using self-learning algorithms and/or artificial intelligence, do not need to cause any problems. For example, in sophisticated M2M environments, it often makes sense for the parties to adopt rules on the (non-)attribution of declarations of intent in the event of software errors ('mistakes' of the software agent) as well as on the nullity or revocation of declarations in such cases. Example: A operates a Smart Factory and agrees with supplier B that A may also place orders via M2M communication. The processes on both sides are highly automated, without any human being involved in either the placement or receipt of orders. The parties agree that, in principle, A is bound without restriction by all declarations made by her software agents, but that she may freely revoke orders placed with B as long as they are still marked as 'unprocessed' in the interconnected ordering system of A and B.

g) ***Data ownership/data access rights***. Data property or a comparable exclusive right is to date not recognised by established law.[3] In Germany, the statutory provisions concerning the ownership of things (cf. § 90 BGB) cannot be applied to data by analogy (Grützmacher 2016, p. 492; Boehm 2016, p. 381; for a different view see Hoeren 2013). A certain protection of industrial data is offered by § 823 (1) BGB in conjunction with the right to the business (*Recht am eingerichteten und ausgeübten Gewerbebetrieb*), the criminal law provisions on data protection (§§ 202a ff., 303a of the German Criminal Code – Strafgesetzbuch, StGB, also in conjunction with § 823 (2) BGB; see Hoeren 2013), the protection of business secrets (Geschäftsgeheimnisgesetz, GeschGehG,[4] transposing the EU Trade Secret Directive 2016/943, OJ EU

[3] See on this in general Denga (2018), p. 1372; Wiebe (2017, 2018); Baum et al. (2017), p. 1826 f.; Thalhofer (2017); Specht (2016); Ehlen and Brandt (2016); Dorner (2014); Peschel and Rockstroh (2014).

[4] The GeschGehG has replaced the former protection of trade secrets under §§ 17-19 of the German Act against Unfair Competition (Gesetz gegen den unlauteren Wettbewerb, UWG).

2016 L157/1) and the *sui generis* right for the creators of databases under §§ 87a ff. of the German Act on Copyright and Related Rights (Urhebergesetz, UrhG, transposing the Database Directive 1996/9/EC, OJ EC 1996 L77/20; Wiebe 2018, p. 100 ff.; Ehlen and Brandt 2016, p. 572 ff.). Apart from these specific provisions, the **factual protection** of data takes centre stage: The data owner may exclude others from using the data under her control by, for example, storing it on secure servers and protecting it against unauthorised access by third parties.[5] The fact that the law does not assign any property rights with respect to data thus favours the person who actually controls the IT systems where the data is stored (this could be a provider of cloud services, for example). If other parties shall be allowed to access and use the data (e.g. machine manufacturers, the plant operator, suppliers, customers), it seems advisable to regulate their rights by contract (Baum et al. 2017, p. 1827). Of course, different **permissions** can be granted to different persons concerning different types of data. It would be conceivable, for example, that a machine manufacturer could be allowed to gain access to the operating data of the machines they supplied (e.g. to detect malfunctions, develop updates or produce improved machines), but not to other production data (e.g. process data and data concerning the utilisation of other machines, or the production plant as a whole) or even customer data (e.g. ordering data, contact details). **Data-sharing agreements** should also contain provisions on data protection (e.g. backups), IT security, deletion rights and obligations and the transfer of data to third parties. Utmost care must be taken when defining data access and usage rights. Otherwise, business or trade secrets could fall into the wrong hands (e.g. to current or potential competitors who might already be part of the same value network at an upstream or downstream level). This would not only be undesirable from a business perspective but could also lead to problems under the competition laws (see also below, under Sect. 3.2). Furthermore, data sharing agreements may conflict with the rights of third parties (e.g. data protection or privacy rights of employees).

h) ***IT security***. IT security is also an important subject of contractual agreements in Industry 4.0. Particular challenges may arise from the fact that the members of the value network must typically cooperate with each other to be able to protect the cyber-physical system (CPS) against IT security risks (Baum et al. 2017, p. 1889). For example, a cloud service provider can only guarantee the security of servers and data storage but cannot prevent the employees of a smart factory from carelessly handling their user account credentials. The operator of the production facilities, on the other hand, can prevent unauthorised persons from gaining access to the hardware that is located on the site, but typically leaves the protection of IT systems against attacks from the outside (by hackers etc.) to the IT service providers. Especially with respect to the operation of IT

[5] Under exceptional circumstances, data access obligations may require the data owner to share his or her data with others, see König (2017, 2018) and Wiebe (2017).

systems and cloud technology, Industry 4.0 is typically based on a broad **division of labour**, which may create special risks. Contractual agreements should therefore precisely regulate who bears which responsibility with regard to IT security.

i) ***Duration and termination***. In Industry 4.0, parties will often have an interest in long-term contractual relationships. For example, if an expensive machine is supplied with a customised software package, regular software updates might be necessary to maintain the interoperability of the plant and to ensure a high level of safety. In such situations, the parties should attach great importance to balancing their respective interests with regard to the **duration of the contract** and the **conditions and modalities of termination** (Kuß 2017, paras. 60 f.). In the example, the buyer of the machine may be interested, on the one hand, in being provided with updates for a certain minimum period, but, on the other hand, she may fear too much dependency if she relies on only one software solution (especially if it is proprietary). The fact that contracts in Industry 4.0 are often economically linked may also influence the parties' interests, for example because the machine in the above hypothetical is just as useless without the software as, conversely, the software is without the machine (see also below, under Sect. 4.3). In such cases, it may be appropriate **to legally link** the contracts. For example, the existence of one contract could be conditioned on the existence of another contract or the parties might agree on rights of revocation, withdrawal or termination that can be exercised not only in relation to the contract in which they were agreed, but also if there is a problem with another contract.

3 Limits to the Freedom of Contract

The lack of specific legal requirements for software agreements, M2M communication, data rights, etc. leaves the parties considerable scope for contractual arrangements. Protected by their freedom of contract rooted in fundamental rights, the parties can basically agree on what they want. Limits result from **mandatory law** which, according to the will of the legislator, is not at the disposal of the parties. In Germany, mandatory law includes almost all administrative law, especially regulatory law; §§ 134, 138 BGB on the nullity of contracts due to illegality or violations of public order; the law on standard terms and conditions (§§ 305-310 BGB); consumer protection law; and large parts of property law. As a matter of principle, the parties are also not allowed to deviate from the protective provisions of data protection law (see the chapter 'Data Protection 4.0' in this book), collective employment law and competition law. In the following, we discuss in more detail the law of general terms and conditions as well as competition law, which both are particularly relevant for Industry 4.0.

3.1 Law and Standard Terms and Conditions

In large parts of the literature, the German law on standard terms and conditions in §§ 305-310 BGB is seen as a big obstacle for contract design in Industry 4.0 (e.g. Groß 2018, p. 7 ff.; Faber 2017, paras. 64 ff.; Baum et al. 2017, p. 1892 f.). However, in our view, the problem should not be overestimated. While it is true that the statutory provisions on the review of standard business terms are a **mandatory part of German contract law** (even if no consumers are involved), the parties can opt out of German contract law altogether under Article 3 (1) of the Rome I Regulation (Regulation (EC) No. 593/2008, OJ EC 2008 L 177/6) (restrictions apply particularly to consumer contracts, see Article 6 of the Rome I Regulation). Nevertheless, smaller companies may shy away from the effort and the cost of applying a contract law that is less familiar to them than the rules laid down in the German Civil Code. These companies then have no choice but to comply with the narrow limits of the German law on standard terms and conditions, as described below.

As already indicated, the **scope of application** of the German law on terms and conditions is broader than that of comparable frameworks. This follows particularly from § 310 (1) BGB, which stipulates that the judicial review of standard business terms also applies to **contracts between businesses**. Thus, deviating from the Unfair Term's Directive (Directive 93/13/EEC, OJ EC 1993 L95/19), German law on standard terms and conditions is not limited to consumer contracts. Under § 305 (1) (1) BGB, standard business terms are 'all contract terms pre-formulated for more than two contracts which one party to the contract (the user) presents to the other party upon the entering into of the contract'. § 305 (1) (3) BGB clarifies that 'contract terms do not become standard business terms to the extent that they have been negotiated in detail between the parties'. The requirements for what constitutes a **detailed negotiation** are high, as the courts require that the user must seriously put the proposed clause up for discussion and must provide the other party with the realistic possibility of influencing the content of the clause.[6] Even in the case of contracts with **high transaction values,** these requirements will often not be met. The main reason is the great complexity of commercial contracts, in which there is often considerable interaction between the individual clauses. This can make it difficult to find a new balance when a single clause is to be changed. Overall, while it is possible to escape the scope of application of the law on standard terms and conditions, it will not be worth the effort in many cases. Contract design in Industry 4.0 will therefore often be subject to judicial review under §§ 305-310 BGB.

Limitations and exclusions of liability in standard business terms prove to be particularly problematic. Companies often seek to increase the efficiency of settling

---

[6] BGH, Judgment of 22 November 2012 – VII ZR 222/12, NJW 2013, 856 f., paras. 10 ff.; BGH, Judgment of 19 May 2005 – III ZR 437/04, NJW 2005, 2543, 2544; BGH, Judgment of 23 January 2003 – VII ZR 210/01, BGHZ 153, 311, 321.

their claims and might, for example, want to simplify the assessment of damages or reduce the requirements for proof. At the same time, they may want to limit compensation to certain maximum amounts to ensure an appropriate distribution of risk. The German law on standard terms and conditions often stands in the way of such efforts. § 309 no. 7a BGB prohibits any limitation of liability in standard business terms with regard to damages resulting from injury to life, body or health. For other types of injuries, § 309 no. 7b BGB stipulates that the parties must not exclude in standard terms any liability arising from gross negligence. For intentional breaches of duty, the same already results from § 276 (3) BGB, which is also applicable to individually negotiated contracts. Although the prohibitions laid down in §§ 308, 309 BGB do not apply directly to contracts between companies under § 310 (1) (1) BGB, the courts assume based on settled case law that similar or even identical restrictions may follow from the general clause in § 307 (1), (2) BGB. In particular, the catalogues listed in §§ 308, 309 BGB may have an indicative effect for the determination of what constitutes an 'unreasonable disadvantage' in terms of § 307 (1), (2) BGB.[7] With respect to § 309 nos. 7a and 7b BGB, the courts have explicitly stated that limitations of liability for personal injuries or gross negligence are also inadmissible in standard business terms between companies (BGH, Judgment of 19 September 2007 – VIII ZR 141/06, BGHZ 174, 1, 5, paras. 13 ff.) Further obstacles arise from the fact that courts typically regard exemptions from liability as conflicting with § 307 (2) no. 2 BGB and thus ineffective if they concern so-called cardinal obligations (*Kardinalpflichten*).[8] These are contractual duties 'which make the performance of the contract possible in the first place, so that the other party may therefore legitimately rely on their fulfilment' (BGH, Judgment of 20 July 2005 – VIII ZR 121/04, BGHZ 164, 11, 36). They include above all (but not only—see also Wurmnest 2019, para. 74 with further details) the main performance obligations. The case law on cardinal obligations is guided, among other things, by the consideration that it would be contradictory to assume duties of great importance for the other party, but not the resulting responsibility. However, not every limitation of liability in standard business terms with regard to significant obligations is automatically ineffective (as is rightly pointed out by Kähler 2020, paras. 125 ff. with further references). The case law leaves some room for differentiation to consider the legitimate interests of the parties. For example, the parties may be able to limit liability with regard to certain types of damage (Kähler 2020, paras. 215 ff.; Boehm 2016, p. 377 f.) or set maximum amounts of compensation (Kähler 2020, paras. 227 ff.), as long as these restrictions do not endanger the purpose of the contract. Nevertheless, utmost care must be taken when formulating liability clauses in standard business terms.

[7] BGH, Judgment of 19 September 2007 – VIII ZR 141/06, BGHZ 174, 1, 4 f., para. 12; BGH, Judgment of 3 March 1988 – X ZR 54/86, BGHZ 103, 316, 328 f.; BGH, Judgment of 3 August 1984 – VII ZR 349/82, BGHZ 90, 273, 278 f.

[8] This is settled case law, see e.g. BGH, Judgment of 20 July 2005 – VIII ZR 121/04, BGHZ 164, 11, 36; BGH, Judgment of 24 October 2001 – VIII ARZ 1/01, BGHZ 149, 89, 95 f.; BGH, Judgment of 27 September 2000 – VIII ZR 155/99, BGHZ 145, 203, 244.

Obstacles for contract design in Industry 4.0 also arise with respect to clauses other than liability clauses. For example, § 309 No. 12 BGB prohibits clauses which change the burden of proof to the disadvantage of the contractual partner particularly because they require her to prove circumstances which lie within the user's area of responsibility (lit. a). The BGH has extended this prohibition via § 307 (1) (1) BGB to standards terms used in contracts between businesses (Judgment of 5 October 2005 – VIII ZR 16/05, BGHZ 164, 196, 207 with further references). On cloud computing contracts, Boehm has argued convincingly that imposing far-reaching **obligations to cooperate** on the other party may conflict with § 307 (2) no. 2 BGB, because the purpose of cloud computing contracts is often the outsourcing of IT resources (Boehm 2016, p. 378 f.) and this purpose would be jeopardised if the contracting party had to continue to take her own precautions. Finally, parties must, of course, also observe the general clause of § 307 (1) (1) BGB, under which standard business terms are void if they unreasonably disadvantage the other party to the contract. Because an important part of the judicial review of standard business terms is a weighing of the parties' interests, the validity of specific clauses can ultimately only be assessed by considering the circumstances of the individual case. For a possible reform of the German law on standard terms and conditions, see below, under Sect. 4.1.

3.2 Competition Law

Furthermore, there are mandatory laws that cannot be opted out of by choosing another jurisdiction under the principles of private international law. These laws can be considerable obstacles for contract design in Industry 4.0, even for larger companies. An important example is competition law (see also Frenz 2016 and the chapter 'Competition Law 4.0' in this book), which is typically applicable to all restraints of competition that affect markets within the geographical scope of the respective competition law (so-called 'effects doctrine'; for German competition law see § 185 (2) of the German Act against Restraints of Competition – Gesetz gegen Wettbewerbsbeschränkungen, GWB; for EU competition law see European Court of Justice (ECJ), Judgment of 6 September 2017 – C-413/14 P, paras. 40 ff. – Intel).

Article 101 (1) TFEU and § 1 GWB prohibit, among other things, agreements and concerted practices which have as their object or effect a restriction of competition, and which are not eligible for exemption under Article 101 (3) TFEU or § 2 GWB. **Horizontal cooperation** between companies at the same market level is considered to be particularly problematic. This must be considered in Industry 4.0 contracts, for example, if such contracts facilitate the creation of value networks. Possible exemptions are described in the European Commission's Guidelines on horizontal cooperation agreements ('Horizontal Guidelines', OJ EU 2011 C 11/1). Examples include the joint development of technical norms and standards ('standardisation agreements'; Horizontal Guidelines, paras. 257 ff.; see also Plattform Industrie 4.0 2018, p. 36) or horizontal cooperations with regard to production, joint purchases or commercialisation (Horizontal Guidelines, paras. 150 ff., 194 ff., 225 ff.; see also Plattform Industrie 4.0 2018, p. 37 ff.) Legally binding block exemption regulations exist for specialisation agreements (Regulation [EU] No. 1218/2010, OJ EU 2010 L 335/43) and research and development agreements (Regulation [EU] No. 1217/2010, OJ EU 2010 L 335/36; Horizontal Guidelines, paras. 111 ff.). The exemptions defined in these regulations also apply to the German prohibition of anticompetition agreements under § 2 (2) GWB. If a company wants to be sure that the competition authority will not take action against a certain practice, it can ask for a non-intervention decision under § 32c GWB, the scope of which has been considerably extended by the 2021 amendments to the GWB.

Vertical cooperation between companies at different market levels is considered to be less problematic under competition law, as such cooperation often leads to efficiency gains. Vertical agreements are therefore exempted more generously under the so-called Vertical Block Exemption Regulation (Regulation [EU] No. 330/2010, OJ EU 2010 L 102/1, see also the Commission's Guidelines on Vertical Restraints, OJ EU 2010 C 130/1). Problems may arise, for example, when long-term dependencies are created. This is quite likely in Industry 4.0, where long-term agreements usually play an important role. For example, machines might be offered with extensive software packages or IT services (e.g. cloud services or data analysis), which may practically tie the buyer to the seller/manufacturer of the machine or highly specialised IT services providers for a longer period. Where such agreements have the effect of restricting competition (e.g. through so-called **lock-in effects**, see also Plattform Industrie 4.0 2018, p. 39; Baum et al. 2017, p. 1831), the parties should consider that Article 3 (1) of the Vertical Block Exemption Regulation makes the exemption conditional on the relevant market shares not exceeding 30%. Furthermore, Article 5 (1) (a) in conjunction with Article 1 (1) (d) of the Regulation excludes long-term commitments with a duration of more than five years from the exemption (which means that they are again illegal) if they concern more than 80% of the buyer's total purchases.

Further competition law barriers to the drafting of contracts in Industry 4.0 exist with regard to the **exchange of data and information** (Horizontal Guidelines, paras. 55 ff.). In particular, companies in a horizontal relationship may infringe Article 101 TFEU or § 1 GWB if they exchange strategic data, e.g. data relating to prices, production costs, quantities, sales, capacities, risks or investments (Horizontal Guidelines, paras. 58, 65 ff.). It should be noted that it is the European Commission's understanding that a horizontal relationship may already exist between potential competitors (Horizontal Guidelines, paras. 1, 10; see also Plattform Industrie 4.0 2018, p. 31 ff.). The ECJ has clarified in the Eturas case that even unilateral contacts via an online platform (e.g. a message sent by the platform operator to all users) may justify a presumption of collusive behaviour if the recipients do not distance themselves sufficiently from the proposal received (Judgment of 21 January 2016 – C-74/14, paras. 26 ff.). A lack of dissociation may even justify the conclusion that the users have coordinated their behaviour horizontally. The platform operator can then be held responsible as an 'accomplice' because Article 101 TFEU and § 1 GWB do not require the undertakings to be active in one of the markets concerned by the anticompetitive collusion (ECJ, Judgment of 22 October 2015 – C-194/1 P, paras. 26 ff. – AC Treuhand). Platform operators must therefore ensure through appropriate design of the platform and, if necessary, additional precautions (organisational unbundling, Chinese walls, etc.), that no unlawful exchange of information can take place on the platform (so-called compliance by design; Plattform Industrie 4.0 2018, p. 21). The German Federal Cartel Office (Bundeskartellamt) has recently emphasised these principles in two cases concerning B2B trading platforms, which are intended to facilitate communication and business transactions between suppliers and customers. In the *ECEMENT* case, the Federal Cartel Office took action to ensure that the platform operator does not offer platform users price information, thus reducing the risk that the platform would serve as a **market information system** restricting competition (Federal Cartel Office, press release of 7 December 2017). In the *XOM Metals* case, the platform operator was part of a group of companies and some of the affiliated companies were themselves active as suppliers in markets that the platform also addressed (Federal Cartel Office, case report of 27 March 2018 – B5-1/18-001). This created a risk that the corporate group to which the platform operator belonged would gain access to strategic information of its competitors via the platform. The Federal Cartel Office has sought to reduce this danger by a series of measures, including an organisational, structural and personal separation of the platform operator from other members of the group.

Where one or more undertakings with a dominant position on the relevant market (or relative or superior market power) are involved, additional rules on unilateral conduct may have to be observed in addition to the prohibitions on restrictive agreements. In particular, §§ 19-21 GWB and Article 102 TFEU do not allow such companies to abuse their market power by imposing disadvantageous contractual terms on dependent business partners.

4 Changes in Contract Law: Contract Law 4.0

Some parts of contract law may be a challenge for Industry 4.0—especially with regard to contract design—but Industry 4.0 is also a challenge for contract law. On the one hand, there are significant shifts within the system—for example, long-term agreements increasingly replace or supplement contracts for one-off exchanges and complex contractual networks take the place of individual contracts. On the other hand, emerging technologies such as artificial intelligence and machine learning raise the question whether the current framework for drafting and concluding contracts is still up-to-date or needs adjustments. Much of the current literature on this topic focusses on exposing perceived regulatory gaps and requesting amendments. In many cases, however, a closer assessment of the underlying issues reveals that there is actually no need for change, at least for now. Instead, contract law proves to be very adaptable in many respects. This section examines three reform discussions by way of example.

4.1 *Standard Terms and Conditions*

The plans for changing the German law on general terms and conditions with respect to **contracts between businesses** are quite advanced. Large parts of the business community consider the case law of the BGH to be too strict, at least as far as it relates to contracts other than consumer contracts (Bitkom et al. 2018). This refers particularly to the high requirements that the BGH places on 'detailed negotiations' in terms of § 305 (1) (3) BGB—one of the criteria that determine whether a particular clause is a standard term. There is also much criticism of the *de facto* application of the prohibitions in §§ 308, 309 BGB to business transactions, which the BGH bases on an extensive interpretation of § 307 (1) (1) BGB, although § 310 (1) (2) BGB expressly states that the catalogues of §§ 308, 309 BGB do not apply between companies. This strict application of the current law by the courts is seen as a major problem for cooperation in Industry 4.0. (Plattform Industrie 4.0 2016, p. 5 f.). A joint initiative by several business associations (*Frankfurter Initiative zur Fortentwicklung des AGB-Rechts*) has therefore proposed to **reduce the judicial control** of standard business terms with regard to contracts between companies (Frankfurt Initiative 2018). The Association of German Jurists (*Deutscher Juristentag*) also called for changes at their convention in 2012 (German Jurists Forum 2013, Part I, p. 90) and reaffirmed this position in 2016 (German Jurists Forum 2017, Part K, p. 197). In the academic literature, supporters of reform also seem to be in the majority (see, among others, Wicker 2014, p. 789; Kaufhold 2012; Berger 2010; for a different view see Basedow 2019, paras. 16 ff. with further references). Leuschner has presented a large-scale study on behalf of the German Federal Ministry of Justice and Consumer Protection (Bundesministerium für Justiz und Verbraucherschutz, BMJV) (Leuschner 2014), in which he argues in favour of legislative changes. He has proposed to explicitly allow liability caps in standard business terms, and to abandon the control of standard terms for contracts with a transaction value of 1 million euros or more, if they are concluded between businesses (Leuschner 2015a, b; for a different view see Westphalen 2015). The parties supporting the former federal government had agreed in their coalition agreement to 'review the law on general terms and conditions for contracts between companies with the aim of improving legal certainty for innovative business models' (coalition agreement of 12 March 2018, paras. 6186 ff., translation by the authors; for a critical view see Westphalen 2018). However, no specific proposal was made. The coalition formed at the end of 2021 does not appear to be pursuing this plan any further.

4.2 *Machine Declarations*

The classification of declarations made by intelligent software agents is one of the most discussed legal issues in the context of Industry 4.0. A practical example would be a fully automated and interconnected production plant in which machines can order materials and other resources from suppliers without the involvement of humans. Due to the fully autonomous ordering processes, it may be questionable whether and, if so, how machine declarations may be attributed to human beings, for example the plant operator or her employees. In cases of very high levels of automation based on artificial intelligence and machine learning, it may even become questionable whether a (human) **declaration of intent** within the meaning of §§ 116 ff. BGB still exists if a machine acts completely independently of human will. However, in our view, the practical relevance of these topics should not be overestimated. As far as typical industry matters are concerned, it is very likely that the parties will be in long-term contractual relations. They will often be linked through **framework agreements** in which they lay down the basic rules of their M2M communication, among other things (see Sect. 2 above). Industrial companies cannot afford to order material 'somewhere on the Internet' as they want to keep stocks low for cost reasons, but at the same time must avoid production downtimes as much as possible. Thus, they typically work with a pool of pre-qualified suppliers from which offers are requested and to whom contracts are awarded. Since a business relationship typically already exists before the first order is placed, it is usually not that difficult to agree on how to deal with machine declarations. For example, the parties can use their existing contractual relationship to agree on the conditions under which machine declarations are accepted or to define the situations in which such declarations may be revoked.

Nevertheless, it is foreseeable that the discussion about machine declarations will also influence **general legal doctrine**. Most of the literature published so far concludes that declarations of automated or autonomous software agents are attributable to the user of the software (e.g. to the owner of a smart factory), even outside existing contractual relationships.[9] Various explanations are offered, as described below.

Most academics conclude that machine declarations are **declarations of intent** made in the name of the **person who uses the software agent**.[10] The subjective element of the declaration, insofar as it is considered necessary at all, is usually seen in the deliberate use of the software that makes the machine declaration (Kitz 2018, para. 51; Spindler 2015, para. 9; Gitter 2007, p. 181; John 2007, p. 74; Cornelius 2002, p. 355). Some authors draw a parallel to computer declarations (Faber 2017, paras. 45 ff.; Spindler 2015, para. 9; Sester and Nitschke 2004, p. 550 f.; Cornelius 2002, p. 355), which are automatically generated by IT systems based on predefined

[9] This is by far the dominant view in the literature, see e.g. Faber (2017), paras. 36 f.; Plattform Industrie 4.0 (2016), p. 7; Spindler (2015), para. 9; Bräutigam and Klindt (2015), p. 1138; Horner and Kaulartz (2015), p. 502 f.; Spindler (2014), p. 64 f.; Medicus (2010), para. 256 and the further references in the subsequent footnotes. For a very critical view on the prevailing opinion, see Teubner (2018), p. 177 ff. Cf. also BGH, Judgment of 16 October 2012 – X ZR 37/12, BGHZ 195, 126, 131, para. 17.

[10] In the case of companies, the situation is more complicated because the declaration of intent must be attributed to the company owner, i.e. typically to a legal person. This generally requires that the declaration can be attributed to a natural person authorised to represent the company.

parameters and data entered by the user. It is settled case law and supported by the literature that such declarations are best understood as prepared human declarations that are merely assisted by technology and for which automation has no formative role (Busche 2018, para. 38; Säcker 2018, para. 189; Singer 2017, para. 57; Köhler 1982, p. 132 ff.). However, this view is contradicted by other authors, at least with respect to declarations made by autonomous software agents whose decisions are no longer predetermined. These authors argue that the previous view no longer fits when decisions are made by intelligent algorithms because the declaration can then no longer be attributed to the will of the user (Grapentin 2018, p. 91; Specht and Herold 2018, p. 42 f.; Nitschke 2010, p. 57 f.; John 2007, p. 100 f.; Bauer 2006, p. 78 ff.; for a sceptical perspective at the traditional view see also Schulz 2015, p. 104). This objection, however, does not sufficiently consider that, according to the prevailing view in Germany, declarations of intent are to be interpreted from the perspective of an objective recipient (*objektiver Empfängerhorizont*). Thus, in cases of doubt the understanding of a neutral observer (the objective appearance of the declaration) prevails over the intent of the person making the declaration (the subjective element of the declaration).[11] However, the objective appearance of a machine declaration will typically clearly refer to the company using the autonomous software agent. For example, if the intelligent ordering software of the X AG orders 25,000 screws from the Y GmbH (for which a unique business identifier will regularly be used), Y will clearly attribute the declaration to X, regardless of whether it was made by a machine or a human being. Just as in cases where a person lacks the consciousness of making a legal declaration (the textbook example being the 'Trier wine auction'; see also BGH, Judgment of 7 June 1984 – IX ZR 66/83, BGHZ 91, 324 ff), it must suffice for the attribution of machine declarations that users of software agents could typically recognise that potential recipients will reasonably assume that the declarations originate from the users.[12]

Viewed from this angle, the attribution of machine declarations raises no greater problems than the attribution of human declarations. A **statutory clarification** in the German Civil Code, as suggested by Working Group 4 of Plattform Industrie 4.0 (Plattform Industrie 4.0 2016, p. 7), is therefore not absolutely necessary, but may be useful to dispel last doubts and increase legal certainty.[13]

---

[11] This is settled case law since BGH, Judgment of 7 June 1984 – IX ZR 66/83, BGHZ 91, 324 ff.; see also BGH, Judgment of 5 October 2006 – III ZR 166/05, NJW 2006, 3777, 3778; BGH, Judgment of 29 November 1994 – XI ZR 175/93, NJW 1995, 953; BGH, Judgment of 2 November 1989 – IX ZR 197/88, BGHZ 109, 171, 177.

[12] For a different view see John (2007), p. 75 f. This author believes, among other things, that the contractual partner is not worthy of protection because he usually knows that the declaration is made by a software agent. The decisive question, however, is whether the contractual partner can recognise that the user of the software agent does not want the agent's declaration to be valid as his own declaration of intent, which is typically not the case.

[13] See also Schirmer (2016), p. 664 who proposes an amendment following the model of Section 14 of the US Uniform Electronic Transaction Act of 1999.

Authors who advocate attributing declarations according to **areas of responsibility** also emphasise the goal of protecting *bona fide* recipients (Wiebe 2002, p. 216 ff., 237 ff.; Nitschke 2010, p. 63 ff.; for critical views see Grapentin 2018, p. 89 f.; Schulz 2015, p. 105) and/or seek to hold the user responsible for the **legal appearance** (*Rechtsschein*) that has been created (see in more detail Schulz 2015, p. 113 ff.; Nitschke 2010, p. 61 ff.; John 2007, p. 110 ff.). The latter authors essentially conclude that the user of a software agent who is able to make legal declarations to third parties is responsible for such declarations because of her control over the software agent. This is because the user could observe what the software agent is doing and could prevent declarations from being made in her name. Ultimately, this line of reasoning also includes those who believe that the agent's declaration should be treated according to the principles that apply to blank deeds (*Blanketturkunden*).[14] According to settled case law of the Federal Court of Justice (BGH), § 172 (2) BGB, according to which the holder of a signed attorney document is presumed to have power of attorney, is applicable by analogy to documents which the issuer has already signed but in which important information is still missing.[15] In fact, the parallels between blank deeds and machine declarations are striking (cf. already Köhler 1982, p. 134): Just like blank deeds, declarations by autonomous software agents seem to be compounds of still incomplete framework declarations by the user (incorporated in the programme code or settings) and their specification in the individual case by the software agent. Nevertheless, this dogmatic approach seems unsatisfactory. In view of the great practical importance that machine declarations will likely have in the future, it would be odd to base their attribution solely on principles that were created for rare and unusual cases. Therefore, the previously discussed approach of objective attribution seems preferable.

Only a minority of authors want to consider machine declarations as **declarations of intent of the software agent** and attribute them to the user only under the **rules of agency**. Most proponents of this view advocate for applying §§ 164 ff. BGB by analogy,[16] whereas a minority within the minority assumes a **partial legal capacity** of the software agent and applies the rules of agency directly.[17] It speaks in favour of these views that the software agent actually forms the relevant intent, and not the

[14] Grapentin (2018), p. 91 ff.; Schulz (2015), p. 109 ff.; Gitter (2007), p. 181 f.; John (2007), p. 102 ff., 121 f.; Sester and Nitschke (2004), p. 550 f.; Gitter and Roßnagel (2003), S. 66; for a similar view see also Wettig (2010), p. 171 f.; for different views see Nitschke (2010), p. 44 f.; Bauer (2006), p. 79.

[15] Settled case law since BGH, Judgment of 11 July 1963 – VII ZR 120/62, BGHZ 40, 65, 67 ff.; see also BGH, Judgment of 29 February 1996 – IX ZR 153/95, NJW 1996, 1467, 1469; BGH, Judgment of 20 November 1990 – XI ZR 107/89, BGHZ 113, 48, 53.

[16] Teubner (2018), p. 177 ff., esp. 181 ff.; Gruber (2012), p. 154 ff.; Teubner (2006), p. 14 ff. (both argue in favour of partial legal capacity, but nevertheless want to apply agency law only by analogy); for critical views see Grapentin (2018), p. 94 ff.; Faber (2017), paras. 38 ff.; Mayinger (2017), p. 70 ff.; Wettig (2010), p. 179 ff.; John (2007), p. 83 ff.; Bauer (2006), p. 68 ff.

[17] Specht and Herold (2018), p. 43; Schirmer (2016), p. 664. For critical views see Grapentin (2018), p. 93 f.; Mayinger (2017), p. 70; Schulz (2015), p. 106 f.; Gitter (2007), p. 177 ff.; John (2007), p. 77 ff.; for comments from a policy perspective see Sorge (2006), p. 118 f.

user on whose behalf the agent is operated (e.g. it is the agent who decides whether to enter into a sales agreement and who determines the order quantity and the delivery time). Those who seek to strengthen the subjective element of declarations of intent, may find a good starting point here. However, if one simply focusses on the **legal consequences** of §§ 164 ff. BGB, the agency solutions do not seem particularly useful or practical (for a similar view see Gitter 2007, p. 179: 'unnecessarily complicated'). Above all, it remains unclear what benefits could be achieved by these approaches.[18] As already described, there is a variety of approaches for the direct attribution of machine declarations to the software agent's user. Therefore, there is simply no need for attributing such declarations via § 164 (1) (1) BGB. The otherwise important rules on whether the agent acts on her own behalf or in the name of the principal cannot serve a meaningful purpose in this context because the software agent cannot be a contracting party anyway. The same problem arises with regard to §§ 164 (2), 179 (1) BGB, which normally would create obligations on the part of the agent under certain conditions. Because it does not make sense to impose legal duties on software agents, these provisions do not fit the specific context of machine declarations. This is also conceded by supporters of agency solutions (Teubner 2018, p. 182 ff.). Holding the software agent personally liable under § 179 (1) BGB would be just as pointless as considering the agent as a contracting party of the agreement. At least according to the current legal situation, a software agent cannot have any assets and therefore cannot be a debtor of claims for damages in any meaningful way.[19]

Finally, it seems doubtful whether there would be any benefit in applying § 166 (1) BGB, which states that, for the purposes of agency law, determinations of intent or knowledge must focus on the agent, and not the principal. This is particularly important for the assessment of mistakes and their legal consequences. However, it seems questionable whether it would make sense to regard errors made by software agents as mistakes within the meaning of this provision. It is not at all clear, which errors should count as a mistake of content (*Inhaltsirrtum*) or a mistake of declaration (*Erklärungsirrtum*) in terms of § 119 BGB (this scepticism is shared by Gitter 2007, p. 178). Moreover, the errors that typically occur in the field of automation (e.g. incorrect data recording by sensors or malfunctions due to programming errors, hardware damage or external attacks) will often already affect the **decision-making process**. Under German law, this would mean that they would

[18]This view is shared by Spindler (2016), p. 816. Teubner (2018), p. 177 ff. and Schirmer (2016), p. 663 f. mainly take issue with the fact that the actual will of the user of the software agent hardly plays a role for the majority because objective attribution replaces the subjective element of the declaration. However, even if one were to assume that the software agent can make its own declarations, these declarations could ultimately only be constructed objectively. This becomes apparent in Teubner (2018), p. 182 ff.

[19]Grapentin (2018), p. 98; Faber (2017), para. 39; Bräutigam and Klindt (2015), p. 1138; Wettig (2010), p. 179 ff.; Gitter (2007), p. 178 f.; John (2007), p. 87 f.; Bauer (2006), p. 70; Sester and Nitschke (2004), p. 550; Cornelius (2002), p. 355; see also Kuhn (1991), p. 66. On possible solutions see Mayinger (2017), p. 71 f. with further references.

have to be classified as mistakes of motive (*Motivirrtümer*), which do not create a right of avoidance (Gitter 2007, p. 199 f.; Medicus 2010, para. 256; cf. also Köhler 1982, p. 135). At most, one could therefore apply § 166 (1) BGB to errors by a software agent that somehow look like human mistakes or otherwise correspond to them. However, a nuanced theory of mistakes by software agents would still need to be developed in detail. Schirmer cites the example of a shopping robot that, due to a sensor error, 'makes a mistake' and puts fish instead of meat in the shopping cart (Schirmer 2016, p. 664). Here, the application of § 166 (1) BGB could indeed lead to a mistake of content in terms § 119 (1) BGB, because the robot declares at the register that it wants to buy fish, although it actually 'wants' to buy meat. However, it is not clear whether it would be appropriate in such a situation to allow the user of the shopping robot to avoid the sales contract. This is not self-evident because it would mean that the **risks** of using intelligent software agents would be partly imposed on the recipients of machine declarations. Against this it can be argued that 'mistakes' made by software agents can often be prevented by technical means (in the example, the robot could additionally scan the product's barcode). Furthermore, it is primarily the users who can ensure a proper functioning of software agents (e.g. through maintenance, but also indirectly by exerting pressure on manufacturers, software developers etc. through their buying behaviour). A right of avoidance would reduce the **incentives** for users to avoid mistakes. Even if the policy discussion on these issues has only just begun (see e.g. Grapentin 2018, p. 99), it can therefore be concluded for the time being that there is no need for applying agency law to software agents.

4.3 Contract Networks

A further challenge for contract law is seen in the fact that contracts in Industry 4.0 are often economically linked to other contracts. For example, a new machine may only function with additional IT services (e.g. cloud services, data analysis) and it may therefore be necessary to conclude further agreements (e.g. service agreements) in addition to the sales contract (see e.g. Spindler 2018, p. 47). Similar situations may arise in B2C relationships, e.g. with regard to so-called 'wearables', such as smart watches or fitness trackers, for which end user license agreements (EULA) are typically concluded in addition to the sales contract even if the additional services are offered free of charge (Heuer-James et al. 2018, p. 2822 ff.; Börding et al. 2017, p. 136 ff.; Bräutigam and Klindt 2015, p. 1138).

In view of the increasing interconnection of contracts and the emergence of contract networks, some authors already proclaim a 'departure from the bilateral understanding of contracts' (see e.g. Wendehorst 2016, p. 2610). In our view, this is exaggerated. On the one hand, contract networks are not a new phenomenon (see e.g. Grundmann 2007; Teubner 2004; Heermann 1998; Rohe 1998; Möschel 1986). On the other hand, contract networks typically consist of classic **bilateral contracts**, which are only economically and not legally connected with each other. However, the economic interactions between the individual contracts can be quite considerable, especially in the context of Industry 4.0. For example, it is discussed whether it may constitute a material defect in terms of § 434 (1) BGB if the purchased good (e.g. a machine) stops to work because a company that is not itself a party to the contract ceases to provide a necessary service (Börding et al. 2017, p. 137; Bräutigam and Klindt 2015, p. 1138). An example could be that necessary IT services are no longer provided or that necessary software is not updated. In such cases, most authors reject the existence of a defect, at least with regard to the current legal framework (Schrader and Engstler 2018, p. 357; Solmecke and Vondrlik 2013, p. 757). One of the reasons is that § 434 (1) BGB requires such a defect to be present already when the risk passes from the seller to buyer, which usually happens when the purchased good is delivered (§ 446 BGB). For the same reason, it is typically concluded that the seller is not required to provide software updates (Schrader and Engstler 2018, p. 356 et seq.; cf. however, for the future legal framework, Article 8(2) of Directive (EU) 2019/770 on contracts for the supply of digital content and digital services, OJ EU 2019 L 136/1, and Article 7(3) of Directive (EU) 2019/771 on contracts for the sale of goods, OJ EU 2019 L 136/28). Some authors, however, stipulate (e.g. by relying on an extensive interpretation of the contractual agreement) that the seller of a smart device may have a secondary obligation to ensure the functionality of the device for a reasonable time even if (as is usually the case) the seller does not provide the necessary IT services herself (Regenfus 2018, p. 81 f.; for a critical view on such approaches see Schrader and Engstler 2018, p. 357 f.). Where no such right can be established, the customer must assert his claims against the IT service providers, etc. This again shows that it is of utmost importance that contracts in Industry 4.0 are carefully drafted (e.g. with regard to duration, performance obligations, warranty). It is also discussed whether contracts can be so closely linked with each other that the voidness of one contract can justify a claim for adjustment or cancellation of the other contract. Grundmann has shown that such situations can best be resolved in accordance with the rules on contract adjustment in the event of substantial changes in the circumstances underlying the contract (*Störung der Geschäftsgrundlage*), § 313 BGB (for details see Grundmann 2007, 2008, p. 741 ff.).

The above examples show that it is typically possible to find solutions for the challenges posed by contract networks based on the **established rules and principles**. Whether these solutions are always convincing as a matter of policy remains to be discussed. Legislative changes are most likely to be expected in the B2C area. With regard to Industry 4.0, it is particularly important that contracts are carefully drafted and that they account for interactions that may arise between individual agreements in a network of contracts.

5 Conclusion

In view of the above considerations, we are confident that the existing contract law is well prepared for the current and future developments in Industry 4.0. It gives companies considerable leeway, but also requires them to take their affairs into their own hands. Much will therefore depend on whether companies succeed in finding appropriate contractual solutions that adequately consider the interests of all the parties involved. As we have underscored, contract design raises difficult questions, for example with regard to data rights, M2M communication and IT security. It is therefore important that the parties exercise utmost care when drafting agreements for Industry 4.0. Mandatory law can prove to be a significant obstacle for contract design. We have highlighted the law on general terms and conditions and competition law as important examples. The latter may, inter alia, prohibit certain data exchanges and make cooperation in Industry 4.0 more difficult. However, it will usually be possible to find arrangements that adequately reflect the interests of the parties without unlawfully restricting competition. Overall, it seems fair to say that contract law 4.0 is characterised above all by the growing complexity of contract design, but does not in principle stand in the way of technical innovations in Industry 4.0.

References

Basedow J (2019) Commentary on § 310 BGB. In: Säcker FJ et al (eds) Münchener Kommentar zum Bürgerlichen Gesetzbuch, vol 2, 8th edn. C.H. Beck, Munich

Bauer MC (2006) Elektronische Agenten in der virtuellen Welt. Dr Kovač, Hamburg

Baum F, Appt S, Schenk IK (2017) Die vernetzte Fabrik: Rechtliche Herausforderungen in der Industrie 4.0. Der Betrieb (DB) 1824-1832 (part 1), 1888-1893 (part 2)

Berger KP (2010) Für eine Reform des AGB-Rechts im Unternehmensverkehr. Neue Juristische Wochenschrift (NJW):465–470

Bitkom et al (2018) AGB-Recht für Unternehmen modernisieren. Accessible at http://t1p.de/1pdr

Boehm F (2016) Herausforderungen von Cloud Computing-Verträgen: Vertragstypologische Einordnung, Haftung und Eigentum an Daten. Zeitschrift für Europäisches Privatrecht (ZEuP):358–387

Börding A, Jülicher T, Röttgen C, Schönfeld M (2017) Neue Herausforderungen der Digitalisierung für das deutsche Zivilrecht. Computer und Recht (CR):134–140

Bräutigam P, Klindt T (2015) Industrie 4.0, das Internet der Dinge und das Recht. Neue Juristische Wochenschrift (NJW):1137–1142

Busche J (2018) Preliminary remarks to §§ 145 ff. BGB. In: Säcker FJ et al (eds) Münchener Kommentar zum Bürgerlichen Gesetzbuch, vol 1, 8th edn. C.H. Beck, Munich

Conrad I, Schneider J (2016) Vertragliche Grundlagen. In: Auer-Reinsdorff A, Conrad I (eds) Handbuch IT- und Datenschutzrecht. C.H. Beck, Munich, pp 367–414

Conrad I, Witzel M (2016) IT-Projektverträge. In: Auer-Reinsdorff A, Conrad I (eds) Handbuch IT- und Datenschutzrecht. C.H. Beck, Munich, pp 721–802

Cornelius K (2002) Vertragsabschluss durch autonome elektronische Agenten. MultiMedia und Recht (MMR):353–358

Denga M (2018) Gemengelage privaten Datenrechts. Neue Juristische Wochenschrift (NJW):1371–1376
Dorner M (2014) Big Data und "Dateneigentum". Computer und Recht (CR):617–628
Ehlen T, Brandt E (2016) Die Schutzfähigkeit von Daten – Herausforderungen und Chancen für Big Data Anwender. Computer und Recht (CR):570–575
Faber T (2017) Vertragsschluss beim IoT Rechtsgeschäft. In: Sassenberg T, Faber T (eds) Rechtshandbuch Industrie 4.0 und Internet of Things. C.H. Beck, Munich, pp 239–266
Forschungsunion Wirtschaft – Wissenschaft und acatech (2013) Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Accessible at http://t1p.de/hlnm
Frankfurt Initiative (2018) Gesetzgebungsvorschlag zur Änderung des AGB-Rechts im unternehmerischen Geschäftsverkehr. Accessible at http://t1p.de/72bh
Frenz W (2016) Industrie 4.0 und Wettbewerbsrecht. Wettbewerb in Recht und Praxis (WRP):671–678
German Jurists Forum (2013) Verhandlungen des 69. Deutschen Juristentages, vol II/1. C.H. Beck, Munich
German Jurists Forum (2017) Verhandlungen des 71. Deutschen Juristentages, vol II/2. C.H. Beck, Munich
Gitter R (2007) Softwareagenten im elektronischen Geschäftsverkehr. Nomos, Baden-Baden
Gitter R, Roßnagel A (2003) Rechtsfragen mobiler Agentensysteme im E-Commerce. Kommunikation & Recht (K&R):64–72
Grapentin J (2018) Vertragsschluss und vertragliches Verschulden beim Einsatz von Künstlicher Intelligenz und Softwareagenten. Nomos, Baden-Baden
Groß J (2018) AGB 4.0: Allgemeine Geschäftsbedingungen im Rahmen autonomer Vertragsschlüsse. Zeitschrift für Innovations- und Technikrecht (InTeR):4–9
Gruber M-C (2012) Rechtssubjekte und Teilrechtssubjekte des elektronischen Geschäftsverkehrs. In: Beck S (ed) Jenseits von Mensch und Maschine. Nomos, Baden-Baden, pp 133–160
Grundmann S (2007) Die Dogmatik der Vertragsnetze. Archiv für civilistische Praxis (AcP) 207: 718–767
Grundmann S (2008) Vertragsnetz und Wegfall der Geschäftsgrundlage. In: Aderhold L et al (eds) Festschrift für Harm Peter Westermann zum 70. Geburtstag. Otto Schmidt, Cologne, pp 227–244
Grützmacher M (2016) Dateneigentum – ein Flickenteppich. Computer und Recht (CR):485–495
Heermann PW (1998) Drittfinanzierte Erwerbsgeschäfte. Mohr Siebeck, Tübingen
Henseler-Unger I (2017) Auf dem Weg zur Industrie 4.0. In: Sassenberg T, Faber T (eds) Rechtshandbuch Industrie 4.0 und Internet of Things. C.H. Beck, Munich, pp 1–26
Heuer-James J-U, Chibanguza KJ, Stücker B (2018) Industrie 4.0 – vertrags- und haftungsrechtliche Fragestellungen. Betriebs-Berater (BB):2818–2832
Hoeren T (2013) Dateneigentum. Versuch einer Anwendung von § 303a StGB im Zivilrecht. MultiMedia und Recht (MMR):486–491
Horner S, Kaulartz M (2015) Rechtliche Herausforderungen durch Industrie 4.0: Brauchen wir ein neues Haftungsrecht? – Deliktische und vertragliche Haftung am Beispiel "Smart Factory". In: Taeger J (ed) Internet der Dinge. OlWIR Publishing Company, Edewecht, pp 501–518
John R (2007) Haftung für künstliche Intelligenz. Dr. Kovač, Hamburg
Kähler L (2020) Commentary on § 307 BGB, keyword "Haftungsklausel". In: Gsell B et al (eds) beck-online.Grosskommentar zum Zivilrecht, Stand 1 October 2020. C.H. Beck, Munich
Kaufhold S (2012) "Echte" und "unechte" AGB in der Klauselkontrolle. Betriebs-Berater (BB):1235–1241
Kitz V (2018) Vertragsschluss im Internet. In: Hoeren T, Sieber U, Holznagel B (eds) Multimedia-Recht, 46th edn. C.H. Beck, Munich, Teil 13.1
Köhler H (1982) Die Problematik automatisierter Rechtsvorgänge, insbesondere von Willenserklärungen. Archiv der civilistischen Praxis (AcP) 182:126–172

König C (2017) Der Zugang zu Daten als Schlüsselgegenständen der digitalen Wirtschaft. In: Hennemann M, Sattler A (eds) Immaterialgüter und Digitalisierung. Nomos, Baden-Baden, pp 89–103
König C (2018) Towards a data sharing economy: the legal framework for access to data. In: Veenbrink M, Looijestijn-Clearie A, Rusu CS (eds) Digital markets in the EU. Wolf Legal Publishers, Oisterwijk, pp 175–192
Kuhn M (1991) Rechtshandlungen mittels EDV und Telekommunikation. C.H. Beck, Munich
Kuß C (2017) Vertragstypen und Herausforderungen der Vertragsgestaltung. In: Sassenberg T, Faber T (eds) Rechtshandbuch Industrie 4.0 und Internet of Things. C.H. Beck, Munich, pp 291–332
Leuschner L (2014) AGB-Recht für Verträge zwischen Unternehmen, Abschlussbericht. Accessible at http://t1p.de/cy7i
Leuschner L (2015a) Reformvorschläge für die AGB-Kontrolle im unternehmerischen Rechtsverkehr. Zeitschrift für Wirtschaftsrecht (ZIP):1045–1052
Leuschner L (2015b) Noch einmal: Reformvorschläge für die AGB-Kontrolle im unternehmerischen Bereich. Zeitschrift für Wirtschaftsrecht (ZIP):1326–1332
Mayinger SM (2017) Die künstliche Person. Deutscher Fachverlag, Frankfurt/Main
Medicus D (2010) Allgemeiner Teil des BGB, 10th edn. C.F. Müller, Heidelberg
Möschel W (1986) Dogmatische Strukturen des bargeldlosen Zahlungsverkehrs. Archiv für civilistische Praxis (AcP) 186:187–236
Nitschke T (2010) Verträge unter Beteiligung von Softwareagenten – ein rechtlicher Rahmen. Peter Lang, Frankfurt/Main
Peschel C, Rockstroh S (2014) Chancen und Risiken neuer datenbasierter Dienste. MultiMedia und Recht (MMR):571–576
Plattform Industrie 4.0 (2016) Industrie 4.0 – wie das Recht Schritt hält. Accessible at http://t1p.de/cucs
Plattform Industrie 4.0 (2018) Industrie 4.0 – Kartellrechtliche Betrachtungen. Accessible at http://t1p.de/21dv
Regenfus T (2018) Deaktivierte Geräte – Rechte des Nutzers bei Einschränkung der Funktionsfähigkeit "smarter" Produkte durch den Hersteller. Juristenzeitung (JZ):79–87
Rohe M (1998) Netzverträge. Mohr Siebeck, Tübingen
Säcker FJ (2018) Introduction. In: Säcker FJ et al (eds) Münchener Kommentar zum Bürgerlichen Gesetzbuch, vol 1, 8th edn. C.H. Beck, Munich
Schirmer J-E (2016) Rechtsfähige Roboter? Juristenzeitung (JZ):660–666
Schrader PT, Engstler J (2018) Anspruch auf Bereitstellung von Software-Updates? MultiMedia und Recht (MMR):356–361
Schulz T (2015) Verantwortlichkeit bei autonom agierenden Systemen. Nomos, Baden-Baden
Sester P, Nitschke T (2004) Software-Agent mit der Lizenz zum ...? Vertragsschluss und Verbraucherschutz beim Einsatz von Softwareagenten. Computer und Recht (CR):548–554
Singer R (2017) Preliminary remarks to §§ 116 ff. BGB. In: Staudingers Kommentar zum Bürgerlichen Gesetzbuch. de Gruyter, Berlin
Solmecke C, Vondrlik S-E (2013) Rechtliche Probleme bei Produkten mit serverbasierten Zusatzdiensten. MultiMedia und Recht (MMR):755–760
Sorge C (2006) Softwareagenten. Universitätsverlag Karlsruhe
Specht L (2016) Ausschließlichkeitsrechte an Daten – Notwendigkeit, Schutzumfang, Alternativen. Computer und Recht (CR):288–296
Specht L, Herold S (2018) Roboter als Vertragspartner? MultiMedia und Recht (MMR):40–44
Spindler G (2014) Zivilrechtliche Fragen beim Einsatz von Robotern. In: Hilgendorf E (ed) Robotik im Kontext von Recht und Moral. Nomos, Baden-Baden, pp 63–80
Spindler G (2015) Preliminary remarks to §§ 116 ff. BGB. In: Spindler G, Schuster F (eds) Recht der elektronischen Medien, 3. Aufl. C.H. Beck, Munich
Spindler G (2016) Digitale Wirtschaft – analoges Recht? Braucht das BGB ein Update? Juristenzeitung (JZ):805–816

Spindler G (2018) Zukunft der Digitalisierung – Datenwirtschaft und Unternehmenspraxis. Der Betrieb (DB):41–50
Teubner G (2004) Netzwerk als Vertragsverbund. Nomos, Baden-Baden
Teubner G (2006) Elektronische Agenten und große Menschenaffen: Zur Ausweitung des Akteursstatus in Recht und Politik. Zeitschrift für Rechtssoziologie (ZfRSoz) 27:5–30
Teubner G (2018) Digitale Rechtssubjekte? Archiv für civilistische Praxis (AcP) 218:155–205
Thalhofer T (2017) Recht an Daten in der Smart Factory. Gewerblicher Rechtsschutz und Urheberrecht. Praxis im Immaterialgüter- und Wettbewerbsrecht (GRUR-Prax) 9:225–227
Wendehorst C (2016) Die Digitalisierung und das BGB. Neue Juristische Wochenschrift (NJW):2609–2613
Westphalen F (2015) Schwellenwert und Neuregelung einer vertraglichen Haftungsbegrenzung – ist das der "Königsweg" einer Reform des AGB-Rechts? Zeitschrift für Wirtschaftsrecht (ZIP):1316–1326
Westphalen F (2018) Reformvorhaben der Koalition zum AGB-Recht bei innovativen Geschäftsmodellen auf dem Prüfstand. Zeitschrift für Wirtschaftsrecht (ZIP):1101–1113
Wettig S (2010) Vertragsschluss mittels elektronischer Agenten. Wissenschaftlicher Verlag, Berlin
Wicker M (2014) Haftungsbegrenzung des Cloud-Anbieters trotz AGB-Recht? MultiMedia und Recht (MMR):787–790
Wiebe A (2002) Die elektronische Willenserklärung. Mohr Siebeck, Tübingen
Wiebe A (2017) Von Datenrechten und Datenzugang – Ein rechtlicher Rahmen für die europäische Datenwirtschaft. Computer und Recht (CR):87–93
Wiebe A (2018) Schutz von Daten im Immaterialgüterrecht. In: Hornung G (ed) Rechtsfragen der Industrie 4.0. Nomos, Baden-Baden, pp 97–118
Wurmnest W (2019) Commentary on § 307 BGB. In: Säcker FJ et al (eds) Münchener Kommentar zum Bürgerlichen Gesetzbuch, vol 2, 8th edn. C.H. Beck, Munich

Liability Law 4.0

Torsten Körber and Carsten König

Contents

The authors thank Hannah Brand for her support in translating and updating this chapter. All weblinks were last accessed on 30 December 2021.

[1]This refers to intelligent production environments in which manufacturing plants and logistics systems largely operate themselves without human intervention, see Wikipedia, 'Smart Factory'.

T. Körber (✉) · C. König
Chair for Civil Law, Competition Law, Regulatory Law, Law of the Digital Economy, University of Cologne, Cologne, Germany
e-mail: ls-koerber@uni-koeln.de; koenig@ls-koerber.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_12

1 Introduction

Liability issues are undoubtedly one of the most discussed legal topics in relation to Industry 4.0. It is widely assumed that the current legal framework does not adequately reflect the special features of digitisation, automation and interconnection of industrial processes, e.g. because it is difficult to clearly identify causation and fault in increasingly complex value networks. For this reason, several proposals to amend the existing legal frameworks have been made, such as the introduction of strict liability for autonomous systems.

Against this background, this chapter examines in detail how liability frameworks for Industry 4.0 present themselves and whether and to what extent there are gaps that need to be closed on the way towards a 'Liability Law 4.0'. A distinction will be made between liability within the Smart Factory,[1] i.e. the mainly contractual liability between the members of the value network (such as the factory owner, machine manufacturers, IT service providers, and suppliers of goods and services—see Sect. 2) and the liability of the network members for the manufactured product, especially towards end customers (Sect. 3). Subsequently, the future development of liability law is discussed from a law and policy perspective (Sect. 4).

2 Liability Within the Value Network

It seems advisable that the companies that are part of the value-added network should regulate their mutual liability by contract (Chirco 2015, p. 527 ff.; see also the chapter 'Contract Law 4.0' in this book). The **potential risks** and **causes for damages** are manifold. Examples include economic losses caused by production disruptions in the Smart Factory, e.g. as the result of software errors, cyber-attacks or incorrect data collection or processing (Marschollek and Wirwas 2017, para. 44). Operating errors, incompatibilities and technical defects may cause physical harm to production equipment or working materials (Horner and Kaulartz 2015, p. 501). Inadequate IT security may result in data losses or data protection violations (Marschollek and Wirwas 2017, para. 47 ff.; Bräutigam and Klindt 2015, pp. 1140–1141), and faulty M2M communication may lead to incorrect orders and resulting supply shortages or unnecessary expenses (Horner and Kaulartz 2015, p. 501). To give a final example, autonomous systems such as self-driving vehicles or drones may cause property damage or physical injury to employees (Heuer-James et al. 2018, p. 2828). In the absence of contractual arrangements, these situations would be subject to **ordinary liability law**. In Germany, for example, §§ 280-283, 276 (1) of the German Civil Code (Bürgerliches Gesetzbuch, BGB) require the debtor to compensate the creditor for all losses the debtor causes by breaching his

[1] This refers to intelligent production environments in which manufacturing plants and logistics systems largely operate themselves without human intervention, see Wikipedia, 'Smart Factory'.

contractual duty. Liability under these provisions is contingent upon the debtor's responsibility for the breach, which is however presumed under § 280 (1) (2) BGB. Claims may also be based on tort law, particularly if there is any damage to property or bodily injury. In German law, the relevant provisions can be found in §§ 823 ff. BGB. It should be noted, however, that claims of employees against employers and claims between employees may be excluded by §§ 104, 105 of the Seventh Book of German Social Security Code (Siebtes Buch des Sozialgesetzbuchs, SGB VII) concerning statutory accident insurance. Liability in tort is fault-based unless the legislator has exceptionally ordered strict liability. The latter applies, for example, to motor vehicles (see § 7 (1) of the German Road Traffic Act, Straßenverkehrsgesetz, StVG) and drones (see § 33 (1) (1) in conjunction with § 1 (2) (2) of the German Air Traffic Act, Luftverkehrsgesetz, LuftVG).

The standard legal situation often does not fit the interests of the parties, for example because liability may already result from ordinary negligence or may be unlimited in amount. This can be remedied by **contractual solutions**, as statutory liability law in Germany is largely dispositive, which means that the parties can opt it out for their specific contractual relationship (see for contractual liability Ernst 2019, paras. 46 ff., for tortious liability Wagner 2017a, paras 87 ff.). Limitations to the parties' freedom of contract result from **mandatory law**, which is not at the discretion of the parties.[2] An example is § 14 of the German Product Liability Act (Produkthaftungsgesetz, ProdHaftG), transposing Article 12 of the European Product Liability Directive 85/374/EEC and stating that the producer's liability for defective products under the Act must not be limited or excluded by contractual agreement. Furthermore, § 276 (3) BGB stipulates that a debtor may not be released in advance from liability for intentional breaches of his or her duties. Important limitations for contractual arrangements also follow from the law on general terms and conditions. For example, § 309 nos. 7 and 12 BGB, which also serve to implement no. 1(a), (b) and (q) of Annex 3 of the European Unfair Contract Terms Directive 93/13/EEC, render ineffective certain exclusions and limitations of liability as well as certain modifications of the burden of proof (see chapter 'Contract Law 4.0' in this book, section 3.1). According to settled case law of the German Federal Court of Justice (Bundesgerichtshof, BGH) concerning the general clause of § 307 (1) BGB, these requirements must also be observed in contracts between businesses, although § 310 (1) (1) BGB seems to exclude the applicability of § 309 BGB in this respect (see chapter 'Contract Law 4.0' in this book, section 3.1). Within these limits of mandatory law, the parties may regulate their affairs however they see fit. Possible contractual arrangements concern the delimitation of areas of responsibility, standards of liability (e.g. no liability for ordinary negligence except in cases of physical injury—cf. § 309 No. 7 BGB), contractual

[2]Mandatory law in Germany includes public (regulatory) law, §§ 134, 138 BGB (stipulating the nullity of contracts in the event of a violation of a statutory prohibition or public policy), the law on standard business terms, §§ 305-310 BGB, consumer protection law, and large parts of the law of property.

penalties (cf. § 309 No. 6 BGB), maximum liability (see also chapter 'Contract Law 4.0', section 3.1) and obligations to collect evidence and share it with the other party (Horner and Kaulartz 2015, p. 512 ff.; see also below, under 3.2.4).

3 Liability for Defective Products

A distinction must be made between the liability of members of the value network towards other members and their liability towards outsiders, especially customers and third parties. The latter liability particularly applies to damage caused by defective products. In general, liability issues are influenced by the fact that industrial production processes are becoming increasingly complex during digitisation and often involve additional players (e.g. software developers, IT service providers), which can make it difficult to clearly assign responsibilities. The contractual liability of the seller and the tortious liability of the producer or manufacturer are of particular importance. The legal term 'manufacturer' is conceivably broad and can also include component suppliers and other members of the value network. In the following, we focus on products from *automated production* (Smart Factory; Industry 4.0 in the narrower sense), not so much on *automated products* (e.g. wearables, autonomous vehicles).[3]

3.1 Contractual Liability

Contractual claims, e.g. following from the non-performance of sales contracts, are essentially subject to the rules that also apply outside Industry 4.0. Despite the complexity of advanced value networks, it is usually not difficult to identify the right debtor, for example. Under ordinary circumstances, the client's contractual partner will be clearly identifiable. In Germany, §§ 434-442 BGB oblige the **seller** to remedy any non-performance of a sales contract. Typically, this will not be the manufacturer (or a member of the value network), but a distributor who sells the final product. If the product proves to be defective, the seller must, under § 437 No. 1 BGB, primarily **provide cure**. This means that the seller must repair the defect or supply a new product that is free of defects. These obligations also apply to contracts for the delivery of products which are manufactured according to the customer's specifications. Such contracts are governed by § 650 BGB, which is predicted to gain importance in Industry 4.0 (Iliou 2017, p. 847) because digitisation and automation

[3] See below Sects. 4.2 to 4.4. For details see also Wagner (2017d); Wende (2017), paras. 37 ff; Gomille (2016); Bodungen and Hoffmann (2016); Borges (2016), p. 274 ff; Kütük-Markendorf and Essers (2016); Spindler (2015).

often allow for a cost-efficient production of small quantities (down to 'batch size 1', i.e. custom-made products).

The seller's obligation to provide cure and the buyer's secondary rights to revocation and reduction of price do not depend on fault, § 437 No. 1, 2 BGB. However, the buyer may only claim **damages** from the seller under §§ 437 No. 3, 280 (1) BGB, if the seller is responsible for her breach of duty. Under § 276 (1) BGB, this normally presupposes that the seller has acted culpably (note that the burden of proof is on the seller, § 280 (1) (2) BGB, i.e. the seller must demonstrate their lack of culpability). In the typical situation that the seller has not manufactured the product (but has purchased it from the manufacturer), the seller will succeed in exculpating herself. The seller is not responsible if the defect was caused by a production error (e.g. a design or manufacturing error) and the seller was not involved in the production. Except in special circumstances, German courts do not qualify manufacturers, suppliers, etc. as persons for whom the debtor is responsible under § 278 BGB.[4] This is true even for contracts for the delivery of customised products under § 650 BGB.[5]

3.2 Product Liability/Manufacturer's Liability

The development towards Industry 4.0 will lead to major changes in the liability of companies involved in production processes for damages caused by defective products. In Germany, the **manufacturer's liability** is regulated in two ways. On the one hand, there are the special provisions of the German **Act on Liability for Defective Products** (Product Liability Act, ProdHaftG), which goes back to the EC Directive concerning liability for defective products of 1985 (Product Liability Directive, OJ EC 1985 L 210/29). The Directive is currently being reviewed, especially concerning the special challenges of new digital technologies (European Commission 2018a, p. 9 ff.; 2018b, p. 17 ff.; see also below under Sect. 4.1). On the other hand, the courts continue to base the manufacturer's liability for defective products on one of the general clauses of German tort law, **§ 823 (1) BGB**, which has already been used for this purpose before the Product Liability Directive came into force. There are some differences in substance, such as that § 1 (1) (2) ProdHaftG, in transposition of Article 9 of the Product Liability Directive, limits the liability for damage to property to items that are ordinarily intended for private use or

[4]BGH, Judgment of 18 October 2017 – VIII ZR 86/16, NJW 2018, 291, para. 24; BGH, Judgment of 29 April 2015 – VIII ZR 104/14, NJW 2015, 244, para. 13; BGH, Judgment of 2 April 2014 – VIII ZR 46/13, BGHZ 200, 337, paras. 31 f. with further references. For a different view see Grundmann (2019) with further references, also with respect to the prevailing opinion; Weller (2012), p. 2312 ff.; Schroeter (2010), p. 497 ff.

[5]BGH, Judgment of 2 April 2014 – VIII ZR 46/13, BGHZ 200, 337, paras. 33 ff.; BGH, Judgment of 21 June 1967 – VIII ZR 26/65, BGHZ 48, 118, 120 ff. For a different view see Schaub (2020), para. 65.2 with further references.

consumption, whereas § 823 (1) BGB does not know such a limitation (Hager 2009, para. 5; see also ECJ, Judgment of 4 June 2009 – C-285/08, paras. 24 ff. – Moteurs Leroy Somer). However, the courts are striving for a uniform interpretation of the relevant provisions, which allows us to consider both liability regimes together.

3.2.1 Product Defects

In general, products from a Smart Factory must meet the same requirements as all other products. Most importantly, every product placed on the market must provide the safety that a reasonable person is entitled to expect (§ 3 (1) ProdHaftG, transposing Article 6 (1) of the Product Liability Directive; Grützmacher 2016, p. 696. On liability under § 823 (1) BGB Wagner 2017b, para. 809 ff.). This safety standard is met if the product has no **design defect** or **manufacturing defect** and if the user has been carefully instructed in the use of the product. Within the scope of § 823 (1) BGB, the manufacturer is also obliged to **observe the product** after it is put into circulation (Wagner 2017b, paras. 836 ff.). Digitisation, automation and the creation of value networks in Industry 4.0 primarily affect the manufacturing process. This production process in the narrower sense runs largely without human intervention in intelligent, technology-driven production environments. However, the impact on the manufacturer's liability will probably not be great. Liability under § 1 (1) (1) ProdHaftG does not require fault but depends only on the product being **objectively defective**. Thus, it does not matter 'whether and, if so, which defects were inherent in the production process [. . .]' (BGH, Judgment of 25 February 2014 – VI ZR 144/13, BGHZ 200, 242, 246, para. 11). The specific cause of the defect within an automated production process is therefore just as irrelevant for liability as any human error that may have played a role. It is the product defect alone that triggers the liability. Furthermore, it is widely recognised that the manufacturer cannot invoke the defence of **development risks** for errors in the production process (Förster 2020, para. 60; Wagner 2017c, § 3 para. 37. On liability under § 823 (1) BGB Wagner 2017b, para. 824). That defence, which is stipulated in § 1 (2) No. 5 ProdHaftG, transposing Article 7 (e) of the Product Liability Directive, exempts manufacturers from liability for design errors that, at the time the product is put on the market, could not be discovered due to the state of scientific and technical knowledge. Liability under § 823 (1) BGB may be narrower in individual cases, as it always presupposes a breach of the duty of care (Wagner 2017b, para. 806).

3.2.2 Responsibility

Primarily, responsibility for defective products lies with the **producer** (§ 4 ProdHaftG, transposing Article 3 of the Product Liability Directive), i.e. 'everyone in whose organisational area the product was created' (BGH, Judgment of 25 February 2014 – VI ZR 144/13, BGHZ 200, 242, 248, para. 16). In Industry 4.0, as elsewhere, this is pre-eminently the operator of the production facilities, in which the final product is manufactured. This can also be a so-called assembler that merely assembles components purchased on the market into the final product (Bundestag document no. 11/2447, p. 19). In addition, liability under § 4 (1) (1) ProdHaftG may also attach to manufacturers of partial products and raw materials. However, such suppliers are only liable under § 1 (1), (3) ProdHaftG if the supplied product or material as such is defective (Wagner 2017c, § 1 para. 60 f.). Moreover, a **partial product** or **raw material** in this sense can only be assumed if the component in question is incorporated into the end product (Spickhoff 2017, paras. 8, 10), i.e. not in the case of services provided in relation to the production process (e.g. IT services, logistics; cf. Wagner 2017c, § 4 para. 10). The supply of defective software may only give rise to product liability if it is to be integrated into the end product (e.g. a smartwatch or fitness tracker) and can therefore be regarded as a partial product. It is however disputed whether software can at all be regarded as a product within the meaning of § 2 ProdHaftG (see e.g. Wagner 2017c, § 2 para. 17 ff.). In any case, liability under the ProdHaftG will not arise if the defective software is only used in the production process. With regard to § 823 (1) BGB, there is even more flexibility in determining the addressees of liability and allocating specific duties of care (see Wagner 2017b, paras. 786 ff.; Wende 2017, p. 60). Digitising and interconnecting the production process will therefore not generally lead to unclear responsibilities as long as it is clear who decides on the organisation of the Smart Factory (see, however, Wende 2017, paras. 64 ff.). In any case, the **overall responsibility** for a flawless end product is always assigned to the final manufacturer (Wagner 2017c, § 4 para. 12).

3.2.3 Recourse

If multiple manufacturers are responsible for the product (e.g. the supplier of a defective component and the final manufacturer), they are jointly and severally liable under § 5 (1) (1) ProdHaftG, transposing Article 5 of the Product Liability Directive. The **internal recourse** including rights to claim contribution from other liable parties is an important part of the contract design in value networks (see the chapter 'Contract Law 4.0' in this book, section 2 d). The main reason is that the liability of several co-producers among each other is primarily governed by their contractual agreements, § 5 (1) (2) ProdHaftG. Their respective shares of responsibility are only of secondary relevance. Furthermore, contractual agreements also allow for regulating recourse and contribution of parties who are not jointly and severally liable under § 5 ProdHaftG (e.g. a dealer who is not a manufacturer and who is neither considered as such, see § 4 (1) (2), (2) and (3) ProdHaftG). Joint and several liability for torts under §§ 823 ff. BGB is also open to contractual regulations concerning rights of contribution or recourse (see §§ 840 (1), 426 (1) BGB).

It is conceivable, for example, that the parties exempt the final manufacturer from her obligation to check the supplied components for defects under § 377 of the German Commercial Code (Handelsgesetzbuch, HGB). Instead, the supplier could assume full responsibility for the safety of the components as part of a **quality**

assurance agreement (Wagner 2017c, § 5 para. 9). However, the requirements of the law on standard business terms would need to be observed, as it also applies to contracts between companies in Germany (see chapter 'Contract Law 4.0' in this book, section 3.1). This may mean, for example, that a complete exclusion of the dealer's recourse against the final manufacturer might be ineffective, as such a provision would deviate considerably from the default rule in § 4 (3) ProdHaftG (which assigns liability to the dealer only in exceptional circumstances). It is argued in the literature that such deviation would conflict with § 307 (2) No. 1 BGB, which prohibits too far-reaching derogations from statutory provisions in standard terms (Wagner 2017c, § 5 para. 8; see also § 478 (2) BGB for the sale of consumer goods). Moreover, since 1 January 2018, additional requirements result from §§ 445a, 445b BGB. These provisions are in principle dispositive (*argumentum e contrario* to § 478 (2) BGB, see Arnold 2020, para. 161) but may nevertheless influence the control of standard business terms under §§ 307 ff. BGB. At least this is advocated for certain situations, e.g. imbalances of power (Arnold 2020, para. 164 ff.; Westphalen 2015, p. 2889; Orlikowski-Wolf 2018, p. 363 f. with further references).

3.2.4 Questions of Evidence

The **burden of proof** in cases covered by the Product Liability Act (ProdHaftG) is regulated in § 1 (4) of the law, transposing Article 4 of the Products Liability Directive. Under the first sentence of § 1 (4) ProdHaftG, the injured party must prove the product defect, the damage, and the causal relationship between the defect and the damage (see more detailed Wagner 2017c, § 1 para. 72). Under the second sentence of the provision, the manufacturer bears the burden of proof for the defences under § 1 (2) and (3) ProdHaftG (resembling Article 7 of the Product Liability Directive). On the manufacturer's liability under § 823 (1) BGB, the burden of proof is distributed in almost the same way. Yet, because the latter provision requires fault, the concept of defect is divided into two parts: the objective defectiveness of the product on the one hand, and the manufacturer's lack of due care on the other (Wagner 2017b, paras. 858 ff.). While the injured party must prove the objective defect,[6] the courts have assigned the burden of proof that reasonable care was taken to the manufacturer. The reason is that the cause of the product defect cannot usually be identified without insight into the production process, which the plaintiff does not have.[7]

[6] BGH, Judgment of 30 April 1991 – VI ZR 178/90, BGHZ 114, 284, 296; BGH, Judgment of 17 March 1981 – VI ZR 191/79, BGHZ 80, 186, 196 f.; BGH, Judgment of 19 June 1973 – VI ZR 178/71, NJW 1973, 1602, 1603; BGH, Judgment of 26 November 1968 – VI ZR 212/66, BGHZ 51, 91, 97 ff.

[7] BGH, Judgment of 11 June 1996 – VI ZR 202/95, NJW 1996, 2507, 2508; BGH, Judgment of 17 March 1981 – VI ZR 191/79, BGHZ 80, 186, 196 f.; both following the leading case BGH, Judgment of 26 November 1968 – VI ZR 212/66, BGHZ 51, 91, 97 ff.

Once again, special challenges arise on the **internal relations** of the members of the value network. The starting point in this context is the general principle of civil procedure that whoever seeks to rely on certain legal consequences must prove that their prerequisites are met (Gottwald 2018, para. 7). In general, the digitisation of the production process can significantly improve the parties' **access to evidence**, since all relevant processes can be logged with the help of software.[8] However, it should be contractually determined who is responsible for the creation and preservation of the log files and who is allowed to access them under which conditions. It is also conceivable, for example, that the parties agree that an independent expert should evaluate the log files in the event of a dispute (for a similar proposal see Horner and Kaulartz 2015, p. 514 f.). They might also agree that the data collection should be carried out by a third-party service provider from the outset to ensure that none of the parties can manipulate the log files to their advantage. In this area, evidentiary issues overlap with questions of **IT security**, which also plays a major role, for example, in the context of the digitisation of the energy industry (see §§ 19 ff. Messstellenbetriebsgesetz, MsbG).

4 Changes in Liability Law: Liability Law 4.0

Recent years have seen a lively legal and political debate on the question of who should be liable for damages caused by automated or autonomous systems (e.g. motor vehicles, robot lawn mowers, robot vacuums, and drones, but also bodiless algorithms that manage online platforms or trade securities). The debate about the appropriate legal framework also affects the liability of manufacturers, as no realistic liability concept could do without the liability for defective products which is firmly established in current EU law. However, the discussion goes far beyond product liability, as it also concerns, for example, the liability of users and holders of autonomous systems. (Mandatory) liability insurance, liability funds and social security systems as well as other **innovative approaches** are also part of the debate. Some of the most important proposals are presented and discussed below. For this purpose, the perspective will be broadened a little, as many of the ideas described do not only affect Industry 4.0, but also other liability situations, e.g. involving consumers. We consider it advisable to regulate the most important principles in a general law so that the future legal framework for automated and autonomous systems does not disintegrate into numerous individual regulations.

[8] See also Spindler (2015), S. 772. For a skeptical view on the economic efficiency of extensive logging see Reichwald and Pfisterer (2016), p. 211 f.; Grützmacher (2016), p. 697.

4.1 Stricter Manufacturer's Liability

The discussion, especially at EU level, about adapting the current product liability law is already well advanced. The European Commission's efforts in this respect, which in all probability will result in a proposal for a **new or amended version of the Product Liability Directive**, are aimed at making the legal framework fit for the challenges created by 'emerging digital technologies' (European Commission 2018a, p. 9 f.). On the Directive's **definition of product**, the Commission emphasises that the boundaries between products and services are becoming increasingly blurred and that a clarification may be particularly required for software (European Commission 2018b, p. 18; 2018c, p. 61). Due to the narrow wording of Article 2 of the Product Liability Directive, it is controversial whether the Directive applies to software (cf. Wagner 2017c, § 2 para. 20 with further references). The prevailing opinion on services, however, is that they are not subject to the Directive (Oechsler 2018, § 2 para. 41 ff.; Wagner 2017c, § 2 para. 1). The Commission also believes that the **definition of 'manufacturer'** should be reviewed because products are becoming increasingly complex and can often be modified or supplemented with additional (software) functions after they are placed on the market (European Commission 2018b, p. 18; 2018c, p. 61). Even more important than the definition might be whether the new or amended Directive should impose obligations on manufacturers that extend beyond the time at which the product is put into circulation. For example, it might be discussed whether the producer should be obliged to guarantee the safety of automated products throughout their life cycle, e.g. with software updates.[9] Up to now, such obligations cannot be construed from the Product Liability Directive (or the ProdHaftG as the transposing act). However, in Germany, post-circulation duties may follow from general tort law, § 823 (1) BGB (Schrader and Engstler 2018, p. 359 with further references).

The Commission has rightly pointed out that it may be difficult for the injured person to prove the **product defect** and the **causal relationship between defect and damage** in cases of complex or adaptable products (European Commission 2018b, p. 18; 2018c, p. 61). In particular, subsequent expansions (e.g. third-party software that is installed after the product is placed on the market) may be an additional source of error for which the producer would not be liable (but perhaps the providers of the additional products or services). This raises the question whether the standard or burden of proof should be changed to the advantage of the injured person. Furthermore, the Commission apparently considers an extension of the legal interests protected (e.g. with respect to property that is intended for professional use, environmental harm, violations of personality rights including privacy rights, perhaps even pure economic losses). Further subjects of review are the so-called development clause (Article 7 (e) Product Liability Directive; § 1 No. 5 ProdHaftG) and the injured party's cost sharing of 500 euros in case of damage to property (Art. 9 lit. b

[9] See Schrader and Engstler (2018), p. 359 f.; Redeker (2017), para. 826; Gomille (2016), p. 81; Droste (2015), p. 107 f.; Orthwein and Obst (2009), p. 3 f.; Spindler (2008), p. 12.

ProdHaftRL; § 11 ProdHaftG) (European Commission 2018b, p. 18 f.; 2018c, p. 61). Therefore, it seems likely now that the review process will result in a stricter manufacturer's liability.

On the other hand, there have been calls in the literature to limit producer liability compared to the current situation. The justification given is that the introduction of new technologies such as automated or autonomous driving should not be restricted by too strict liability (Lutz 2015, p. 121). It is true that manufacturers can pass on liability risks to customers via product prices, which may have a negative impact on demand. This may well make market penetration more difficult. However, '**subsidising**' innovative products at the expense of injured parties seems arbitrary and is inferior to more targeted instruments of economic policy (e.g. real subsidies for the improvement of product security). It is also obvious that manufacturers have considerable influence on the risks posed by their products because they determine the level of care in production. From a policy perspective, liability law should therefore provide effective incentives for manufacturers (Wagner 2017d, p. 762; Sosnitza 2016, p. 772; Borges 2016, p. 279; Spindler 2015, p. 774).

In line with the considerations of the European Commission, a tightening of manufacturer's liability should therefore be considered. This particularly applies—but not only—if policy makers decide against strict liability for owners of autonomous systems (see below, under Sect. 4.2). Two examples of a possible tightening of liability shall be highlighted. First, harmonised product liability does not yet apply to damage caused to the **product itself** (Article 9 (1) (b) of the Product Liability Directive; § 1 (1) (2) ProdHaftG). Thus, if an autonomous vehicle crashes into a tree due to a defect in the vehicle control system, the owner has no claim against the manufacturer under § 1 (1) (1) ProdHaftG. Depending on the circumstances of the case, the owner might be able to assert a claim under § 823 (1) BGB, but the case-law on the limitation of non-contractual liability would need to be observed (*Weiterfresserschaden*; see Wagner 2017d, p. 723 f.). The dealer's obligation to remedy the non-performance under §§ 434, 437 BGB is only a small consolation, as it expires two years after the delivery. Of course, manufacturers could offer long-term warranties or comprehensive insurance to convince customers to buy an automated or autonomous vehicle notwithstanding the limited manufacturer's liability (for a similar consideration see Lutz 2015, p. 121). But the problem could also be solved by extending that liability accordingly and this may be even more suitable to strengthen the trust of customers than voluntary measures. What could prove to be even more problematic in times of automation and interconnectedness is that the EU's product liability regime so far only covers damage to property that is ordinarily intended for **private use** (Article 9 (1) (b) of the Product Liability Directive; § 1 (1) (2) ProdHaftG). This can lead to arbitrary results, especially regarding the question whether primarily liable owners of autonomous products such as autonomous vehicles can seek recourse from the manufacturer (see e.g. Wagner 2017d, p. 760 f.; Borges 2016, p. 280). If an autonomous vehicle, because of a defect, collides with another vehicle that is used privately, the owner can take recourse against the manufacturer under §§ 840 (1), 426 BGB, because the manufacturer is liable under § 1 ProdHaftG. If the damaged vehicle is used professionally, this possibility would only exist if the manufacturer were liable under § 823 (1) BGB as liability under § 1 (1) ProdHaftG is out of the question. If the harmonised product liability law is to provide convincing solutions in itself, an adjustment of the manufacturer's liability is therefore necessary.

The future liability framework for automated, autonomous and interconnected technologies will be closely related with questions of **insurability** as well as with questions of the scope and financing of insurance protection. These issues should therefore also be taken into account in all legal policy considerations. However, apart from the question of mandatory insurance, the focus should be on enabling appropriate contractual arrangements.

4.2 Liability of Users and Owners

An important part of the current legal policy debate concerns the question of how the liability of users and owners of autonomous systems should be structured. As far as the **user** is concerned (e.g. the driver of an autonomous vehicle), it can first be observed that automation means that it is increasingly unlikely that the user will become liable. The main reason is that errors of the automated system (e.g. due to programming errors, defective sensors, etc.) are usually neither predictable nor avoidable for the user (Horner and Kaulartz 2015, p. 509). If only the user were to be considered as a possible subject of liability, the shift towards automation could therefore lead to poorer protection for victims. A reversal of the burden of proof (see e.g. §§ 280 (1) (2) BGB, § 18 (1) (2) of the German Road Traffic Act – Straßenverkehrsgesetz, StVG) would be of little help because the user would typically succeed in proving the exoneration.[10] In some cases, however, the liability of the user is accompanied by the liability of the **owner**.[11] Where it already exists, the owner's liability is typically a strict liability independent of fault. In Germany, this applies, for example, to motor vehicles (§ 7 (1) StVG) and drones (§ 33 (1) (1) in connection with § 1 (2) (2) of the German Air Traffic Act – Luftverkehrsgesetz, LuftVG). Thus, at least in these two important fields of applying automation technology, there is no threat of liability gaps even under the current legal framework. The policy debate should therefore focus on whether the liability is properly distributed between user, owner and, if applicable, manufacturer.

[10] See also Pieper (2016), p. 197; Borges (2016), p. 273; Schrader (2015), p. 3541; for an illustrative example see OLG Frankfurt, Judgment of 14 December 2017 - 11 U 43/17, NJW 2018, 637, para. 17.

[11] It should be noted that many relevant provisions in German law (e.g. § 7 StVG and § 833 BGB) technically do not refer to the 'owner' but to the 'holder' or 'keeper' (*Halter*), i.e. the person who bears the costs and reaps the benefits of the dangerous thing or animal and who may decide about its use. For simplicity, however, we will speak only of the owner, who in the vast majority of cases is also the holder or keeper.

The European Commission currently weighs the advantages and disadvantages of creating an EU-wide **strict liability framework for owners of autonomous systems** (European Commission 2018b, p. 19 ff.). Many academics and practitioners are also in favour of such an approach (Schirmer 2016, p. 665; Spindler 2016, p. 816; 2015, p. 775 f.; Horner and Kaulartz 2015, p. 507 ff.; apparently also Riehm 2014, p. 114). Indeed, strict liability would have many advantages, as the current system of motor vehicle liability shows. It guarantees a strong **protection of victims**, since the owner is typically easy to identify and the requirements for successfully claiming damages are rather low (König 2017, p. 329 f.; Borges 2016, p. 278). To establish liability under § 7 (1) StVG, for example, the injured party only has to prove the damage, the operation of the opposing vehicle, and the causal relationship between the vehicle's operation and the damage (instead of many see Burmann 2020, para. 28). In addition, strict liability provides important **incentives** to prevent damage. While the owner will normally not be able to influence the precise functioning of the autonomous system, he or she will still decide on the system's use as such and on the frequency of use (so-called activity level; König 2017, p. 330 f.; Borges 2016, p. 278; Spindler 2015, p. 775). It is therefore important that the liability framework incentivises the owner to use the system mindfully.

Strict liability is usually statistically well calculable because its requirements tend to be simple and strict liability frameworks typically limit the maximum amount of compensation. The resulting predictability is a basic condition for the **insurability** of associated liability risks. Strict liability for autonomous systems could be combined with a compulsory insurance system, again following the examples of motor vehicle liability (§ 1 of the German Obligatory Car Insurance Law – Pflichtversicherungsgesetz, PflVG; see also Article 3 of the Motor Insurance Directive 2009/103/EC, OJ EC 2009 L 263/11) and liability for drones (§ 43 (2) (1) LuftVG in connection with § 1 (2) (2) LuftVG as well as §§ 101 ff. of the German Air Traffic Licensing Regulation – Luftverkehrszulassungsordnung, LuftVZO; for more details see e.g. Schäfer 2017). Such a system of compulsory insurance could ensure that liability risks from emerging technologies would not unequally affect individual operators but would be distributed among all owners of objects covered by the scheme. The involvement of insurers could also professionalise the handling of claims, which would be an advantage especially to **recourse claims** against the manufacturer(s) of the autonomous system (end manufacturers, but also suppliers, e.g. of faulty control software or sensors, see Sect. 3.2.2 above).

Some authors fear that a strict owner liability could hinder the—socially desirable—market penetration of autonomous systems because potential buyers might shy away from liability risks (for considerations in this direction, see e.g. Hanisch 2014, p. 36; cf. also Gless and Janal 2016, p. 571 f.; Albrecht 2005, p. 374 f.). It is true that liability law can 'subsidise' activities or technologies by allocating the costs of damages resulting from errors or malfunctions to the group of injured parties. However, utmost care is required when assessing such **distributional effects** from a policy perspective, as they are often difficult to determine and are sometimes offset by countervailing effects (König 2017, p. 331 f.).

For example, any assessment of the distributional effects of introducing new liability rules for the owners or operators of autonomous systems would need to account for interactions with other existing liability frameworks. In particular, one must consider the implications of **producer liability**, which is often relevant with regard to typical sources of error in autonomous systems (programming errors, defective hardware, insufficient IT security, etc.; for more details see above, under Sects. 3.2 and 4.1). On the one hand, this means that, as long as the product liability regime is not weakened (which seems very unlikely at present—see already above, under Sect. 4.1), the damage costs cannot be assigned to the injured parties, because the injured parties have the right to take action against the manufacturer (§ 1 (1) (1) ProdHaftG, § 823 (1) BGB). On the other hand, liability risks for owners under a new liability rule for autonomous systems may be lower than they appear at first sight, because the owner could often take recourse against the manufacturer (§§ 840 (1), 426 BGB). Even with these considerations, however, the distribution effects are not yet sufficiently described. The manufacturer will not remain inactive either but will include the expected damage costs or the costs of product liability insurance in the **product price**. As the product price is usually paid by the owner (assuming he or she is also the owner, which is often the case, but not always), this means that the owner would normally bear the typical damage costs anyway even without the introduction of a new strict liability rule. The reason is simply that, even with perfect competition, all costs of value creation (including the costs of liability and insurance) are inevitably passed on to the end customer. However, for many automation technologies (e.g. automated and autonomous vehicles), it is expected that automation will increase the level of safety. Thus, the overall costs from damages should be lower than they are today (Spindler 2015, p. 775), which under competitive conditions *ceteris paribus* leads to lower product prices and insurance premiums. This is likely to favour the market penetration of autonomous systems.

The previous reflections on distributional effects show that the primary purpose of creating a strict liability rule for the owners of autonomous systems should not be seen in the allocation of additional liability risks or costs. The advantages of such a system lie in its **practicability**, which promises considerable efficiency gains compared to alternative models. In particular, the combination of a clear liability regime with rather low requirements and maximum amounts of compensation facilitates the **professionalisation of claims settlement**, especially through the participation of insurance companies. In addition, strict liability of easily identifiable persons ensures a high level of victim protection.

In the academic literature, it is sometimes considered whether a new liability rule for owners of autonomous systems should include a due diligence defence as it already exists in Germany with respect to the liability of professional animal keepers (§ 833 (2) BGB; Bräutigam and Klindt 2015, S. 1139), building owners (§ 836 (1) (2) BGB; Grützmacher 2016, S. 698) or principals for harm caused by their vicarious agents (§ 831 (1) (2) BGB; Riehm 2014, p. 114). Yet, such a possibility of exoneration would make the settlement of claims more complicated and might even jeopardise insurability without providing any significant benefit. In a typical situation, the owner would probably try to exonerate herself by pointing to the manufacturer's area of responsibility as the origin of the damage (or to the areas of responsibilities of the manufacturer's suppliers). Indeed, this is where most defects in autonomous systems are likely to be caused. Again, however, the cost and price effects mentioned above would need to be considered. Because the costs of producer liability are ultimately passed on to the users via product prices, little would be gained. The **costs of the damage** will be socialised in the group of owners either through insurance premiums or through product prices. Because of the lower complexity, the total costs of settling damage claims through a system of strict liability and mandatory insurance are likely to be lower compared to a system that allows for due diligence defences. Furthermore, a solution that relies heavily on insurance would have the additional advantage that liability costs could be distributed among owners and users according to their respective levels of activity (i.e. higher premiums could be charged for intensively used autonomous systems, whereas lower premiums could be charged for systems which are only rarely used; on this aspect see also Wagner 2017d, p. 763 f.). Other individual circumstances, e.g. a particularly prudent behaviour of the user in the past could also be considered in the design of insurance contracts. Such differentiation according to the intensity of use or the user's individual situation could be reached with the help of graduated tariff structures as they are already known today from motor vehicle liability insurance.

There is, however, considerable leeway in the design of strict liability rules and the details still need to be discussed. For example, it must be clarified whether and, if so, which **liability exclusions** should be accepted. The German Road Traffic Law, for example, which also applies to automated and autonomous cars, provides for restrictions if an accident is due to force majeure (§ 7 (2) StVG) or an unavoidable event (§ 17 (3) (1) StVG). Through appropriate 'fine-tuning', the liability for autonomous systems can be appropriately located in the wide spectrum between liability based on causality and liability based on fault.

This is also recognised by the European Commission, which is considering whether and, if so, to what extent it should influence the owner's liability that he or she could have avoided the damage (European Commission 2018b, p. 19). Other issues that are relevant in this context are **the standard and the burden of proof**, which can be used to further differentiate the legal allocation of risks among the owner and the injured persons. As already explained, however, there are good reasons for not making the liability framework too complex but keeping it simple and predictable (e.g. efficiency of claims settlement, insurability).

The Commission has also launched a discussion on the **interests** that should be protected by the new strict liability rule (European Commission 2018b, p. 20). Classical cases such as the liability of animal keepers (§ 833 BGB), vehicle holders (§ 7 (1) StVG) and aircraft holders (§ 33 (1) LuftVG) limit the protected interests to damage to life, body and property. Of course, the legislator is not bound by this rather narrow limitation. But any extension to less clearly defined framework rights, such as personality or privacy rights, carries the risk of making the liability unpredictable and could thus jeopardise the efficient handling of claims. This would be particularly the case if liability were to be extended to pure economic losses. Such a wide definition of the protected legal interests would also deviate significantly from the existing frameworks of strict liability, so that the extensive experience gained with these systems could only be used to a limited extent. **Limitations on the maximum amount of liability** (European Commission 2018b, p. 20), on the other hand, ensure that liability risks remain calculable and facilitate insurability. Thus, their introduction seems appropriate. Their exact amount, however, is only of secondary importance. In light of the considerable interaction discussed above, it is very much to be welcomed that the Commission emphasises the need for coordinating any new liability rule with the existing liability frameworks (European Commission 2018b, p. 20 f.), especially with manufacturer's liability under the Product Liability Directive.

4.3 Liability for Autonomous Agents

With a strict liability of the owners of autonomous systems as described above under Sect. 4.2, liability gaps in particularly sensitive areas (death, personal injury, substantial damage to property) could be largely avoided. Alternatively or additionally, it is sometimes proposed to hold the owners of such systems liable according to the rules concerning the liability of principals for their agents (either directly or by analogy).[12] Whether such liability could make a significant contribution to the future liability framework cannot be said with certainty at present, as the answer also depends on whether a new strict liability rule will actually be created and, if so, how exactly it will be designed.

For the field of **contract law,** it has been discussed for some time whether **autonomous systems** should be recognised as **vicarious agents** under § 278 BGB (*Erfüllungsgehilfen*) to attribute their misconduct to the debtor regardless of the latter's own fault. The primary motivation of such considerations is often to find an appropriate solution in case the autonomous system harms interests of the creditor that the debtor is obliged to protect under § 241 (2) BGB. For example, Hanisch has created the hypothetical example of a craftsman who is carrying out repairs in a department store (commissioned by the store owner) and is then injured by a cleaning robot under the store owner's control. The malfunctioning of the robot and the injury may have been unforeseeable and thus unavoidable for the store owner (i.e. the contractual partner of the craftsman) (Hanisch 2010, p. 20). The requirements for liability under §§ 280 (1), 241 (2) BGB could then not be proven due to the store owner's lack of fault (ibid., p. 21 f.). However, cases like these are most likely to fall under the new liability rule for owners of autonomous systems

[12] See e.g. Teubner (2018), p. 185 ff.; Kluge and Müller (2017), p. 28 f.; Schirmer (2016), p. 664 f.; Wulf and Burgenmeister (2015), p. 407. For the contrary view see among others Heuer-James et al. (2018), p. 2829 f.; Horner and Kaulartz (2015), p. 505.

(as discussed above under Sect. 4.2). The craftsman would then not be without protection if he suffered bodily injury.

The question whether to apply § 278 BGB might therefore be of greater importance where autonomous systems are used to fulfil other contractual obligations, particularly the debtor's main **duty of performance**. Consider, for example, the case of a provider of winter services who uses a salting robot to secure icy pavements. If the robot does not drive out during winter weather because of a sensor failure, the provider will breach his or her duty to perform. If someone is injured and demands compensation from the owner of the property under § 823 (1) BGB, the question arises as to whether the owner would have a claim for indemnification against the winter service provider under §§ 634 No. 4, 280 (1) and (3), 283 BGB. As in the previous example, this claim would likely fail due to the lack of fault on the part of the service provider who, in ordinary circumstance, could not foresee and avoid the failure of the sensors, which might have been caused e.g. by a programming error in the control software. A possible future strict liability rule would probably not apply because the property owner's obligation to compensate the victim is a matter of pure economic loss.

Attributing the salting robot's 'fault' to the winter service provider under § 278 BGB would eliminate the problem of establishing fault on the part of the service provider. But such a shortcut would also mean a far-reaching **departure from the principle of fault** because the service provider in the example has actually observed reasonable care. Spindler correctly points out that, under German contract law, particularly §§ 276, 278, 280 BGB, the debtor is generally not liable for machine failures and technical defects, and it is not obvious why this should be different for autonomous systems (Spindler 2016, p. 816). Precisely because these systems are less controllable even for their users or owners, it seems particularly necessary to allow the debtor to exonerate himself or herself under § 280 (1) (2) BGB. The fact that this favours the debtor seems largely unproblematic, at least in contract law, because the parties are free to distribute the liability risks differently by means of guarantees, etc.

For the **field of torts**, Teubner has shown that holding owners of autonomous systems strictly liable would not solve all problems (Teubner 2018, p. 191 ff.). This is mainly due to the fact that the scope of such liability would probably be relatively narrow. Although the exact limits still need to be defined and recent remarks by the European Commission indicate a preference for a 'broad' approach (European Commission 2018b, p. 19 ff.), it is likely (and for the reasons stated also correct) that a new strict liability rule would initially follow familiar examples. This should apply, for example, to the **protected interests**, which often only include life, body, and property. However, while motor vehicles, drones, animals, dangerous installations, etc. in most cases only cause bodily injury or property damage, the harm potential of autonomous systems is greater because they are used in virtually all areas of life (Teubner 2018, p. 194). Two examples may illustrate this point. If, for example, Google's search suggestion algorithm violates **personality rights** because it suggests defamatory terms in addition to the terms typed by the user, this would not be covered by strict liability modelled after traditional rules, as these do not extend to violations of personality rights. In a famous German case, the Federal Court of Justice was able to rely on § 823 (1) BGB, because Google developed the prediction algorithm itself, and it is therefore easy to assume that Google has breached its duty of care (BGH, Judgement of 14 May 2013 – VI ZR 269/12, BGHZ 197, 213 ff.) But what if the algorithm was not developed by the user, but by a third party? In such a situation, one could hardly blame the user of the algorithm, because he or she usually does not know the details of the programming. The injured party could then only turn to the producer. Similar problems arise if the algorithm 'merely' causes **pure economic losses**, which are also not covered by traditional strict liability. For example, one could think of automated trading with securities (so-called algorithmic trading). A strict liability rule limited to physical damage according to the traditional model would be of little use.

It is doubtful, however, whether a '**liability for digital assistance**' (*digitale Assistenzhaftung*), as proposed by Teubner (Teubner 2018, p. 191 ff.; similarly already Hanisch 2014, p. 54) would be the appropriate solution. Teubner see this as a strict liability for autonomous systems according to the principle of *respondeat superior*. This approach differs from the proposal of strict liability as discussed above in that no new liability rule would need to be established. Instead, the owner of the autonomous system would become liable whenever the system 'commits a tort', e.g. under § 823 (1), § 823 (2) or § 826 BGB. Thus, the owner could, for example, become liable for the system's violations of personality rights or the causing of pure economic losses, insofar as these interests are at all protected by tort law. A major disadvantage in comparison with the strict liability rule previously discussed, however, would be that the liability requirements and thus the entire process of claims settlement would be more complicated and uncertain. For example, it must be clarified which systems should be considered as capable of committing torts (Kluge and Müller 2017, p. 26 f.). It would also be difficult to determine whether an autonomous system acts intentionally or negligently (Heuer-James et al. 2018, p. 2829 f.) or whether the system has the necessary knowledge of circumstances that qualify a conduct as contrary to public policy under § 826 BGB. For now, it therefore seems more appropriate to rely on a clear rule of strict liability as previously discussed and to close any problematic gaps in liability by means of special regulations e.g. in securities law or in the area of platform regulation.

4.4 Autonomous Systems' Own Liability

At present, it would have no practical benefit to assign liability to the autonomous system as such, as is sometimes suggested in the literature.[13] The system's current

[13] Kluge and Müller (2017), p. 29 f.; Beck (2013), p. 256; Beck (2009), p. 229 f.; Gruber (2012), p. 156. See also Bodungen and Hoffmann (2015), p. 525. For a different view see Schirmer (2016), p. 665; Spindler (2015), p. 774 f.; Hanisch (2014), p. 39 f.

inability to commit torts and possess assets is not even the problem, because such a (partial) legal capacity could be established by the legislator (see also Schirmer 2016, p. 665). But where should the assets assigned to the autonomous system come from? Ultimately, attention would have to turn again to the usual suspects, particularly users, owners and manufacturers. Legally, one could easily oblige these parties to make assets available to the autonomous system. But what would the advantage be compared to direct liability of the aforementioned groups of persons? To open a separate bank account for each autonomous system and deposit a minimum amount for liability (Hilgendorf 2012, p. 128) would mean a lot of dead capital (Hanisch 2014, p. 40). Such an approach would be extremely inefficient because it lacks any risk pooling. In contrast, a **compulsory insurance system**, as is sometimes proposed in combination with the personal liability of autonomous systems (Hilgendorf 2012, p. 128), is a good solution. However, any insurance requirement could also be linked to the liability of the owner, especially as they would remain the contact persons for all processes and activities in which the autonomous system needs human assistance anyway. A certain potential for approaches that rely on the personal liability of the autonomous system could exist to **fund solutions**. Such liability schemes are characterised by the fact that several persons contribute (voluntarily or by virtue of legal obligation) to the asset stock that is then available for liability. For example, the contributing parties could all be members of the value network in which the autonomous system is manufactured or used. However, where such funds are considered beneficial in practice, they can easily be created contractually. For now, the further development can thus be left to the markets, and it does not seem advisable for the legislator to interfere.

4.5 Systemic Liability

A major challenge for liability law is seen in the fact that autonomous systems will increasingly interact with each other in the future. Therefore, it may no longer be possible to assign liability to the owner or operator of a specific system. Spiecker gen. Döhmann speaks of '**systemic digitisation**' (*systemische Digitalisierung*) and has impressively described the resulting problems for liability. These risk result above all from the fact that in strongly interconnected, 'collectivised' systems it may become increasingly difficult to identify which component or process was responsible for the specific defect, and to what extent (Spiecker gen. Döhmann 2016). Teubner has picked up these considerations under the keyword '**interconnection risk**' (*Vernetzungsrisiko*). As an example, he cites the so-called flash crash of 6 May 2010 (Teubner 2018, p. 201 ff.), in which the Dow Jones stock market index plummeted by almost 1000 points within a few minutes, for which knowledgeable observers have blamed the interaction of algorithms in high-frequency trading in addition to market manipulation by individuals.[14] At least on the interplay of algorithms, it is difficult to assign responsibility to individual parties. Since the main problem of systemic digitisation is that it can no longer be established who contributed to what extent to the damage caused by multiple systems, the solutions discussed so far, such as owner's liability or personal liability of the autonomous systems themselves, do not help (Spiecker gen. Döhmann 2016, p. 702 f.).

Spiecker gen. Döhmann has therefore put up for discussion a proposal of a '**gradual joint and several liability**' (*graduelle Gesamtschuld*), in which all parties involved in a highly interconnected system would be jointly and severally liable, but to different extents. Specifically, liability could be differentiated according to the role in which parties are involved in the system (e.g. as entrepreneurs or as consumers; Spiecker gen. Döhmann 2016, p. 703 f.). Within groups consisting of companies, liability ratios could, for example, be based on market shares (ibid.; see also Hanisch 2014, p. 41 ff.). Pointing in a similar direction, Teubner has proposed the creation of **risk pools** 'by virtue of authoritative order'. According to his proposal, these pools would not be determined by cooperative, organisational or technical structures, but solely on the respective pool's 'ability to manage risk' (Teubner 2018, p. 203).

The discussion about systemic digitisation, which admittedly is still in its infancy, thus seems to be tending towards collective security systems, which may also include **insurance** and **fund solutions** as well as **tax-financed compensation systems**. Such approaches have the advantage that they allow for compensating the injured parties even without identifying the actual perpetrators. A major disadvantage is the lack or inadequate incentives for careful behaviour. If companies contribute to liability funds according to their turnover, market shares, etc., but not depending on the actual damage they cause, they have little incentive to avoid the harm. This results in a classic moral hazard problem. Collective systems can therefore only be the *ultima ratio*. Thus, even in areas of systemic digitisation, policymakers should attempt to preserve as long as possible the **individual attribution of responsibility**. For example, regulations could require that, wherever possible, highly interconnected systems be designed in such a way that responsibilities can still be determined (e.g. by always using clear identifiers and log files; see also Spiecker gen. Döhmann 2016, p. 703).

5 Conclusion

Overall, it seems fair to say that digitisation, automation and interconnection—in Industry 4.0 as well as in other areas—increase the complexity of liability law. They can make it more difficult to determine causality and to attribute fault. Moreover, at least in some cases, greater effort may be necessary to successfully establish a claim.

[14] See Wikipedia, '2010 Flash Crash'.

However, the discussion of these issues has not revealed any major liability gaps. The legal policy discussion on automation technologies primarily concerns the question of whether liability is still properly assigned under the changed actual circumstances (e.g. whether it is still appropriate, even in the case of a highly automated or autonomous vehicle, to make the vehicle holder liable for accidents regardless of fault), or whether liability should be redistributed to the manufacturer.

Given the relevance of emerging technologies to the internal market, legislative proposals, e.g. on the further development of product liability, are best located at the EU level. In the industrial sector, businesses can also create clarity among themselves through contractual regulations. In this way, the parties can also ensure that pragmatic solutions in line with their respective business and industry interests are used. Questions of recourse deserve special attention, as they will almost certainly gain in importance in increasingly complex value networks of the future. To avoid difficulties in collecting and assessing evidence, it is advisable not only to allocate responsibilities legally, but also to make them traceable through technical solutions (identity certificates, log files, etc.). The Liability Law 4.0 will likely be characterised by a changed legislative and regulatory framework, but especially in the area of Industry 4.0 also and above all by a growing importance of contractual solutions.

References

Albrecht F (2005) “Fährt der Fahrer oder das System?” – Anmerkungen aus rechtlicher Sicht. Straßenverkehrsrecht (SVR):373–376

Arnold S (2020) Commentary on § 445a BGB. In: Gsell B et al (eds) beck-online.Grosskommentar zum Zivilrecht, 15 December 2020. C.H. Beck, Munich

Beck S (2009) Grundlegende Fragen zum rechtlichen Umgang mit der Robotik. Juristische Rundschau (JR):225–230

Beck S (2013) Über den Sinn und Unsinn von Statusfragen. In: Hilgendorf E, Günther J-P (eds) Robotik und Gesetzgebung. Nomos, Baden-Baden, pp 239–260

Bodungen B, Hoffmann M (2015) Belgien und Schweden schlagen vor: Das Fahrsystem soll Fahrer werden! Neue Zeitschrift für Verkehrsrecht (NZV):521–526

Bodungen B, Hoffmann M (2016) Autonomes Fahren – Haftungsverschiebung entlang der Supply Chain? Neue Zeitschrift für Verkehrsrecht (NZV):449–454 (Part 1), 503–509 (Part 2)

Borges G (2016) Haftung für selbstfahrende Autos. Computer und Recht (CR):272–280

Bräutigam P, Klindt T (2015) Industrie 4.0, das Internet der Dinge und das Recht. Neue Juristische Wochenschrift (NJW):1137–1142

Burmann M (2020) Commentary on § 7 StVG. In: Burmann M et al (eds) Straßenverkehrsrecht, 26th edn. C.H. Beck, Munich

Chirco C (2015) Industrie 4.0 in der Praxis – Die Auswirkungen der Vernetzung von Wertschöpfungsketten auf die anwaltliche Beratung. In: Taeger J (ed) Internet der Dinge. OlWIR Publishing Company, Edewecht, pp 519–535

Droste J (2015) Produktbeobachtungspflichten der Automobilhersteller bei Software in Zeiten vernetzten Fahrens. Corporate Compliance Zeitschrift (CCZ):105–110

Ernst W (2019) Commentary on § 280 BGB. In: Säcker FJ et al (eds) Münchener Kommentar zum Bürgerlichen Gesetzbuch, vol 2, 8th edn. C.H. Beck, Munich

European Commission (2018a) Report on the Application of the Council Directive on the approximation of the laws, regulation, and administrative provisions of the Member States concerning liability for defective products, COM(2018) 246 final

European Commission (2018b) Liability for emerging digital technologies, SWD(2018) 137 final

European Commission (2018c) Evaluation of Council Directive 85/374/EEC, SWD(2018) 157 final

Förster C (2020) Commentary on § 1 ProdHaftG. In: Bamberger HG et al (eds) Beck'scher Online-Kommentar zum Bürgerlichen Gesetzbuch, 1 October 2020. C.H. Beck, Munich

Gless S, Janal R (2016) Hochautomatisiertes und autonomes Autofahren – Risiko und rechtliche Verantwortung. Juristische Rundschau (JR):561–575

Gomille C (2016) Herstellerhaftung für automatisierte Fahrzeuge. Juristenzeitung (JZ):76–82

Gottwald P (2018) Die Beweislast. In: Rosenberg L, Schwab KH, Gottwald P (eds) Zivilprozessrecht. C.H. Beck, Munich, § 116

Gruber M-C (2012) Rechtssubjekte und Teilrechtssubjekte des elektronischen Geschäftsverkehrs. In: Beck S (ed) Jenseits von Mensch und Maschine. Nomos, Baden-Baden, pp 133–160

Grundmann S (2019) Commentary on § 278 BGB. In: Säcker FJ et al (eds) Münchener Kommentar zum Bürgerlichen Gesetzbuch, vol 2, 8th edn. C.H. Beck, Munich

Grützmacher M (2016) Die deliktische Haftung für autonome Systeme – Industrie 4.0 als Herausforderung für das bestehende Recht? Computer und Recht (CR):695–698

Hager J (2009) Commentary on § 823 BGB, Teil F – Die deliktische Produkthaftung. In: Staudingers Kommentar zum Bürgerlichen Gesetzbuch. de Gruyter, Berlin

Hanisch J (2010) Haftung für Automation. Cuvillier, Göttingen

Hanisch J (2014) Zivilrechtliche Haftungskonzepte für Robotik. In: Hilgendorf E (ed) Robotik im Kontext von Recht und Moral. Nomos, Baden-Baden, pp 27–61

Heuer-James J-U, Chibanguza KJ, Stücker B (2018) Industrie 4.0 – vertrags- und haftungsrechtliche Fragestellungen. Betriebs-Berater (BB):2818–2832

Hilgendorf E (2012) Können Roboter schuldhaft handeln? In: Beck S (ed) Jenseits von Mensch und Maschine. Nomos, Baden-Baden, pp 119–132

Horner S, Kaulartz M (2015) Rechtliche Herausforderungen durch Industrie 4.0: Brauchen wir ein neues Haftungsrecht? – Deliktische und vertragliche Haftung am Beispiel "Smart Factory". In: Taeger J (ed) Internet der Dinge. OlWIR Publishing Company, Edewecht, pp 501–518

Iliou C (2017) Industrie 4.0: Contract Splitting und werkvertraglicher Erfolg – ist beides möglich? Betriebs-Berater (BB):847–849

Kluge V, Müller A-K (2017) Autonome Systeme. Zeitschrift für Innovations- und Technikrecht (InTeR):24–31

König C (2017) (Vorerst) Keine Haftungsbefreiung für Halter und Fahrer automatisierter Kraftfahrzeuge. Juristische Rundschau (JR):323–332

Kütük-Markendorf ME, Essers D (2016) Zivilrechtliche Haftung des Herstellers beim autonomen Fahren. MultiMedia und Recht (MMR):22–26

Lutz LS (2015) Autonome Fahrzeuge als rechtliche Herausforderung. Neue Juristische Wochenschrift (NJW):119–124

Marschollek JP, Wirwas M (2017) Digitalisierung der Elektroindustrie (Smart Factory). In: Sassenberg T, Faber T (eds) Rechtshandbuch Industrie 4.0 und Internet of Things. C.H. Beck, Munich, pp 420–436

Oechsler J (2018) Commentary on the ProdHaftG. In: Staudingers Kommentar zum Bürgerlichen Gesetzbuch. de Gruyter, Berlin

Orlikowski-Wolf SM (2018) Auswirkungen der Änderungen der kaufrechtlichen Mängelhaftung zum 1.1.2018 auf AGB im B2B-Verkehr. Zeitschrift für Wirtschaftsrecht (ZIP):360–364

Orthwein M, Obst J-S (2009) Embedded Systems – Updatepflichten für Hersteller hardwarenaher Software. Computer und Recht (CR):1–4

Pieper F-U (2016) Die Vernetzung autonomer Systeme im Kontext von Vertrag und Haftung. Zeitschrift für Innovations- und Technikrecht (InTeR):188–194

Redeker H (2017) IT-Recht, 6th edn. C.H. Beck, Munich

Reichwald J, Pfisterer D (2016) Autonomie und Intelligenz im Internet der Dinge. Computer und Recht (CR):208–212
Riehm T (2014) Von Drohnen, Google-Cars und Software-Agenten. IT-Rechtsberater (ITRB):113–115
Schäfer C (2017) Drittschäden durch Drohnen. Versicherungsrecht (VersR):849–855
Schaub R (2020) Commentary on § 278 BGB. In: Gsell B et al (eds) beck-online.Grosskommentar zum Zivilrecht, 15 December 2020. C.H. Beck, Munich
Schirmer J-E (2016) Rechtsfähige Roboter? Juristenzeitung (JZ):660–666
Schrader PT (2015) Haftungsrechtlicher Begriff des Fahrzeugführers bei zunehmender Automatisierung. Neue Juristische Wochenschrift (NJW):3537–3542
Schrader PT, Engstler J (2018) Anspruch auf Bereitstellung von Software-Updates? MultiMedia und Recht (MMR):356–361
Schroeter U (2010) Untersuchungspflicht und Vertretenmüssen des Händlers bei der Lieferung sachmangelhafter Ware. Juristenzeitung (JZ):495–499
Sosnitza O (2016) Das Internet der Dinge – Herausforderung oder gewohntes Terrain für das Zivilrecht? Computer und Recht (CR):764–772
Spickhoff A (2017) Commentary on § 4 ProdHaftG. In: Gsell B et al (eds) beck-online. Grosskommentar zum Zivilrecht, 15 December 2020. C.H. Beck, Munich
Spiecker gen. Döhmann I (2016) Zur Zukunft systemischer Digitalisierung – Erste Gedanken zur Haftungs- und Verantwortungszuschreibung bei informationstechnischen Systemen. Computer und Recht (CR):698–704
Spindler G (2008) IT-Sicherheit – Rechtliche Defizite und rechtspolitische Altenrativen. Computer und Recht (CR):7–13
Spindler G (2015) Roboter, Automation, künstliche Intelligenz, selbst-steuernde Kfz – Braucht das Recht neue Haftungskategorien? Computer und Recht (CR):766–776
Spindler G (2016) Digitale Wirtschaft – analoges Recht: Braucht das BGB ein Update? Juristenzeitung (JZ):805–816
Teubner G (2018) Digitale Rechtssubjekte? Archiv für civilistische Praxis (AcP) 218:155–205
Wagner G (2017a) Preliminary remarks on §§ 823 ff. BGB. In: Säcker FJ et al (eds) Münchener Kommentar zum Bürgerlichen Gesetzbuch, vol 6, 7th edn. C.H. Beck, Munich
Wagner G (2017b) Commentary on § 823 BGB. In: Säcker FJ et al (eds) Münchener Kommentar zum Bürgerlichen Gesetzbuch, vol 6, 7th edn. C.H. Beck, Munich
Wagner G (2017c) Commentary on the ProdHaftG. In: Säcker FJ et al (eds) Münchener Kommentar zum Bürgerlichen Gesetzbuch, vol 6, 7th edn. C.H. Beck, Munich
Wagner G (2017d) Produkthaftung für autonome Systeme. Archiv für civilistische Praxis (AcP) 217:707–765
Weller M (2012) Die Verantwortlichkeit des Händlers für Herstellerfehler. Neue Juristische Wochenschrift (NJW):2312–2317
Wende S (2017) Haftungsfragen bei vernetzten und autonomen Systemen. In: Sassenberg T, Faber T (eds) Rechtshandbuch Industrie 4.0 und Internet of Things. C.H. Beck, München, pp 69–84
Westphalen F (2015) Der Referentenentwurf zur Änderung der kaufrechtlichen Mängelhaftung – Licht und Schatten. Betriebs-Berater (BB):2883–2892
Wulf HM, Burgenmeister C (2015) Industrie 4.0 in der Logistik – Rechtliche Hürden beim Einsatz neuer Vernetzungs-Technologien. Computer und Recht (CR):404–412

Copyright Law 4.0

Gerald Spindler

Contents

G. Spindler (✉)
University Göttingen, Göttingen, Germany
e-mail: lehrstuhl.spindler@jura.uni-goettingen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_13

1 Introduction and Terminology

Industry 4.0 has become a catchword in the German discussion, describing a variety of phenomena associated with increasing digitalization in the industry. These include the direct communication of machines with each other across the boundaries of companies or legal entities, the mass production and use of non-personal data, the close (digitally supported) networking of companies, and even the use of "artificial intelligence" (AI).

Of course, general copyright law also claims validity in this connected world and does not need to be described separately, e.g., as to the form of license agreements. What is of interest, however, are the special aspects raised by the characteristic phenomena of Industry 4.0 from a copyright perspective. From the perspective of "Copyright 4.0," the protection of data and databases appears to be particularly controversial (see Sect. 2); but also, in particular, the legal admissibility of text and data mining (TDM) under copyright law, which must be considered from the perspective of both current and future law (see Sect. 3). The increasing cooperation of various partners beyond the boundaries of the company also requires a greater participation in joint intellectual property rights (see Sect. 4). With the use of a wide variety of cloud applications, but especially cloud computing, copyright issues are ultimately also becoming a growing concern for a company's infrastructure (see Sect. 5).

2 (No) Copyright Protection of Data

Whether and, if so, how data and data collections are legally protected has been and continues to be the subject of intensive discussions both *de lege lata* and *de lege ferenda*. In view of their proximity to economic goods, which are classically protected by intellectual property law, it is obvious that this question must also be addressed for data. In line with the focus on intellectual property law, the following section will not discuss in detail problems relating to the protection of data in tort (for example, under § 823 (1) of the German Civil Code or their treatment under property law or contract law, nor will it address issues relating to the data protection of personal data, which are the subject of the GDPR (on the definition of machine-generated, non-personal data).

2.1 Data and Copyright Protected Works

Under § 2 (2) of the Copyright Act, the prerequisite for copyright protection is an intellectual, individual creation of a work, without § 2 (1) of the Copyright Act containing an exhaustive list. Admittedly, case law has considerably reduced the requirements for an intellectual level of creation over the course of time by also granting recognition to the so-called "*small coin*" (see e.g., BGH, decision of 27.1.1983 – I ZR 177/80, GRUR 1983, p. 378 – Brombeer-Muster; on the alignment of utility art: BGH, decision of 13.11.2013 – I ZR 143/12, GRUR 2014, p. 175 paras 26 ff. – Geburtstagszug; Wandtke and Bullinger/*Bullinger* 2019, § 2 UrhG paras 23 f.; Dreier and Schulze/*Schulze* 2018, § 2 UrhG, paras 24 ff.); but it is obvious that this does not apply to pure data, since these are not created by a creative work. Furthermore, data that is only created using algorithms and thus through novel combinations of already existing data is only the product of this software or algorithms, but not an independent intellectual **creation** (EU Commission (Hartmann et al.) 2020, p. 28; Ramalho 2017, p. 13 ff.; Grimmelmann 2016, pp. 408 ff. with further references). And also the protectability of the collection under § 4 (1) of the Copyright Act or the database work under § 4 (2), (1) of the Copyright Act extends in any case only to its structure, but not to the elements (works and data) it contains, (CJEU, decision of 1.3.2012 - C-604/10, ECLI:EU:C:2012:115, para 30—Football Dataco; Dreier and Schulze/*Dreier* 2018, § 4 UrhG paras 11 f.; Möhring and Nicolini/*Ahlberg* 2018, § 4 UrhG para 11), meaning no protection of the data can be achieved by this either. Consequently, protection of data under these provisions of copyright law is not possible.

2.2 Database Protection (§§ 87a ff. Copyright Act)

However, the legal situation is more differentiated on the protection of databases under §§ 87a ff. of the Copyright Act. For the protection of databases under §§ 87a ff. of the Copyright Act, no intellectual creation is required; rather, a substantial investment in the database is sufficient, although, in the opinion of the ECJ, not all costs can be considered (see e.g., CJEU, decision of 9.11.2004 – C 203/02, ECLI:EU:C:2004:695, paras 31, 42 – The British Horseracing Board; CJEU, decision of 9.11.2004 – C-338/02, ECLI:EU:C:2004:696, para 24 – Fixtures-Fußballspielpläne I; CJEU, decision of 9.11.2004 – C-444/02, ECLI:EU:C:2004:697, para 40 – Fixtures-Fußballspielpläne II; CJEU, decision of 9.11.2004 – C-46/02, ECLI:EU:C:2004:694, paras 31 ff. – Fixtures-Marketing). The individual datum itself is not protected by §§ 87a ff. of the Copyright Act; the legislator resp. the legislator of the Directive has deliberately decided in favor of the public domain of the data itself (see Recital 46 of the Directive 96/9/EC of the European Parliament and the Council of. 11 March 1996 on the legal protection of databases, OJ L 77 p. 20; Spindler/Schuster/*Wiebe* 2019, § 87a UrhG para 1; Wandtke and Bullinger/*Hermes* 2019, §

87a UrhG para 5; Schmidt and Zech 2017, p. 418; Wiebe 2017, p. 338; Ehmann 2014, p. 395). Indirect protection may arise, nevertheless.

2.2.1 Requirements for Protection

The prerequisite for protection under §§ 87a ff. of the Copyright Act is the independence of the elements contained in the database from each other, particularly their meaning, without this meaning only arising in connection with other elements (Schricker and Loewenheim/*Vogel* 2020, § 87a UrhG paras 12, 14; Wandtke and Bullinger/*Hermes* 2019, § 87a UrhG paras 12 f. ; Fromm and Nordemann/ *Czychowski* 2018, § 87a UrhG para 9; Dreier and Schulze/*Dreier* 2018, § 87a UrhG para 6; Wiebe 2017, p. 339). The decisive factor is whether any third party can still extract an—albeit possibly reduced—information value from the element detached from the collection (see CJEU, decision of 9.11.2004 – C-444/02, ECLI: EU:C:2004:697, para 33—Fixtures-Fußballspielpläne II; CJEU, decision of 1.3.2012 – C-604/10, ECLI:EU:C:2012:115, Rn. 26 – Football Dacato i.a.), as for example geographic data within a topographic map (see CJEU, decision of 29.10.2015 – C-490/14, ECLI:EU:C:2015:735, Rn. 29 – Verlag Esterbauer). For most applications in the field of the Industry 4.0, the **independence of the elements** will be given without any further concern, since they are collected separately and generate new insights only through their combination (correctly so for big-data appliances Schmidt and Zech 2017, p. 419). However, the independence of elements in the context of AI systems is excluded, since its (machine-generated) parameters do not convey any information value in their isolated form (Hartmann and Prinz 2018, p. 786). The situation may be different for the AI generated data, i.e., the AI products (Hetmank and Lauber-Rönsberg 2018, p. 578).

Furthermore, **the systematic or methodical arrangement** is required for database protection—without any particular differences here; the only essential factor is that the individual date must be retrievable with the assistance of a link (CJEU, decision of. 9.11.2004 – C-444/02, ECLI:EU:C:2004:697, para 30 – Fixtures-Fußballspielpläne II; see also Wandtke and Bullinger/*Hermes* 2019, § 87a UrhG para 24: de minimis-criteria). Only completely unorganized "data piles" are thus excluded (OLG Köln, decision of 15.12.2006 - 6 U 229/0, MMR 2007, p. 444 – DWD-Wetterdaten; OLG München, decision of 9.11.2000 – 6 U 2812/00, GRUR-RR 2001, p. 228 – Übernahme fremder Inserate; KG, decision of 26.5.2000 – 5 U 1171/00, GRUR-RR 2001, p. 102 – Stellenmarkt*;* Dreier and Schulze/*Dreier* 2018, § 87a UrhG para 7; Schricker and Loewenheim/*Vogel* 2020, § 87a UrhG para 24; Wiebe 2017, p. 340). As *Schmidt/Zech* rightly point out, however, even this criterion is questionable in the age of Big Data applications, since raw data can always be found through the query software—which is not part of database protection (Schmidt and Zech 2017, p. 420). Nevertheless, at least the individual accessibility of the respective elements does not represent a major hurdle; neural networks or the files created through training an artificial intelligence are however excluded from the database protection (correctly Schmidt and Zech 2017, p. 420); but not the training data itself.

Decisive factor for the question of protection is—if the other requirements of database protection are met—the focus of §§ 87a ff. of the Copyright Act as investment protection, particularly the extent to which the very generation of data is capable of protecting it. A prerequisite for the protection of databases is the investment of substantial resources in the acquisition, verification and presentation of the database and its contents (see CJEU, decision of 9.11.2004 C-338/02, ECLI: EU:C:2004:696, paras 22 f.—Fixtures-Fußballspielpläne I; CJEU, decision of 9.11.2004 – C-444/02, ECLI:EU:C:2004:697, paras 38 f. – Fixtures-Fußballspielpläne II; Dreier and Schulze/*Dreier* 2018, § 87a UrhG para 12; Wandtke and Bullinger/*Hermes* 2019, § 87a UrhG paras 34 f.; Wiebe 2017, p. 341), but not investments in the creation of the database (data generation) itself, according to the case law of the European Court of Justice (CJEU, decision of 9.11.2004 – C-203/02, ECLI:EU:C:2004:695 – The British Horseracing Board; CJEU, decision of 9.11.2004 – C-46/02, ECLI:EU:C:2004:694; CJEU, decision of 9.11.2004 - C-338/02, ECLI:EU:C:2004:696 – Fixtures-Fußballspielpläne I; CJEU, decision of 9.11.2004 – C-444/02, ECLI:EU:C:2004:697 – Fixtures-Fußballspielpläne II; Wandtke and Bullinger/*Hermes* 2019, § 87a UrhG para 36). The reason for this restriction—not necessarily plausible— is to keep self-generated data free from monopolization (so-called sole-source databases) (Schricker and Loewenheim/*Vogel* 2020, § 87a UrhG paras 43 ff; Wandtke and Bullinger/*Hermes* 2019, § 87a UrhG para 44; Leistner 2007, p. 458; Ehmann 2014, pp. 397 f.; Wiebe 2014, p. 4; but see Schmidt and Zech 2017, pp. 421 f.; concerning Industry 4.0-Appliances Sattler 2017, § 2 paras 20 ff.). Accordingly, the investment necessary for the protection of the database does not depend on whether the database originator first obtained the data through elaborate measurement or observation procedures, but only on the investment in the generation of the database itself (Dreier and Schulze/*Dreier* 2018, § 87a UrhG Rn. 13; Wandtke and Bullinger/*Hermes* 2019, § 87a UrhG para 49; Wiebe 2016, p. 879; Schmidt and Zech 2017, pp. 421 f.). Others, in contrast, focus on whether third parties could also collect the data with comparable effort which, on the one hand, would also capture the data of many observations and measurements (Leistner 2007, p. 460; Grützmacher 2016, p. 488; Schmidt and Zech 2017, pp. 421 f.; Wiebe 2017, p. 341; Hetmank and Lauber-Rönsberg 2018, p. 578), but on the other hand—and apart from a very abstract delimitation (see Wiebe 2017, p. 341)—would further exclude exclusively temporarily observable data (such as weather phenomena), since a third party cannot repeat this observation (consequently Schmidt and Zech 2017, p. 422). This seems consistent, at least considering the otherwise threatening issue of monopolizing “volatile” data.

The investment must also be substantial in nature or scope, for instance, in having expended a significant amount of time, energy, and labor, or in having used novel combination and categorization procedures (Wandtke and Bullinger/*Hermes* 2019, § 87a UrhG paras 59 ff.; Fromm and Nordemann/*Czychowski* 2018, § 87a UrhG para 15). In contrast, obtaining information that is publicly and easily accessible is considered immaterial (LG Düsseldorf, decision of 7.2.2001 – 12 O 492/00, ZUM

2002, 66; Dreier and Schulze/*Dreier* 2018, § 87a UrhG para 15). In general, however, the case law does not make high demands (BGH, decision of 1.12.2010 – I ZR 196/08, GRUR 2011, p. 724 with annotations *Sendrowski* paras 20 ff. – Zweite Zahnarztmeinung II; Fromm and Nordemann/*Czychowski* 2018, § 87a UrhG Rn. 16; Wandtke and Bullinger/*Hermes* 2014, § 87a UrhG para 54; see Dreier and Schulze/*Dreier* 2018, § 87a UrhG para 15).

For Industry 4.0 designs, this means that investments in the creation of data, e.g., by robots as part of a production line, do not become part of the investments required for the database resp. cannot justify the protection—but they can if these data are licensed to a third party (even to a subsidiary company (BGH, decision of 6.5.1999 – I ZR 199/96, GRUR 1999, p. 925—Tele-info.CD; Wandtke and Bullinger/*Hermes* 2019, § 87a UrhG Rn. 45 ff.; similarly on manufacturer data Wiebe 2017, p. 342)) so that the database protection then arises with this third party, because the third party has then expended effort for the acquisition of the data. The buyer of a robot with data generation and systematization therefore invests, according to the case law, "only" in the data generation, whereas he could be the database generator if he acquired the same data from the robot supplier (correctly Sattler 2017, § 2 para 37; Wiebe 2017, p 342). Especially furthermore problematic are the cases typical for "Industry 4.0" designs, in which the data only arise as "by-products" of the activity, but the creation of a database is precisely not the purpose of the investment, so called "spin-off" (Wandtke and Bullinger/*Hermes* 2019, § 87a UrhG para 41). Nevertheless, the judicature considers it as sufficient if the data are at least also covered by the purpose of the investment (BGH, decision of 25.3.2010 – I ZR 47/08, GRUR 2010, p. 1006 para 37—Autobahnmaut; Dreier and Schulze/*Dreier* 2018, § 87a UrhG para 13; Wandtke and Bullinger/*Hermes* 2019, § 87a UrhG para 41; Schmidt and Zech 2017, p. 423; Ehmann 2014, pp. 396 ff.; Grützmacher 2016, p. 488).

2.2.2 Subject of Protection

The object of protection of the sui generis property right, however, also does not refer to the individual data (Art. 8 Directive 96/9/EC, Recitals 46, 49 Directive 96/9/EC), neither on the structure of the database (Recital 15 Directive 96/9/EC, explicitly Recital 58; but see, Wandtke and Bullinger/*Hermes* 2019, § 87a UrhG para 3, § 87b UrhG para 9, § 87d UrhG para 3) or the investment made as such, but "only" to the database as a result of the investment (Wandtke and Bullinger/*Hermes* 2019, Vor §§ 87a ff. UrhG para 27, sowie § 87a UrhG Rn. 2 f.; Schricker and Loewenheim/*Vogel* 2020, Vor §§ 87a ff. UrhG para 29, § 87a UrhG para 31). Accordingly, only the extraction or use of at least a substantial part of the database can be prohibited under § 87b (1) sentence 1 of the Copyright Act; therefore, not individual data, but only its (substantial) entirety, which also constitutes the database content, is subject to protection (Schmidt and Zech 2017, p. 423). For the question of when a use or withdrawal of essential parts exists, the objective of the directive (the protection of the investment and thus primarily the amortization of the investment) must be considered as well (Wandtke and Bullinger/*Hermes* 2019, § 87b UrhG paras 26 f.;

Schricker and Loewenheim/*Vogel* 2020, § 87b UrhG paras 9 f.; Dreier and Schulze/*Dreier* 2018, § 87b UrhG para 6; Schmidt and Zech 2017, p. 425). In this context, the significance of the withdrawal can be determined both qualitatively—with regard to the investments—and quantitatively—with regard to the volume of data (CJEU, decision of 9.11.2004 – C-203/02, ECLI:EU:C:2004:695, paras 70, 71 – The British Horseracing Board). Following the latter, a withdrawal of 10% should not yet be considered significant, but more than 50% should be (BGH, decision of 1.12.2010 – I ZR 196/08, GRUR 2011, p. 724 with further annotations *Sendrowski* paras 10, 29 – Zweite Zahnarztmeinung II; BGH, decision of 21.07.2005 I ZR 290/02, BGHZ 164, p. 37; Dreier and Schulze/*Dreier* 2018, § 87b UrhG para 7; Wandtke and Bullinger/*Hermes* 2019, § 87b UrhG para 15; Grützmacher 2016, p. 488).

However, the fact that it is permissible to extract (and exploit or use) insubstantial parts of the database does not mean that this exception could be used to undermine the limits of substantial extraction, for example, with repeated extractions below the substantiality limit (see CJEU, decision of 9.11.2004 – C-203/02, ECLI:EU:C:2004: 695, para 26—The British Horseracing Board; Dreier and Schulze/*Dreier* 2018, § 87b UrhG para 11). § 87b (1) sentence 2 Copyright Act prevents such circumvention strategies if insubstantial parts of the database are repeatedly and systematically reproduced, disseminated, or publicly displayed. However, since this is a circumvention protection, this exception is interpreted restrictively (Schricker and Loewenheim/*Vogel* 2020, § 87b UrhG para 60; Wandtke and Bullinger/*Hermes* 2019, § 87b UrhG para 66; Dreier and Schulze/*Dreier* 2018, § 87b UrhG para 13; Schmidt and Zech 2017, p. 425). Accordingly, it must be targeted, systematic uses that, in sum, amount to the removal of substantial parts (Wandtke and Bullinger/*Hermes* 2019, § 87b UrhG para 69; Schmidt and Zech 2017, p. 426). In the opinion of the CJEU, an extraction or reuse within the meaning of the Directive also exists not only in the case of a physical reproduction of data, but already when the "origin" of the stored data can be traced back to the protected database (Hetmank and Lauber-Rönsberg 2018, p. 579; on the loose interpretation of the CJEU see CJEU, decision 19.12.2013 – C 202/12, ECLI:EU:C:2013:850, Rn. 37 – Innoweb; CJEU, decision of 9.11.2004 – C-203/02, ECLI:EU:C:2004:695, para 67 – The British Horseracing Board).

The covered acts of **exploitation** refer to the acts known under German law, such as reproduction under § 16 Copyright Act, even if the Directive uses a different terminology on this issue (extraction) (Dreier and Schulze/*Dreier* 2018, § 87b UrhG para. 4). As copies always have to be made in digital processing, there is practically always a reproduction present (on Big Data see, e.g., Schmidt and Zech 2017, p. 426). The same applies to the making available to the public and distribution—the usual criteria are applied here, specific questions for the Industry 4.0 do not arise, except on the question of when a new public is present.

2.2.3 Protection of Data by Contractual Agreements

From a copyright point of view, data can therefore if anything be protected via §§ 87a ff. Copyright Act—but not individually, but only in a significant entirety. It is therefore not surprising that the current focus is rather on the contractually secured protection of data, in analogy to licenses under intellectual property law. For this purpose, categorizations of the accumulating data are relevant, along with the company's own analysis and further processing options. The contractual arrangements can be based on comparable contracts in the area of know-how, since this also involves information that is important for the company, but not information that is protected under intellectual property law, whereby an important component will concern the confidentiality agreements or non-disclosure agreements.

3 Text and Data Mining As (One) Requirement for Industry 4.0

3.1 Relevance of Text and Data Mining for Industry 4.0

Text and data mining (TDM, for short) is a particularly prominent technology in the broad context of Industry 4.0, networks and "Big Data." Machine learning methods are often used for the development of **AI**-supported technologies. Machine learning is a general term for various training methods in which knowledge generation takes place based on experience (see Meys 2020, p. 457; detailed: EU Commission (C. Hartmann et al.) 2020, pp. 25 ff.). The foundation for these experience-based training methods are large amounts of data—at this point, text and data mining becomes relevant. It enables the algorithm to improve its performance with each additional data set, which means that the further development of many AI technologies stands and falls with the availability and usability of data. The technical functionality of text and data mining has already been described several times (Spindler 2016, pp. 1112 f.; Spindler 2018, pp. 290 f.; generally Clark 2012; Hippner and Rentzmann 2006;[1] in depth on the technical procedures of TDM see Feldman and Sanger 2006, pp. 15 ff.; Rosati 2019, pp. 7 ff), so that it is only to be briefly mentioned here that the respective data sets are made machine-readable, reworked into structured data sets including meta information (see also Heyer et al. 2006, pp. 4 ff.; referring to this de la Durantaye 2014,[2] pp. 7 ff.) and finally transformed into a so-called corpus, which forms the actual basis for the application of the mining software. The original material is not replaced or changed in the process (see LIBER,

[1] Available at: https://gi.de/informatiklexikon/text-mining.

[2] Available at: https://edoc.hu-berlin.de/bitstream/handle/18452/19416/wissenschaftsschranke_durantaye.pdf?sequence=1.

Text and Datamining Factsheet[3]). Ultimately, it is always a form of data analysis, although the term mining is used rather as a catchword and must be understood in a broad sense (correctly Triaille et al. 2014, pp. 8 f., 17 f.). Characteristic is the automated exploitation of data in the broadest sense (texts, images, data, etc.) (Triaille et al. 2014, p. 17). Algorithms can therefore be used to determine new correlations in existing data records or texts (as emphasized at first by Hearst 1999, pp. 3 ff.). As the name suggests, this "mining" is not limited to any type of work but can extend to anything that is and can be digitized (cf. Recital 8 of the Digital Single Market Directive).

The permissibility of text and data mining under copyright law has given rise to a few questions in the past, in which the interests of the rights holders or users have naturally been and still are opposed to those of research, but also of industry in the form of commercial research. However, the development of **artificial intelligence (AI)** in particular also depends as stated before on the largest possible data volumes, as the AI needs to be trained.

As shown, however, the extraction of data from a database (even successively) can exceed the threshold of substantiality, so it stands to reason that AI use or text and data mining could infringe database rights (see, e.g., Raue 2017a, pp. 13 f.; on the sui generis-protection, the restrictive interpretation of the CJEU must be considered, CJEU decision of. 9.11.2004 – C-203/02, ECLI:EU:C:2004:695, para 87 – The British Horseracing Board; see Spindler 2016, p. 1114; Triaille et al. 2014, pp. 78 f.). In addition, the infringement of the exploitation rights to the analyzed works also comes into consideration. It is therefore even more important to have relevant limitations that only enable TDM and the training of AI as one of the keystones of Industry 4.0.

After some countries such as Great Britain, France, or Japan have already introduced certain limitations in their copyright laws to put text and data mining on a legal ground (in detail Rosati 2019, pp. 15 ff.; Spindler 2016, 1117 f.), Germany has also introduced regulations for text and data mining for the first time in the context of the **reform of the highly disputed "Act to Align Copyright Law with the Current Demands of the Knowledge-based Society"** (UrhWissG, Law of 1.9.2017 (BGBl. I p. 3346)), without conclusively ending the disputes between rights holders and users (cf. criticism on § 60d Copyright Act: Schack 2017, p. 807; Raue 2017b, pp. 657 ff.). At the EU level, the fight for limitations on text and data mining has ended for the time being; on 17 April 2019, the DSM Directive was adopted, introducing mandatory limitations for text and data mining in the scientific and **commercial** sectors. Whether this will actually strengthen the Digital Single Market in the way announced by the EU in its Digital Single Market Strategy[4] is however not entirely certain (see Sects. 3.4 and 3.5 below for more details).

[3] Available at: https://libereurope.eu/document/tdm-factsheet-v-2/.

[4] European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A Digital Single Market Strategy for Europe, COM(2015) 192 final, 4.1.

3.2 Relevant Exploitation Rights

As already described in detail earlier (Spindler 2016, p. 1113 f.), the TDM affects both the right of reproduction under § 16 of the Copyright Act and Article 2 of the InfoSoc Directive (Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, OJ L 167, 22.6.2001, p. 10) and the right of making available to the public under § 19a of the Copyright Act resp Article 3 of the InfoSoc-D. This is because usually, PDF texts or other data, for example, are converted into analyzable XML files (Triaille et al. 2014, p. 30; on the conversion of the collected texts into XML-format in general see Weiss et al. 2015, p. 15 f.), thus reproductions are made. However, there are also TDM applications that only search ("crawl") data records and do not initiate any reproductions. Excluded are only marginal reproductions, e.g., text excerpts with less than 8-11 words (CJEU, decision of 16.7.2009 – C-5/08, ECLI:EU:C:2009:465, para 51 – Infopaq; on this Schulze 2009, p. 1019; Handig 2012, p. 976; Metzger 2012, p. 120 f.).

The right to making available to the public (§ 19a Copyright Act, Art. 3 InfoSoc-D), on the other hand, does not concern the original work, which is not published in parts either, but the corpus, which still contains recognizable parts of the work (together with meta-information). The publication of this corpus is often necessary if research results are to be widely discussed and verified. Despite the transformation into a machine-readable format, however, there is no editing in the sense of § 23 Copyright Act, since it is not a change of the original work (more extensively Spindler 2016, p. 1113 f.). Thus, parts of the work are merely transferred into another format; the meta-information does not touch the actual work either, but only enriches it, so to speak, and systematizes it according to its own scheme; the production of the corpus does not constitute an editing, either. However, this does not eliminate the need for granting rights or the application of limitations on reproduction or making available to the public (see above).

Since not only individual data are searched and analyzed, databases could also be affected in addition to individual works, whether as collective works protected under copyright law, § 4 (2) Sentence 1 Copyright Act,(Art. 5 Database Directive respectively (Directive 96/9/EC of the European Parliament and the Council of March 1996 on the legal protection of databases, OJ L 77, 27.3.1996, pp. 20 ff.) or as databases protected under sui generis law under §§ 87a ff. Copyright Act (or Article 7 ff. Database-D). However, text and data mining **does not affect the logical arrangement or selection** protected under § 4 (2) sentence 1 Copyright Act (Ahlberg/Götting UrhG/*Ahlberg* 2020, § 4 UrhG para 26; Dreier and Schulze/*Dreier* 2018, § 4 UrhG para 19; Wandtke and Bullinger/*Marquardt* v, § 4 UrhG para 8; Fromm and Nordemann/*Czychowski* 2018, § 4 UrhG para. 31), nor does it involve the extraction or reuse of substantial parts of the database in qualitative or quantitative terms (Art. 7, 8 (1) Database-D) (more in detail Spindler 2016, p. 1114; Raue 2017b, p. 659 with further references) – under the condition that the license permits such uses as "normal" (CJEU, decision of 9.11.2004 – C 203/02, ECLI:EU:C:2004:

695, para 84 – The British Horseracing Board citing Recital 42 of Directive 96/9/EC). However, as already shown (see Sect. 2.2.2 above), Article 7 (5) of the Database-D again restricts the extraction of non-substantial parts: A repeated and systematic extraction of even non-substantial parts or their re-use that harms the legitimate interests of the right holder is therefore not permitted (CJEU, decision of 9.11.2004 – C 203/02, ECLI:EU:C:2004:695, para 86 – The British Horseracing Board). In this case, the CJEU requires a "restoration" of the actual database, so that TDM, which as such it is aimed at gaining knowledge and structure from data and not at the (re-)production of the original database (structure), will seldom fall under this re-exception (see CJEU, decision of 9.11.2004 – C-203/02, ECLI:EU:C:2004:695, para 87 – The British Horseracing Board; correctly so Triaille et al. 2014, p. 79). Nevertheless, an action relevant under exploitation law remains in effect—at least with regard to the specific works.

3.3 European Legal Context

The current TDM limitation in force is based on Art. 5 (3) (a) InfoSoc-D (for a comprehensive analysis of the European regulation, see Dusollier 2013, p. 359 ff.) which, as a general scientific restriction, covers both the right of reproduction and the right of making available to the public. However, Art. 5 (3) (a) of the InfoSoc-D explicitly restricts the privilege to non-commercial purposes, as is similarly provided for in Art. 6 (1) (c) of Directive 96/6/EC (Database-D) for scientific research for non-commercial purposes. Nevertheless, other (European) principles cannot be applied. Article 5 (1) (a) of the InfoSoc-D which privileges so-called ephemeral copies as a mandatory limitation for all Member States (in Germany § 44a Copyright Act), is of no help, because the ephemeral (i.e., only temporary) copies may only serve the purpose of technical transmission and have no economic value of their own (Dreier and Schulze/*Dreier* 2018, § 44a UrhG paras 4 ff.; Wandtke and Bullinger/*v. Welser* 2019, § 44a UrhG paras 2, 7, 21; Spindler/Schuster/*Wiebe* 2019, § 44a UrhG paras 3, 6; Fromm and Nordemann/*Dustmann* 2018, § 44a UrhG paras 8 ff.). However, most copies made in the context of TDM can hardly be regarded as only temporary copies and a fortiori not only serving transmission purposes; this particularly applies to the corpus that serves as the basis for mining. Doubts would also be appropriate with regard to economic exploitability, since the corpus could be used for further mining purposes beyond the TDM itself (doubts also in Triaille et al. 2014, p. 47; on the idea of obtaining a profit from the lawful use of the work see CJEU, decision of 17.1.2012 – C-302/10, ECLI:EU:C:2012:16, paras 50 ff. – Infopaq II), the same applies to data converted into machine-readable formats. Therefore, the limitation of Art. 5 (1) (a) of the InfoSoc-D or Art. 44a of the Copyright Act can hardly be applied to TDM (Triaille et al. 2014, p. 50). The situation is comparable to the exception for quotations under Article 5 (3) (d) of the InfoSoc-D and § 51 of the Copyright Act. The citation must always serve to guide the author's own thoughts or to illustrate the author's own work (Ahlberg/

Götting UrhG/*Schulz* 2020, § 51 UrhG para 13; Schricker and Loewenheim/*Spindler* 2020, § 51 UrhG paras 14 ff. (“subjective circumstance”), 39; Dreier and Schulze/*Dreier* 2018, § 51 UrhG para 3), which does not apply to TDM, since it does not serve to support the own work.

§ 60d Copyright Act Draft is based on Art. 3 DSM-D. Its legal predecessor was oriented to the Draft Directive (Proposal for a Directive of the European Parliament and of the Council on copyright in the Digital Single Market COM(2016) 593 final) available in 2016. To quickly end the legal uncertainty on text and data mining (Begr RegE BT-Drs. 18/12329, p. 24) and to realize its own digitization strategy (see Legislative Report: Digital Agenda 2014–2017, pp. 42, 90), the German legislator introduced the limitation even before the DSM-D was finally adopted. The general education and science limitation in Art. 5 (3) (a) InfoSoc-D as well as Art. 10 (1) Rental and Lending Directive (Directive 2006/115/EC of the European Parliament and of the Council of 12 December 2006 on rental right and lending right and on certain rights related to copyright in the field of intellectual property) and Art. 6 (2) and Art. 9 (b) Database-D served as a basis. Data-driven commercial research in private companies was not covered by the scope of application of the limitation and therefore remained unregulated. On commercial research, however, the legislator did not have much discretion at the time, as the InfoSoc and the Database Directive were fully harmonizing and only allowed exceptions in favor of scientific research, cf. Art. 5 (3) (a) InfoSoc Directive and Art. 6 (2) (b) Database-D. The initial Commission draft (see Proposal COM(2016) 593 final) was also still limited to exempting non-commercial research with data. It did not include a general TDM limitation, although data-driven research in companies was naturally subject to the same hurdles as non-commercial research. It was foreseeable that a lack of a limitations to commercial text and data mining would lead to significant competitive disadvantages for the Data economy in the European Union and thus contradict the EU’s Digital Single Market Strategy (DSM Strategy COM(2015) 192 final, para 2.4). This disparity in the Commission’s draft was recognized in the legislative process (see Recital 18 DSM-D) and eliminated at the very last moment. Therefore, the new Art. 4 DSM-D was included in the adopted Directive, which provides for a mandatory limitation of **any** form of text and data mining (see Sect. 3.5). From the European side, it was correctly assumed that Big Data and the analysis of copious amounts of data are of considerable importance not only for scientific research, but also for all private and public institutions (see DSM Strategy COM(2015) 192 final, para 4, Recital 18 DSM-D). It can serve as a basis for any kind of decision whether it be in the business or governmental sphere, as well as for the development of new systems, applications or technologies that can be crucial for the progress and existence of an economy on the international market (see Recital 18 DSM-D). For this reason, the EU has rightly taken its goal of an information society and a “flourishing” European data economy (Communication COM(2014) 442 final, para 1) seriously, also in favor of the Industry 4.0, and has launched with Arts. 3 and 4 DSM-D a first attempt to actually create legal certainty for the use of text and data mining techniques at all levels.

3.4 *The Limitation of § 60d Copyright Act*

Based on Art. 3 DSM-D, § 60d Copyright Act will soon be adapted. Although the introduction of the general TDM limitation in Art. 4 DSM-D is to be seen as an important achievement for the further development of Industry 4.0, the upcoming changes due to Art. 3 DSM-D are also not to be despised. Above all, for scientific research, the "right to read" now indispensably includes the "right to mine" (to this dispute Spindler 2020, p. 307). Art. 3 DSM-D obligates all Member States to introduce such a copyright limitation on a mandatory basis. It thus creates legal certainty for scientific research throughout the Union.

3.4.1 Scope: Non-Commercial Scientific Research

In accordance with the basis of EU law in Art. 5 (3) (a) of the InfoSoc-D, § 60d of the Copyright Act restricts the exception to non-commercial scientific research (Begr RegE BT-Drs. 18/12329, p. 23), which means that Industry 4.0 applications were in the past generally excluded for a number of reasons (for details: Spindler 2020, pp. 308 ff.).

In detail, § 60d (1) Copyright Act does not determine the group of privileged persons based on their status, but solely on the activity performed. As stated earlier, the restriction to scientific research, it being understood as a methodical and knowledge-based activity (Raue 2017b, p. 657; Jarass 2021, Art. 13 para 8; Calliess and Ruffert 2016, Art. 13 GRCh para 6), and no pursued commercial purposes is crucial. This understanding of scientific research includes the actual research activity as well as the presentation of its results (Raue 2017b, p. 657; Jarass 2021, Art. 13 para 8; Calliess and Ruffert 2016, Art. 13 GRCh para 6). This includes practitioners, private scholars and even students, regardless of institutional affiliation to a university or other research institution, provided they work independently scientifically (Spindler 2018, p. 279). The decisive factor for assessing whether commercial purposes are being pursued is the specific research activity, which means that the source of funding or the orientation of the institution that carries out the research are not crucial (Spindler 2018, p. 279). Overall, § 60d Copyright Act is only applicable if the research activity is not for profit, making it of no use to commercial research, e.g., Big Data analyzes of companies of the Industry 4.0. Consequently, the examination of the individual case is now particularly necessary and difficult concerning private third-party funded research at public universities and private-public partnerships (Spindler 2018, p. 280). However, whether a fee is paid for a later publication of research results from TDM projects stays out of consideration (RegE, BT-Drs. 18/12329, 39; Raue 2017b, p. 657).

The mandatory limitation is introduced for non-commercial research only, Art. 3 (1) DSM-D, which—under Art. 2 No. 1 DSM-D—includes only (European) research organizations and cultural heritage institutions, not—unlike § 60d Copyright Act (Raue 2017b, p. 657; Spindler 2018, pp. 279 f.)—individual researchers, doctoral students, or research communities outside their institutions. Under Art. 2 No. 1 DSM-D, research organizations include universities and their libraries as well as institutions whose primary goal is research—also in connection with teaching. The demarcation from commercial research is particularly relevant for Industry 4.0 purposes, as the lack of profit orientation or acting on a state-recognized mandate in the public interest are decisive (Recital 12). If fees are charged or there are expenses to access research results, the decisive factor is whether these are only intended to cover the costs incurred (Raue 2017b, p. 657; Spindler 2018, p. 280). As in § 60d Copyright Act, privately financed research should continue to be covered as long as its purpose is not aimed at commercial exploitation. Moreover, private-public partnerships between commercial companies and research institutions should still be possible as long as there is no commercial exploitation and there is no predominant influence of the private third party (Recitals 11, 12). Although the Directive does not explicitly deal with a privileged access to research results by the private party, the denial of an exemption for these cases is in line with Recital 12 and Art. 2 No. 1 (Raue 2019, p. 691; Spindler 2019, p. 279) Overall, private-public ownership can now be considered for Industry 4.0, although the whole concept must be understood rather restrictively.

3.4.2 Covered Works and Affected Exploitation Rights

§ 60d (1) Copyright Act refers to the utilization of a large number of works. The nature of these works is irrelevant; all works protected by copyright are covered, whether works of visual art, texts, or software. Pure ancillary copyrights are also covered by the referencing of ancillary copyrights to the exceptions, such as § 87c (1) No. 2 Copyright Act or § 72 (1) Copyright Act (Raue 2017b, p. 658). Similarly, it does not matter in which format the works are available, whether analog or digital; even the first digitization is covered by the limitation (Begr RegE BT-Drs. 18/12329, p. 41; Raue 2017b, p. 658). Art. 3 DSM-D does not deviate significantly in this respect. According to the provision, all acts of **reproduction** or **extraction** of copyrighted material from a database can be based on the TDM exception. However, they must be done with the goal of subsequent automated analysis in digital form. Researchers are therefore allowed to compile, store, and retain data bundles from different sources (Geiger et al. 2018, pp. 33 f.; Bottis et al. 2019, pp. 370 ff.). In addition to the acts of reproduction just mentioned, acts that are nationally considered as **adaptations** § 23 Copyright Act, are also authorized (Geiger et al. 2018, p. 29; Bottis et al. 2019, p. 386). In particular, the data material may be normalized. Normalization is a process of structuring and reorganising a database (Brumm 2017;[5] Watts 2020[6]). The aim is to avoid and eliminate redundancies, inconsistencies, and anomalies. Such normalization, as well as standardization and

[5] Available at: https://www.databasestar.com/database-normalization/.

[6] Available at: https://www.bmc.com/blogs/data-normalization/.

pre-structuring, is of enormous importance, especially for high-quality data analyses. It is therefore crucial for the Industry 4.0 that these processes can be carried out without major hindrances (transaction costs).

However, in the past it was controversial of whether the restriction applies only to works to which the researcher has legitimate access (in detail on this debate: Spindler 2020, pp. 309 ff). This question has now been finally clarified by the DSM Directive: in line with the prevailing opinion (Federal Government in the explanatory memorandum to § 60d Copyright Act: Begr RegE BT-Drs. 18/12329, S. 41; in detail on the question: Spindler 2020, pp. 309 f.) the exception does in fact only apply to works to which the researcher has legitimate access (Recital 18 DSM-D), such as works acquired under license or works in a library. However, it also applies to works freely and generally accessible via the Internet (cf. Recital 9 DSM-D).

Unlike German law, however, Art. 3 (1) DSM-D refers only to reproductions and extractions from databases (see also Recital 11). While German copyright law in its current version explicitly mentions the use of the corpus as a technique in § 60d, the DSM-D refrains from more detailed circumscriptions here; Art. 2 No. 2 DSM-D merely refers to automated analyses of texts and data in digital form. The national legislator has followed this path in § 60d Copyright Act Draft. The use of the corpus is no longer mentioned in § 60d Copyright Act Draft—however, it is made clear that the creation of a corpus is still possible despite the lack of a specific reference (Copyright Act Draft Bill[7] of 02.09. 2020, p. 93). In principle, all works protected by copyright are covered, including “other subject matter” or ancillary copyrights, but apparently not software. As already mentioned, it does not matter in which format the works are available, only the reproduction after or by text and data mining must be digital.

Under Art. 3 (2) DSM-D, the reproductions may also be kept without a time limit as long as it only serves scientific research including the verification by third parties which is indispensable for research purposes. In this respect, Art. 3 DSM-D deviates from the current TDM limitation in national law, which required the deletion of the corpus and original material immediately after the completion of the research work and the termination of any making available to the public, § 60d (3) Copyright Act. The draft bill shows that the national legislator wants to adapt § 60d according to the Directive, the obligation to delete the original material and the corpus was deleted without replacement, cf. § 60d Copyright Act Draft.

3.4.3 Author Designation and Source Citation

Apart from the limitation, the requirement to indicate the source was also a hurdle for TDM in Industry 4.0, Art. 5 (3) (a) InfoSoc-D, § 63 (1) sentence 1, (2) sentence

[7] Available at: https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/RefE_Urheberrecht.pdf;jsessionid=BAE5097EA5EFF51529D243F15115DD79.2_cid334?__blob=publicationFile&v=7.

2 Copyright Act. In relation to data or databases, the owner of the database had to be named, however this owner was not always clear (see below Sect. 4.3). § 63 (1) sentence 3 Copyright Act (unfortunately) does not address questions of reasonableness. Therefore, an interpretation in conformity with European law has rightly been demanded, especially with regard to § 60d (1) Copyright Act, which considers the automation and the then often missing reasonableness (Raue 2017b, p. 659). This problem has now been remedied, as Art. 3 DSM-D does, in contrast to Art. 5 (3) (a) InfoSoc-D, not demand a source citation. This is being considered at the national level—the Draft bill for the Copyright Act states that TDM will be exempted from the obligation under § 63 of the Copyright Act (Draft Bill 2020, p. 16).

3.4.4 Limitation for Database Rights

As shown, the property rights of databases are of particular importance for Industry 4.0 designs, but especially for the extraction of data. Similar to Art. 5 (3) (a) InfoSoc-D, Art. 6 (2) (b) Database-D contains an optional limitation for Member States. These can introduce limitations on copyrighted databases for (solely) **non-commercial research purposes**, as long as the **source** is mentioned (however, without any restriction on reasonableness) (see Walter and von Lewinski/*v. Lewinski*/*Walter* 2010, paras 11, 5, 46, who see no difference in this; other opinion Triaille et al. 2014, p. 70). Only a few Member States, such as Belgium, Spain, the UK, and Italy, have made use of this right (for the whole see Triaille et al. 2014, p. 68). In addition, the right to extract for scientific purposes (not for reuse) can also be found in Article 9 of the Database-D and § 87c (1) sentence 1 No. 2 of the Copyright Act. Again, in the past the non-commercial purpose and the indication of the source are required (§ 87c (1) sentence 2 Copyright Act). This cannot be upheld due to Arts. 3 and 4 DSM-D, scientific as well as commercial Text and data mining is now explicitly mentioned in § 87c Copyright Act Draft Bill.

Once again, it should be noted that generated data does not per se lead to database protection (see above Sect. 2.2.1), so that the question of limitations is naturally not relevant in the absence of database protection.

3.4.5 Digital Rights Management (DRM) or Technical Protection Measures and Restrictions on Access

Despite the far-reaching protection granted by the InfoSoc Directive to technical protection measures, Art. 6 (4) InfoSoc-D stipulates that the Member States must provide for the enforcement of the limitations mentioned in Art. 6 (4) (1) InfoSoc-D which also includes Art. 5 (3) (a) InfoSoc-D in favor of the scientific limitation. The Database Directive is also referred to here accordingly, so that the exemption also applies to it. However, Art. 6 (4) subpara. 4 InfoSoc-D immediately restricts this again, in that Art. 6 (4) subpara. 1 InfoSoc-D does not apply if the work has been made publicly accessible by contractual agreement, which applies in the majority of cases. Due to Art. 7 (2) DSM-D, the following picture now emerges:

The limitation now also prevails against technical protection measures, Art. 7 (2) DSM-D. Member States must oblige right holders to provide the beneficiary with the necessary means to use the exception despite technical protection measures, Art. 6 (4) subpara. 1, 3 and 5 InfoSoc-D. This also applies to content that is made publicly accessible, for example in databases, and technical protection measures prevent mining activities, hence § 95b Copyright Act has to be adapted. According to this provision, limitations do not prevail over technical protection measures if content is made available to the public based on contractual agreement and the user makes the individual choice regarding the time and place of use. § 95b (3) Copyright Act Draft reveals that the provision will be adapted in accordance with the requirements of Art. 7 DSM-D.

For Industry 4.0 designs, this ultimately means that any access limitation to data and databases leads to the limitation of § 60d Copyright Act as inapplicable—even if non-commercial research is involved.

3.4.6 TDM-Limitation and License Models

One of the most intensely disputed provisions of the reform concerned the relationship between limitations and licenses. In the end, the regulation of § 60g (1) Copyright Act prevailed, according to which licenses cannot fall short of the limitations resp. the limitations are semi-mandatory. Such a provision still seems necessary, since TDM researchers can hardly acquire each individual license for a large number of works, which are also subject to different conditions; also, the individual rights holder may not be traceable (as here Raue 2017b, p. 661; Truyens and van Eecke 2014, p. 167). This changes with the DSM Directive:

With Art. 7 (1) DSM-D, the decision was made to make the limitation mandatory and not disposable in a license contract. Under Art. 7 (1) DSM-D, the rightholder cannot enforce contractual provisions that contravene the text and data mining exception of Art. 3 DSM-D. This avoids the high transaction costs that arose under the previous legal situation as a result of litigation with rightholders. In the past, the problem could arise that the same content had to be licensed several times. Although mining was only permitted for content to which one had legitimized access anyway—so that a license already existed—for mining purposes a different file format (often XML) is often required than is provided, for example, in "normal" campus licenses or similar (PDF), so that a license then had to be obtained again.

Closely related to this is one of the Achilles' heels of all mandatory rights in copyright, namely enforcement in an international context, especially where there is a choice of law in international contracts (ignoring this, for example, Bruch and Pflüger 2014, pp. 394 f.). Therefore, the question arises whether the binding national law is still valid even in the case of a choice of law. The only mandatory rules recognized in European conflict of laws under Article 9 of the Rome I Regulation are those that serve compelling public interests. Thus, the Federal Court of Justice (still for Art. 34 Introductory Act to the German Civil Code) has rejected (BGH, case of 24.9.2014 – I ZR 35/11, GRUR 2015, p. 267 paras 45 ff. – Hi Hotel II; see also Loewenheim 2014, pp. 891 f.) the view that **the doctrine of transfer for a specific purpose** was mandatory law and thus immunized against a choice of law (of this opinion: Dreier and Schulze/*Dreier* 2018, Vor §§ 120 ff. UrhG paras 55; Schricker and Loewenheim/*Katzenberger* 2020, Vor §§ 120 ff. UrhG, paras 166 f., in this direction also Staudinger 2018, Art. 9 Rom I-VO para 30). Due to the unification intended by the conflict of laws, it must be assumed that in case of doubt, these are not binding mandatory rules (BGH, case of 24.9.2014 – I ZR 35/11, GRUR 2015, p. 267, paras 47 – Hi Hotel II with reference to BGHZ 165, p. 256 ff.). However, the TDM exception under § 60d Copyright Act does not concern contractual relationships between the author and the rights holder, but with users, which also serve the public interest, namely the promotion of research and development. Consequently, the limitation is likely to prevail over a choice of law in this case as well. Finally, the *Igmar* CJEU case can be cited, in which the court qualified the protection of the commercial agent as a binding mandatory norm, since this was intended to serve the unification of law (CJEU, case of 9.11.2000 – C-381/98, ECLI:EU:C:2000:605 – Ingmar GB).

3.4.7 Remuneration

In contrast to § 60d (in conjunction with § 60h) Copyright Act, Art. 3 DSM-D no longer provides a claim for remuneration (Recial 13; Stieper 2019, p. 213; Raue 2017b, p. 661). This is justified in Recital 17 by stating that the caused damage is minimal anyway (approving: Stieper 2020, p. 4; Spindler 2019, p. 281). Accordingly, § 60h, under which the scientific TDM was not an unremunerated use, was in the Draft Bill amended in line with the DSM Directive.

3.5 The New Limitation of § 44b Copyright Act

As already indicated, there has been a fierce debate about limitations for text and data mining (in detail: de la Durantaye 2017, pp. 561 f.; Geiger et al. 2018, pp. 13 ff.; Raue 2017b, p. 656; Spindler 2016, pp. 1118 f.) in the commercial sector and for non-research purposes, which—as already mentioned under Sect. 3.1—can have a considerable significance for Big Data applications and artificial intelligence. The compromise finally provides a limitation in Art. 4 DSM-D; but commercial TDM is however, in contrast to scientific TDM, subject to other licensing agreements, Art. 7 (1) DSM-D, as well as corresponding reservations of rights and rights of use Art. 4 (3) DSM-D (see below Sect. 3.4.2). In contrast to scientific TDM, it is hence not unrestrictedly valid that the right to read includes the right to mine (agreeing: Geiger et al. 2018, p. 30). The introduction of such a mandatory limitation for the Member States is nevertheless highly welcomed especially with regard to the Industry 4.0, as it can eliminate legal uncertainties, as Recital 18 of the Copyright Directive itself points out. According to the Draft Bill of the Copyright Act, the requirements of Art. 4 DSM-D will be considered by also introducing a new TDM limitation on commercial mining, § 44b Copyright Act Draft.

3.5.1 Covered Works and Exempted Acts of Exploitation

As with § 60d Copyright Act draft, the general TDM limitation allows **reproductions** for text and data mining. The limitation applies to works (including databases and **computer programs**), sui generis databases and other subject matter of copyright protection (Copyright Act Draft Bill of 02.09.2020, p. 93.). In the Discussion draft on the implementation of the Copyright Directive, § 44b still contained the addition that these reproductions must be necessary. This necessity criterion rightly met with criticism (Raue 2020, p. 172). On the one hand, because such a restriction does not arise from Art. 4 DSM-D and, on the other hand, because the dependence on this legal concept, which is up for interpretation, would again lead to legal uncertainty, which is precisely what the DSM Directive and its implementation are intended to prevent (Recital 10, 11, 18 DSM-D). The legislator has implemented the justified criticism and removed the necessity criterion in the current draft bill. Under Art. 4 (3) DSM-D, the copies may only be kept for as long as necessary for text and data mining—this is implemented nationally by a deletion obligation in § 44b (2) sentence 2 Copyright Act Draft Bill.

3.5.2 Opt-Out

In contrast to scientific TDM, it is possible for the rightholder to declare a reservation of use. As also follows from Art. 7 DSM-D, commercial TDM is therefore not mandatory and can be contracted around by a license contract. The corresponding reservation of use from Art. 4 (3) DSM-D was implemented in § 44b (3) Copyright Act Draft Bill. However, such a reservation of use is under Art. 4 (3) DSM-D only effective if it is "expressly declared in an appropriate manner"—this somewhat mitigates the severe restriction of commercial TDM. This mitigates the severe restriction at least to a certain extent (critical in this respect: Chiou 2020, p. 409). The fact that it is necessary to elaborate individually and manually whether there is a reservation of use would have made the use of TDM in the commercial sector extremely difficult—but the EU has stated that an appropriate manner for works published online means in a machine-readable form (Recital 18 DSM-D), so that the algorithms that are supposed to search the data can directly and automatically in advance determine whether there is a reservation of use. The specification of appropriateness was initially not included in the discussion draft (Discussion

Draft[8] of 15.01.2020, p. 5), which was rightly criticized (Raue 2020, p. 173) as this would cause unnecessary difficulties for the possibility of TDM. The national legislator has considered this circumstance and has now inserted the specification of appropriateness, cf. § 44b (3) sentence 2 Copyright Act Draft Bill. In the case of works that are not published online, the reservation must also be made in a manner that is appropriate to the automated processes of text and data mining, but this can also be provided by contractual provisions or in other ways (Draft Bill, p. 94; Recital 18 DSM-D).

3.5.3 Remuneration

The text and data mining limitation does not provide for remuneration, since Art. 4 DSM-D (in contrast to Art. 3 DSM-D) does not contain an equivalent declaration or any other possibility of introducing such a remuneration for Member States (Draft Bill of 02.09.2020, p. 93). The draft justifies this on the one hand with the fact that the rightholder is already able to forbid the general use of his content under Art. 4 (3) DSM-D (Draft Bill of 02.09.2020, p. 93). In addition, the intended legal certainty for users of text and data mining would be undermined, if it always had to be determined whether only an ephemeral copy was created or any other kind of duplication, having a direct influence on the obligation to pay (Draft Bill of 02.09.2020, p. 93).

3.6 Outlook

Art. 3 DSM-D has created the urgently needed union-wide legal certainty for text and data mining in scientific research. This is especially true on the scope of application of the limitation, which is further illuminated by detailed recitals guiding the interpretation. The enormous competitive disadvantage suffered by the European Data Economy due to the lack of a TDM limitation for commercial purposes has been remedied by the EU through Art. 4 DSM-D. Now, data analyses are possible in the private and also the scientific sector, enabling the advancement and further development of Industry 4.0. Although commercial TDM remains restricted in that a reservation of use or a contractual waiver can be declared, but these must at least be designed in such a way that they consider the automated processes of TDM, so that the transaction costs will nevertheless be significantly lower compared to the previous legal situation. It remains open to what extent the rights holders will actually make use of the possibility of reservation of use and whether they will adhere to the

[8] Available at: https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/DiskE_Anpassung%20Urheberrecht_digitaler_Binnenmarkt.pdf?__blob=publicationFile&v=1.

requirements to do so in an appropriate manner; should this be the case, Industry 4.0 technologies will benefit greatly.

4 Co-Authorship in Networks

As emphasized several times, the Industry 4.0 represents one of the most intensive interconnections within the economy without removing the boundaries of the legal person or contractual relationships and without crossing the boundary to corporate law (see chapter @ para @). However, new questions of allocation and partition of emerging rights arise, particularly from the sharing and creation of data.

4.1 Prerequisites According to § 8 Copyright Act (Analogue)

The development process is decisive for the copyright question about the moment and the circumstances a joint copyright in the database work can arise under § 4 (2) Copyright Act, or—probably more practically—a joint sui generis right of the database producer in data communitized in networks. A joint contribution to the further development of (online learning) AI systems is also conceivable (see Hartmann and Prinz 2018, p. 785). In all these cases, it is possible that various partners in an Industry 4.0 network will work simultaneously on a database and, over time, supply the data from their manufacturing and other processes. It is also feasible for a database to be created successively, quasi "one after the other" in the value chain. According to these different development models, it is also determined who can become the rights owner to a database.

If it is a joint concept, such as a database that should be shared, all data suppliers can be considered as a co-creator for a unified work (Schricker and Loewenheim/ *Loewenheim/Peifer* 2020, § 8 UrhG paras 8 f.; Möhring and Nicolini/*Ahlberg* 2018, § 8 UrhG paras 4 ff; Wandtke et al. 2019, § 8 UrhG paras 25 ff.; Fromm and Nordemann/*Wirzt* 2018, § 8 UrhG paras 2 f.; Omsels 2000, p. 141, 166 f.; Plaß 2002, p. 672). However, this requires that the contribution is made to a work that has been designed according to a joint plan or idea (see BGH, decision of 14.7.1993 - I ZR 47/91, BGHZ 123, p. 213 – Buchhaltungsprogramm; BGH, decision of 9.5.1985 - I ZR 52/83, NJW 1985, pp. 195 ff.; BGH, decision of 14.11.2002 – I ZR 199/00, GRUR 2003, p. 233 – Staatsbibliothek; see also BGH, decision of 13.6.2002 – I ZR 1/00, BGHZ 151, p. 92 – Mischtonmeister; OLG Düsseldorf, decision of 8.9.2015 - I-20 U 75/14, ZUM-RD 2016, p. 368, 372; OLG Frankfurt, decision of 17.9.2002 - 11 U 67/00 MMR 2003, pp. 46 f.; OLG Köln, decision of 14.10.1952 – 4 U 82/52, GRUR 1953, p. 499 – Kronprinzessin Cäcile I; Schricker and Loewenheim/ *Loewenheim/Peifer* 2020, § 8 UrhG para 9; Möhring and Nicolini/*Ahlberg* 2018, § 8 UrhG paras 4 ff.; Waldenberger 1991, pp. 26 ff.), whereby the individual contributions are not separately exploitable (BGH, decision of 14.11.2002 – I ZR 199/00, GRUR 2003, p. 233—Staatsbibliothek; in detail and in favor of a restrictive concept of co-authorship Waldenberger 1991, pp. 25 ff.). It should be noted again that (unfortunately) with regard to the sui generis protection under §§ 87a ff. Copyright Act, the generation of own data and the supply to the common database itself does not constitute a protected investment and therefore cannot establish database protection (see above Sect. 2.2.2 Schutzgegenstand). Accordingly, only the supply of "third-party" data or the development of the database structure, etc. by the Industry 4.0 partners is subject to database protection; the supply of own data is not.

Even the fact that a committee or an individual Industry 4.0 partner decides on the design and use of the database (detailed information on the mostly basic-democratic decision-making processes Grassmuck 2004, pp. 239 f.) does not change the fact that the individual participants in the Industry 4.0 network are and remain the "authors" of their respective contributions. However, it is still necessary that the plan is joint and that substantial investments are made together.

Less problems are posed by the implicit requirement of § 8 (1), (3) Copyright Act that co-authorship generally refers to different contributions of the same kind in the same category of works (such as for film works and the participation of musical authors Schricker 1986, pp. 76, 79; Fromm and Nordemann/*Wirtz* 2018, § 8 UrhG para 11). Because this is primarily about data that is merged into one database, but not other categories of work. If such co-authorship exists, § 8 (2) sentence 1 Copyright Act stipulates—possibly analogously for the sui generis right (Dreier and Schulze/*Schulze* 2018, § 87a UrhG, para 21; without express analogy Schricker/Loewenheim/*Vogel* 2020, § 87a UrhG para 75)—that a **community of joint ownership** arises between the co-authors of the database by operation of law, but regulates this only for the right of publication as well as the right of exploitation of the work, with the resulting problems of the joint decision on the exploitation of the joint property (näher Waldenberger 1991, pp. 63 ff.; Dreier/Schulze/*Schulze* 2018 § 8 UrhG paras 13 ff.; Eichelberger et al. *Seifert/Wirth* 2020, § 8 para 3). The regulations of the German partnership under §§ 705 ff. German Civil Code apply mutatis mutandis (see also Waldenberger 1991, pp. 39 ff. w. f. m.): Thus—in contrast to §§ 705 ff. German Civil Code—the community cannot be dissolved by termination, but ends only with the expiry of the term of protection, which is calculated after the death of the longest-living co-author, § 65 Copyright Act (Wandtke/Bullinger/*Thum* 2021, § 8 UrhG paras 51 ff. and Wandtke and Bullinger/*Lüft* 2019, § 65 UrhG paras 2 f; Möhring and Nicolini/*Freudenberg* 2018, § 65 UrhG para 4; Schricker and Loewenheim/*Loewenheim/Peifer* 2020, § 8 UrhG para 12.), respectively after the end of the protection period of § 87d Copyright Act. In the case of legal entities, which can be considered as "creators" of databases within the framework of §§ 87a ff. Copyright Act due to the investment protection (and not the creative) idea (Schricker and Loewenheim/*Vogel* 2020, § 87a UrhG para 72; Dreier/Schulze/*Schulze* 2018, § 87a UrhG, para 20), the relationship depends on the underlying agreement; which will usually also establish the German Civil Code company, or at least a fractional shared joint ownership under §

741 German Civil Code (Schricker and Loewenheim/*Vogel* 2020, § 87a UrhG para 75; Dreier and Schulze/*Schulze* 2018, § 87a UrhG, para 21).

In any case, on the rights of exploitation bound by joint ownership, the co-authors should also be subject to a solidary liability, for example, on the rights of third parties under § 97 Copyright Act. This is because they often act—precisely as required by § 8 Copyright Act—according to a joint plan, in that the database is developed together, so that the (common) connection in pursuit of a common purpose, which is the only decisive factor for a partnership (Habersack MünchKommBGB/*Schäfer* 2020, § 705 German Civil Code paras 95 f.; Bec Bamberger/Roth/Hau/*Schöne* 2020, § 705 German Civil Code para 62) can be assumed to exist. In contrast, a specific management, joint assets, or a general meeting of partners taking place in reality are not required for the assumption of a (partnership) company, even if such a company may no longer correspond to the structure which the user of the law traditionally imagines to be a partnership with its personal ties in the case of larger networks. At the same time, mass phenomena, such as the famous public limited partnerships, are by no means unfamiliar to partnership law.

4.2 Connected Work According to § 9 Copyright Act

However, it is also possible to create a so-called connected work, in which several independent works that were created independently of each other are linked together. This is also conceivable for databases. With the connected work, a civil law partnership is also created due to the purpose of joint exploitation of the works connected in this way (so F. A. Koch 2000, p. 277 f.; BGH, decision of 2.10.1981 – I ZR 81/79, GRUR 1982, pp. 42 f. – Musikverleger III; in detail von Becker 2002, p. 581 ff.; differentiating Fromm and Nordemann/*Wirtz* 2018, § 9 UrhG para 12; in general Schricker and Loewenheim/*Loewenheim/Peifer* 2020, § 9 UrhG paras 3, 9; Wandtke and Bullinger/*Thum* 2019, § 9 UrhG Rn. 56 ff.; Möhring and Nicolini/*Ahlberg* 2018, § 9 UrhG Rn. 12 ff.), without, however, the copyrights to the respective works thereby becoming jointly held property. However, the sustainability of such a construction stands and falls with the independence of the individual databases that are connected to each other (see for this Wandtke and Bullinger/*Thum* 2019, § 8 UrhG paras 7 ff.; Möhring and Nicolini/*Ahlberg* 2018, § 8 UrhG paras 10 ff.; Schricker and Loewenheim/*Loewenheim/Peifer* 2020, § 8 UrhG para 5; Fromm and Nordemann/*Wirtz* 2018, § 8 UrhG paras 16 ff.).

4.3 Consequences: Authority to Dispose of and Take Legal Action

Both co-authorship under § 8 Copyright Act and joint exploitation under § 9 Copyright Act are (at least) a joint ownership that holds the copyrights or sui generis property rights and—unlike the German Civil Code partnership – cannot be terminated or dissolved (see already Sect. 4.1).

This has numerous consequences for the right of disposal and the right of action. Although the company constituted under civil law (German Civil Code partnership) as a typical form of joint ownership has for some time even been recognized as capable of being a party and thus capable of partial legal capacity by case law (BGH, decision of 29.1.2001 – II ZR 331/00, NJW 2001, pp. 1056 ff. – ARGE Weißes Ross); this does not necessarily apply to every joint ownership and, moreover, does not change anything about the question of who can declare the granting and transfer of rights for the joint ownership. Thus, even in the case of a German Civil Code partnership and its partial legal capacity, the transfer of rights must in principle take place unanimously (BGH, decision of 02.10.1981 – I ZR 81/79, GRUR 1982, p. 43 – Musikverleger III). In addition, the filing of a lawsuit in a German Civil Code partnership also requires the consent of all shareholders, unless a special arrangement has been made for management and representation (on the GbR with partial legal capacity in civil proceedings: Kemke 2002, p. 2218 ff.; Wertenbruch 2002, p. 324 ff.; Habersack 2001, pp. 477 ff.; Ulmer 2001, pp. 585 ff.); an actio pro societate can only be considered in exceptional cases (compare to actio pro societate Habersack MünchKommBGB/*Schäfer* 2020, § 705 Rn. 210 ff.; Palandt/*Sprau* 2021 § 714 German Civil Code para 9). On copyright and co-authorship, § 8 (2) sentence 3 Copyright Act, as a special legal norm, also requires the suit for performance to all co-authors, even if only the individual author sues, so that the legal capacity of the German Civil Code partnership (BGH, decision of 29.1.2001 – II ZR 331/00, NJW 2001, pp. 1056 ff. – ARGE Weißes Ross) does not affect this either.

In practice, it is therefore crucial for a joint database whether the co-author as plaintiff can name the names of all other co-authors or those involved in the network of Industry 4.0; in this respect, procedural difficulties arise that have already arisen, for instance, in the development of open source software (compare to this Omsels 2000, p. 141 ff., 168; Koch 2000, p. 273, 279; Jaeger and Metzger 2020, paras 33 ff.). Only in cases where the data supplier and co-author leave his name or in the case of small communities a lawsuit will be possible. Although the presumption of authorship under § 10 (1) Copyright Act can help to a certain extent if several people can be named as co-authors, but it depends very much on the circumstances of the individual case whether enough co-authors can be named, so that a counter-evidence would fail (for parallel problems with Open-Source, Omsels 2000, p. 141, 168 f.).

In contrast, in the case of an injunction suit under § 8 (2) sentence 3 Copyright Act, an action by an individual co-author is also possible, even if the other co-authors cannot be named (Möhring and Nicolini/*Ahlberg* 2018, § 8 UrhG para 42). § 8 (2) sentence 3 subsection 2 Copyright Act refers in this respect only to actions for performance, whereas infringements of the joint copyright can generally be pursued by each author independently without obtaining the consent of the other joint authors—for example, in the context of an action for injunctive relief (likewise to understand Schricker and Loewenheim/*Loewenheim/Peifer* 2020, § 8 UrhG para 20). This already follows from § 8 (2) sentence 3 subsection 1 Copyright Act, according to which each co-author is entitled to assert claims arising from infringements of the joint copyright. In the case of a claim for injunctive relief, it is also not necessary to sue for payment to all co-authors, since there is no risk of overreaching in this case (Möhring and Nicolini/*Ahlberg* 2018, § 8 UrhG para 42; Wandke and Bullinger/*Thum* 2019, § 8 UrhG para 122; in contrast the prevailing opinion for § 1004 German Civil Code: Assertion in accordance with the joint ownership relationship thus collaborative, see MünchKommBGB/*Baldus* 2017, § 1004 para 50; Palandt/*Herrler* 2021, § 1004 German Civil Code para 14).

In a comparable manner, in the case of § 9 Copyright Act, either § 8 (2) sentence 3 Copyright Act analogously (Fromm and Nordemann/*Wirtz* 2018, § 9 UrhG para 13) or § 744 (2) German Civil Code (emergency administration) is applied (Schricker and Loewenheim/*Loewenheim/Peifer* 2020, § 9 UrhG Rn. 11), so that in this case, as well, the action must be filed naming all other authors.

5 Copyright Issues Related to Cloud-Based Applications

In Industry 4.0 applications, typical **cloud applications** will often arise in the context of networking, possibly also using a shared cloud between the partners. In addition to the contractual issues (see Chap @ Rn @), from a copyright point of view the various exploitations in the form of the necessary reproductions as well as the making available of works of any kind are the focus. Thus, software services as well as data or databases can be made available to the partners in the cloud, in the case of software in the form of so-called cloud—(comprehensive Bräutigam 2013; Marly 2018, paras 1117 ff.; from a technical perspective see also Liesegang 2015, pp. 776 ff.) or GRID-computing (on typology and delimitation: Argyriadou and Bierekoven 2018, §14 para 2 ff; see furthermore Marly 2018, para 1122).

In this context, the term **cloud computing** is understood as an IT delivery model based on virtualization, in which resources both in the form of infrastructure and applications and data are provided as a distributed service over the Internet by one or more service providers (hereto Pohle and Ammann 2009a, pp. 273 ff.; Niemann and Paul 2009, pp. 444 ff.; Schulz 2009, pp. 403 ff.; Karge and Sarre 2009, pp. 427 ff.; Schuster and Reichl 2010, pp. 38 ff.; see also Böhm et al. 2009, p. 8; Burgelnig and Mulholland 2009; Dunkel et al. 2008, pp. 270 ff.; Loewenheim/*Lehmann/Spindler*, § 82 UrhG paras 43 ff.), whereby these services are flexibly scalable as required and can be billed according to usage (definition from Böhm/Leimeister/Riedl/Kremar IM 02/2009, p. 8; Mann 2012, pp. 500 f.; Lehmann and Giedke 2013a, p. 608; Grützmacher 2011, p. 703). However, given the variety of products and approaches offered by cloud providers, a common definition for the totality of services offered under the term "cloud computing" has not yet emerged (Niemann and Paul 2009, p. 445; Henneberger et al. IM 02/2009, p. 20; Bisges 2012, p. 574; Lehmann and Giedke 2013b, p. 610; in detail about the different approaches Giedke 2013, pp. 36 ff.; The National Institute of Standards and Technology 2011,[9] p. 2). Similar to the ASP model, users no longer operate their IT infrastructure themselves in cloud computing, but obtain these resources via the Internet from a provider who operates both for them and other users in one or more data centers (Pfirsching IM 02/2009, p. 34). However, the services from the "cloud" go far beyond what an ASP provider makes available (in cloud computing: bundling of IT services), by providing not just application software but also hardware resources and system software as required (Schulz 2009, p. 404; Schuster and Reichl 2010, pp. 39 f.). In detail, cloud computing currently includes infrastructure services (Infrastructure-as-a-Service, IaaS, provision of computing power and storage space), the provision of application and development platforms (Platform-as-a-Service, PaaS) and also the services that were previously provided under the term Software-as-a-Service (SaaS) (The National Institute of Standards and Technology 2011,[10] pp. 2 f.; for an overview Marly 2018, paras 1117 ff.; in detail Giedke 2013, end of p. 27 and ff.; Fromm and Nordemann/*Czychowski* 2018, § 69c UrhG para 76a; Bräutigam and Thalhofer 2013, part 14 paras 12 ff.). In addition to this, Business Process as a Service (BPaaS) is sometimes mentioned (Sujecki 2012, p. 313).

5.1 Reproduction

The use of the application software (in the case of SaaS) or of data, databases, etc. by the cloud provider initially always constitutes reproduction as defined by §§ 16, 69c No. 1 Copyright Act, as the software (data, etc.) is installed and stored on the provider's server (Schneider 2017, chapter U para 24; Niemann and Paul 2009, p. 448; Spindler and Schuster/*Wiebe* 2019, § 69c UrhG para 61; Fromm and Nordemann/*Czychowski* 2018, § 69c UrhG para 76a in conjunction with para 76; Leupold/Glossner/*Wiebe* 2013, part 3 para 134; Leupold/Glossner/*Doubrava et al.* 2013, part 4 para 112; Grützmacher 2015, p. 785; Bisges 2012, p. 575). In principle, any storage of works, etc. in the cloud is a reproduction under §§ 16, 69c No. 1 Copyright Act, both the upload and the download (instead of many, Lehmann 2016, § 14 para 9).

This is only different if the cloud provider just provides the platform (operating system, etc.—Platform as a Service (PaaS)); in this case, the user retains control over any software, data, etc. used and thus the reproductions (Niemann 2009, pp. 662 ff.). However, it is argued that the provider has technical control over the system landscape, so that both the provider and the user are responsible for copying

[9] Abrufbar unter: http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.

[10] Abrufbar unter: http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.

processes (Giedke 2013, pp. 382 ff.; following this Grützmacher 2015, pp. 782 f.), in accordance with the statements of the Federal Court of Justice in the Internet video recorder decision, this is a normative attribution that also considers the legal potential influence and not just the technical control (BGH, decision of 22.4.2009 – I ZR 215/06, ZUM-RD 2009, 511 paras 16 f. – Internetvideorekorder I; in similar cases, also on the assumption of reproduction by the user: CJEU decision of 29.11.2017 – C-265/16, ECLI:EU:C:2017,913, Rn. 37 – VCAST/RTI.).

From a copyright perspective, the relationship between the cloud provider and the customer (in the case of SaaS) does not constitute software leasing requiring consent within the meaning of § 69c No. 3 Copyright Act, since leasing as a physical right of exploitation in copyright law requires the (physical) transfer of a copy (Marly 2018, para 1100 with further mentions; Wandtke and Bullinger/*Grützmacher* 2019, § 69c UrhG paras 43 f.; Grützmacher 2015, p. 784; Spindler and Schuster/*Wiebe* 2019, § 69c UrhG para 64; Leupold/Glossner/*Wiebe* 2013, part 3 para 137; Bisges 2012, p. 578; different view Lehmann 2016, § 14 para 11 with further mentions). The same applies to other works or to data (if these are protected by copyright at all). A subsequent provision for use by the user falls rather under § 19a Copyright Act (Wandtke/Bulinger/*Bullinger* 2014, § 19a UrhG paras 10, 12, 18, 23; Pohle and Ammann 2009a, p. 276; Niemann and Paul 2009, p. 448; Schuster and Reichl 2010, pp. 40 f.; Spindler and Schuster/*Wiebe* 2019, § 69c UrhG para 63, which directly refers to § 69c No. 4 Copyright Act; Leupold/Glossner/*Doubrava et al.* 2013, part 4 paras 116 ff., 119; Leupold/Glossner/*von dem Bussche and Schelinski* 2013, part 1 para 392; Bisges 2012, pp. 576 f.; different view: Schneider 2017, chapter G paras 233 ff., on the grounds that the characteristic “public” is lacking, because there is no majority of members of the public as customers, which is required after § 15 (3); according to Fromm and Nordemann/*Czychowski* 2018, § 69c UrhG para 76a in conjunction with para 76 the provider does not need a right of public reproduction), to § 69c No. 4 Copyright Act refers (Dreier and Schulze/*Dreier* 2018, § 69c UrhG para 28; Möhring and Nicolini/*Kaboth/Spies* 2018, § 69c UrhG para 28; different view Bisges 2012, p. 576 para 8: lex specialis).

5.2 Making Available to the Public

The right to making available to the public also covers the use of the software (Wandtke and Bullinger/*Grützmacher* 2019, § 69c UrhG para 81) or the protected work, so that it can be relevant for cloud services. However, the right is only affected by cloud services if the software or the work is not only made available to one person, but to a majority of members of the public (On this differentiating approach see Dietrich 2010, p. 568; see also Dreyer et al. *Kotthoff* 2018, § 69c UrhG para 24). Whether files are also transferred to the user’s computer in the process is irrelevant, because no such restriction can be inferred from the standard. Rather, the right to making available to the public, as a subcategory of the right of communication to the public, is based only on the **increased usage** by the public (Marly 2018, para 240;

Bisges 2012, p. 576; Spindler/Schuster/*Wiebe* 2019, § 19a UrhG para 2; different Wandtke and Bullinger/*Grützmacher* 2019, § 69c UrhG paras 85, 99, who wants to think about § 69c No. 4 depending on the technical situation with streaming; Paul and Niemann 2014, part 3, para 107). The right of making available to the public therefore already covers cases in which customers are given the opportunity to access the software as desired. The decisive factor is therefore that the program or work is made available to the public generally; technical coincidences, on the other hand, are just as irrelevant as the actual access by the cloud customers (OLG München, decision of 7.2.2008 - 29 U 3520/07, CR 2009, p. 502 paras 53 ff.; Giedke 2013*,* p. 398; Bräutigam and Thalhofer 2013, part 14 para 122; Marly 2018, paras 1101 ff., esp. 1104; Argyriadou and Bierekoven 2018, § 14 paras 17 ff.; Bettinger and Scheffelt 2001, p. 735; Söbbing 2015, p. 175; Möhring and Nicolini/*Kaboth/Spies* 2018, § 69c UrhG para 28; different opinion Grützmacher 2015, pp. 784 f.; Wandtke and Bullinger*/Grützmacher* 2019, § 69c UrhG para 99, which assumes § 69c No. 4 UrhG only if program parts and not merely graphic data are transferred; Grützmacher 2011, p. 705; Fromm and Nordemann/*Czychowski* 2018, § 69c UrhG para 76, if not the program itself, but only the screen mask is played back; also *Czychowski/Siesmayer* 2020, section 1, part 2, 20. 4, para 146). This result also applies to all forms of SaaS by the cloud provider. It is true that there are also technically different ways of accessing software via remote applications: Either a virtual machine is created on which the desired software is installed and run, or a program installed in the cloud is made available virtually to several cloud customers (Giedke 2013*,* pp. 399 ff., esp. 401). His differentiation is secondary, when it comes to the question of making available to the public in the sense of § 69c No. 4 as well as § 19a Copyright Act because it can be affirmed for both variants (Giedke 2013*,* pp. 400 ff.).

If the software is provided to a general public, the provider needs the **right to make it available to the public**. It is irrelevant, however, whether the user interface itself is protectable or whether program parts such as Java applets run on the client system (Wandtke and Bullinger/*Grützmacher* 2019, § 69c UrhG para 99; Nägele and Jacobs 2010, pp. 287, 290; Paul and Niemann 2014, part 3, paras 107 ff.; agreeing probably: Bräutigam and Thalhofer 2013, part 14 para 122; Bisges 2012, pp. 576 f.). The question of whether this is a type of streaming (Pohle and Ammann 2009a, p. 276; Pohle and Ammann 2009b, p. 629; Bierekoven 2010, pp. 43 f.; Splittgerber and Rockstroh 2011, p. 2179, who wants to apply §§ 44a or 69d), irrespective of the fact that technically there is usually no streaming of program data, but only of graphics data (Lehmann and Giedke 2013b, p. 682; Wandtke and Bullinger/*Grützmacher,* 2019, § 69c UrhG para 99) and no copies are made on the user's computer, but only at the provider's site (F. A. Koch 2011, p. 43; Wandtke and Bullinger/*Grützmacher* 2019, § 69c UrhG para 99) is also irrelevant. This also applies when it comes to the use of an operating system software as part of the Infrastructure Cloud Service (IaaS) or Platform as a Service (PaaS) (Pohle and Ammann 2009a, p. 276; Niemann and Paul 2009, p. 448; Giedke 2013, pp. 402 ff.; different opinion Wandtke and Bullinger/*Grützmacher* 2019, § 69c UrhG para 99: only one copy at supplier; also, Nägele and Jacobs 2010, p. 287).

5.3 *Type of Use and Intended Use*

From the perspective of the provider, the intended use of the software for cloud computing is a separate type of use, since the software is used to a particular extent and the use is technically as well as economically separable (Paul and Niemann 2014, part 3 paras 82 f., 146; Nägele and Jacobs 2010, p. 290; Pohle and Ammann 2009a, p. 276; Dorner 2011, p. 785; Wandtke and Bullinger/*Grützmacher* 2019, § 69d UrhG Rn. 13; Grützmacher 2011, p. 705; Giedke 2013, pp. 409 ff.; leaving open Bräutigam and Thalhofer 2013, part 14 para 123; affirming a separate type of use specifically for SaaS in relation to ASP Leupold/Glossner/*Doubrava et al.* 2013*,* part 4 para 120). Similarly the provision of other types of works in the cloud is handled. In contrast, the necessary processing and copying actions during the provision of the appropriately licensed software in the cloud by the provider already constitute a use for the intended purpose within the meaning of § 69d (1) Copyright Act (Paul and Niemann 2014, part 3 para 142). No deviation results if the reproduction is triggered by the user (Paul and Niemann 2014, part 3 paras 94, 142). Whether corresponding copyright-relevant **acts of use** also occur on the part of the user depends on whether the software is executed or the work is loaded at least in the user's main memory (Pohle and Ammann 2009a, p. 276; Spindler/Schuster/*Wiebe* 2019, § 69c UrhG para 65; Leupold/Glossner/*Wiebe* 2013 part 3 para 138; different opinion Niemann and Paul 2009, p. 448, on the grounds that any acts of reproduction take place exclusively in the cloud; Schuster and Reichl 2010, pp. 40 f.). A **justification** under § 69d (1) Copyright Act on the grounds of intended use of the computer program generally fails in the case of the end user due to his lack of authorization (Pohle and Ammann 2009a, p. 276; different opinion Fromm and Nordemann/*Czychowski* 2018, § 69c UrhG para 76a), since he has not originally acquired any contractual rights of use to the software in question (Wandtke and Bullinger*/Grützmacher* 2019, § 69d UrhG para 29; Grützmacher 2011, p. 704). However, if it is only a temporary intermediate storage of the software without an economic value of its own, a teleological reduction must also be made for software following the idea of § 44a Copyright Act, even if § 44a Copyright Act is based on the InfoSoc Directive, which is precisely not intended to apply to computer programs (leaving open BGH, order from 3.2.2011 - I ZR 129/08, GRUR 2011, p. 419 para 17 – UsedSoft I; Wandtke and Bullinger*/Grützmacher* 2019, § 69a UrhG para 85; Leupold/Glossner/*von dem Bussche and Schelinski* 2013, part 1 para 391; affirming the applicability of § 44a Fromm and Nordemann/*Czychowski* 2018, § 69a UrhG para 43, § 69c UrhG para 9; vague Dreier and Schulze/*Dreier* 2018, § 69a UrhG para 34 (leaving open), § 69c para 9 (affirmative), § 69d para 3 (§§ 69d Abs. 1–3, 69e are legis specialis); considering the applicability Pohle and Ammann 2009a, p. 276; negating § 44a UrhG in favor of § 69c Nr. 1 S. 2 UrhG Bisges 2012, p. 577; Hoeren 2006, pp. 576 f.).

As far as the **mere display on the screen** is concerned, the question of justification does not arise since, due to the lack of embodiment of the computer program therein, there is no reproduction within the meaning of § 69c No. 1 Copyright Act anyway. If the display on the screen is nevertheless based on a reproduction of another work protected by copyright in the main memory of the user system, § 44a Copyright Act may be applicable in this respect (also Bisges 2012, p. 577). If the cloud user himself provides the software (within the framework of IaaS or PaaS), the resulting copyright-relevant actions—whether those of the user or also of the cloud provider—can in any case unproblematically be assumed to be intended use within the meaning of § 69d (1) Copyright Act if the underlying license expressly permits the use of the software within the framework of IaaS and/or PaaS (Paul and Niemann 2014, part 3 paras 94, 115, 143 f.). However, if the cloud user is only the holder of a single-user, network or terminal license, use for the intended purpose within the meaning of § 69d (1) Copyright Act only exists if the license agreement does not contain any provisions to the contrary (e.g., hosting clauses) and the use in the cloud is also otherwise consistent with the license, e.g., there is no expansion of the user group (Paul and Niemann 2014, part 3 paras 81, 146, 194; see also Nägele and Jacobs 2010, p. 290; Grützmacher 2011, pp. 704 f.; Giedke 2013, p. 425).

6 Conclusion

In summary, the networked industry also poses a variety of challenges to copyright law. Individual, non-personal (possibly machine-generated) datum is not subject to its own copyright protection—despite its proximity to intellectual property law (see Sect. 2.1). However, indirect protection of data in “Copyright 4.0” is particularly provided by database protection under §§ 87a ff. of the Copyright Act, although this is subject to numerous restrictions in terms of the requirements for protection and its scope (see Sect. 2.2). Reliable protection of generated data is therefore left *de lege lata* to contractual agreements.

The in-depth use and analysis of existing data sets (e.g., by and for the use of artificial intelligence) collides with the right of reproduction under§ 16 and the right of making available to the public under § 19a Copyright Act (see Sect. 3.2). Legally, TDM applications are now generally permitted for all purposes. This is very positive for Industry 4.0—although commercial TDM is subject to the restriction that a reservation of use can be declared, the transaction costs that the previously existing legal uncertainty entailed have been significantly reduced.

The increasing networking of economic actors also entails the growth of co-authorship relationships (e.g., between the various producers of a common database) (see Sect. 4.1), but this is manageable, also in view of the problems of connection under civil law and civil procedure (see Sect. 4.3).

On the other hand, the copyright assessment of cloud-based applications in Industry 4.0 poses fewer problems, as the acts of use to be ascertained (see Sect. 5.1, 2) can in any case be permitted with corresponding license agreements, apart from the legal justification (Sect. 5.3).

References

Ahlberg H, Götting H-P (eds) (2020) BeckOK Urheberrecht Gesetz über Urheberrecht und verwandte Schutzrechte, 29th edn, Munich

Argyriadou I, Bierekoven C (2018) Application service providing/SaaS (Cloud). In: Intveen M, Gennen K, Karger M (eds) Handbuch des Softwarerechts, Bonn, pp 371 ff

Bamberger HG, Roth H, Hau W, Poseck R (Hrsg.) (2020) BeckOK BGB, 56. Edition, München

Bettinger T, Scheffelt M (2001) Application Service Providing: Vertragsgestaltung und Konflikt-Management, CR 729 ff

Bierekoven C (2010) Lizenzierung in der Cloud, ITRB 42 ff

Bisges M (2012) Urheberrechtliche Aspekte des Cloud Computing, MMR 574 ff

Böhm M, Leimeister S, Riedl C, Krcmar H (2009) Cloud Computing: Outsourcing 2.0 oder ein neues Geschäftsmodell zur Bereitstellung von IT-Ressourcen?, IM 6 ff

Bottis M, Papadopoulos M, Zampakolas C, Ganatsiou P (2019) Text and data mining in Directive 2019/790/EU enhancing web-harvesting and web-archiving in libraries and archives. Open J Philosophy 369 ff

Bräutigam P (2013) IT-outsourcing und cloud computing, Berlin

Bräutigam P, Thalhofer T (2013) Cloud-computing. In: Bräutigam P (eds) IT-outsourcing und cloud computing, Berlin, pp 1191 ff

Bruch C, Pflüger T (2014) Das Zweitveröffentlichungsrecht des § 38 Abs. 4 UrhG, ZUM 389 ff

Burgelnig E, Mulholland A (2009) Cloud Computing: Hype, Wegbereiter oder Spielerei, IM 15 ff

Calliess C, Ruffert M (eds) (2016) EUV/AEUV Das Verfassungsrecht der Europäischen Union mit Europäischer Grundrechtecharta Kommentar, 5th edn. Munich

Chiou T (2020) Copyright lessons on machine learning: what impact on algorithmic art?, JIPITEC, p 398

Clark J (2012) Text mining and scholarly publishing, Loodsrecht

Czychowski C, Siesmayer L (2020) In: Kilian W, Heussen B (eds) Computerrechts-Handbuch, Loseblatt, Stand: 07/2020

de la Durantaye K (2014) Allgemeine Bildungs- und Wissenschaftsschranke, Münster

de la Durantaye K (2017) Neues Urheberrecht für Bildung und Wissenschaft - eine kritische Würdigung des Gesetzentwurfs, GRUR 558 ff

Dietrich N (2010) ASP – öffentliche Zugänglichmachung oder unbenannte Nutzungsart?, ZUM 567 ff

Dorner M (2011) Umfassende Nutzungsrechteeinräumung gegen Pauschalabgeltung – Ende für "Buy-outs"?, MMR 780 ff

Doubrava C, Münch I, Leupold A (2013) In: Leupold A, Glossner S (eds) Münchner Anwalts Handbuch IT-Recht, 3rd edn. Munich

Dreier T, Schulze G (eds) (2018) Urheberrechtsgesetz, 6. Aufl., München

Dreyer G, Kotthoff J, Meckel A, Hentsch MA (Hrsg.) (2018) Urheberrecht, 4. Aufl., Heidelberg

Dunkel J, Eberhart A, Fischer S, Kleiner C, Koschel A (2008) System-Architekturen für verteilte Anwendungen, Munich

Dusollier S (2013) In: Triaille J-P (eds) Study on the application of Directive 2001/29/EC on copyright and related rights in the information society (the "Infosoc Directive"), European Union Brussels

Ehmann T (2014) Big Data auf unsicherer Grundlage - was ist "wesentlich" beim Investitionsschutz für Datenbanken?, K&R 394 ff

Eichelberger J, Wirth T, Seifert F (Hrsg.) (2020) Urheberrechtsgesetz, Handkommentar, 3. Aufl., Baden-Baden

Feldman R, Sanger J (2006) The text mining handbook. Cambridge University Press, Cambridge

Fromm FK, Nordemann W (eds) (2018) Urheberrechtsgesetz, 12th edn. Stuttgart

Geiger C, Frosio G, Bulayenko O (2018) The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects. Centre for International Intellectual Property Studies (CEIPI) Research Paper No. 2018-02

Giedke A (2013) Cloud Computing: eine wirtschaftliche Analyse, Munich
Grassmuck V (2004) Freie Software zwischen Privat- und Gemeineigentum, Bonn
Grimmelmann J (2016) There's no such thing as a computer-authored work – and it's a good thing, too. Colum J Law Arts 39:403 ff
Grützmacher M (2011) Lizenzgestaltung für neue Nutzungsformen im Lichte von § 69d UrhG (Teil 2), CR 697 ff.
Grützmacher, M (2015) "Software aus der Datendose" - Outsourcing, Cloud, SaaS & Co., CR 779 ff
Grützmacher M (2016) Dateneigentum - ein Flickenteppich, CR 485 ff
Habersack M (2001) Die Anerkennung der Rechts- und Parteifähigkeit der GbR und der akzessorischen Gesellschafterhaftung durch den BGH, BB 477 ff
Habersack M (ed) (2020) Münchener Kommentar zum Bürgerlichen Gesetzbuch Band 6 Schuldrecht – Besonderer Teil IV §§ 705 – 853 Partnerschaftsgesellschaftsgesetz Produkthaftungsgesetz, 8. edition, Munich
Handig C (2012) Durch "freie kreative Entscheidungen" zum europäischen urheberrechtlichen Werkbegriff, GRUR Int. 973 ff
Hartmann C, Allan J, Hugenholtz PB, Quintais J, Gervais D (2020) Trends and Developments in Artificial Intelligence – Challenges to the Intellectual Property Rights Framework – Final Report of the European Commission
Hartmann F, Prinz M (2018) Immaterialgüterrechtlicher Schutz von Systemen Künstlicher Intelligenz, DSRITB, 769 ff
Hearst M (1999) Untangling text data mining. In: Association for Computational Linguistics (editor): 37th Annual Meeting of the Association for Computational Linguistics - Proceedings of the Conference, Maryland
Henneberger M, Ruess C, Sessing G (2009) Innovationsberatung am Beispiel Cloud Computing, IM 18 ff
Hetmank S, Lauber-Rönsberg A (2018) Künstliche Intelligenz – Herausforderungen für das Immaterialgüterrecht, GRUR 574 ff
Heyer G, Quasthoff U, Wittig T (2006) Text Mining: Wissensrohstoff Text, Dortmund
Hoeren T (2006) Der urheberrechtliche Erschöpfungsgrundsatz bei der Online-Übertragung von Computerprogrammen, CR 573 ff
Jaeger T, Metzger A (2020) Open source software, 5th edn. Munich
Jarass H (2021) Charta der Grundrechte der Europäischen Union: GRCh, 4th edn. Munich
Karge M, Sarre F (2009) Wird Cloud Computing zu neuen juristischen Herausforderungen führen? In: Taeger J, Wiebe A (eds) Inside the Cloud – Neue Herausforderungen für das Informationsrecht, Oldenburg, pp 427 ff
Kemke A (2002) Die Gesellschaft bürgerlichen Rechts im Prozess - Parteibezeichnung und Kostendrittwiderklage gegen die Gesellschafter, NJW 2218 ff
Koch FA (2000) Urheber- und kartellrechtliche Aspekte der Nutzung von Open-Source-Software, CR 273 ff
Koch FA (2011) Client Access License – Abschied von der Softwarelizenz?, ITRB, 42 ff. Koch: Muss
Lehmann M (2016) In: Borges G, Meents J (eds) Cloud computing. C.H. Beck, Munich
Lehmann M, Giedke A (2013a) Cloud Computing - technische Hintergründe für die territorial gebundene rechtliche Analyse, CR 608 ff
Lehmann M, Giedke A (2013b) Urheberrechtliche Fragen des Cloud Computings, CR 681 ff
Leistner M (2007) "Last exit" withdrawal?, K&R 457 ff
Liesegang W (2015) Technische Aspekte der Fernnutzung von Software im Kontext urheberrechtlicher Fragen, CR 776 ff
Loewenheim U (2014) Eingriffsnormen ins Urheberrecht. In: Büscher W, Erdmann W, Haedicke M, Köhler H, Loschelder M (eds) Festschrift für Joachim Bornkamm zum 65. Geburtstag, pp. 887 ff

Mann M (2012) Vertragsgestaltung beim IT-Outsourcing - Besonderheiten und Fallstricke, MMR 499 ff
Marly J (2018) Praxishandbuch Softwarerecht, 7th edn. C.H. Beck, Munich
Metzger A (2012) Der Einfluss des EuGH auf die gegenwärtige Entwicklung des Urheberrechts, GRUR 118 ff
Meys R (2020) Data Mining Under the Directive on Copyright and Related Rights in the Digital Single Market: Are European Database Protection Rules Still Threatening the Development of Artificial Intelligence?, GRUR Int. 457 ff
Möhring P, Nicolini K. (Begr.) (2018) Urheberrecht, 4th edn, Munich
Nägele T, Jacobs S (2010) Rechtsfragen des Cloud Computing, ZUM 281 ff
Niemann F (2009) Shift der urheberrechtlichen Verwertungsrechte in der arbeitsteiligen digitalen Welt, CR 661 ff
Niemann F, Paul J-A (2009) Bewölkt oder wolkenlos - rechtliche Herausforderungen des Cloud Computings, K&R 444 ff
Omsels H-J (2000) Open Source und das deutsche Vertrags- und Urheberrecht. In: Schertz C, Omsels H-J (Hrsg.): Festschrift für Paul W. Hertin zum 60. Geburtstag, Munich
Palandt O (ed) (2021) Bürgerliches Gesetzbuch: BGB mit Nebengesetzen, 80th edn. Munich
Paul J-A, Niemann F (2014) In: Hilber M (ed) Handbuch Cloud Computing, Cologne
Pfirsching V (2009) Wertbeitrag des Cloud Computing: für die meisten Unternehmen noch gering, IM 34 ff
Plaß G (2002) Open Contents im deutschen Urheberrecht, GRUR 670 ff
Pohle J, Ammann P (2009a) Über den Wolken. . . - Chancen und Risiken des Cloud Computing, CR 273 ff
Pohle J, Ammann P (2009b) Software as a Service - auch rechtlich eine Evolution?, K&R 625 ff
Ramalho A (2017) Will robots rule the (artistic) world? A proposed model for the legal status of creations by artificial intelligence systems, J Internet Law 12 ff
Raue B (2017a) Das Urheberrecht der digitalen Wissen(schaft)sgesellschaft, GRUR 11 ff
Raue B (2017b) Text und Data Mining, CR 656 ff
Raue B (2019) Rechtssicherheit für datengeschützte Forschung, ZUM 684 ff
Raue B (2020) Die geplanten Text und Data Mining Schranken (§§ 44b und 60d UrhG-E), ZUM 172 ff
Rosati E (2019) Copyright as an Obstacle or an Enabler? A European perspective on text and data mining and its role in the development of AI creativity. Asia Pac Law Rev 27:198
Sattler A (2017) In: Sassenberg T, Faber T (eds) Rechtshandbuch Industrie 4.0 und Internet of Things, Munich
Schack H (2017) Das neue UrhWissG – Schranken für Unterricht, Wissenschaft und Institutionen, ZUM 802 ff
Schmidt K J, Zech H (2017) Datenbankherstellerschutz für Rohdaten?, CR 417 ff
Schneider J (Hrsg.) (2017) Handbuch EDV-Recht, 5th edn. Cologne
Schricker G (1986) In: Poll G (Hrsg.): Videorecht, Videowirtschaft, München, p 76
Schricker G, Loewenheim U (eds) (2020) Urheberrecht, 6th edn. Munich
Schulz C (2009) Rechtliche Aspekte des Cloud Computing im Überblick. In: Taeger J, Wiebe A (eds) Inside the Cloud – Neue Herausforderungen für das Informationsrecht, Oldenburg, pp 403 ff
Schulze G (2009) Schleichende Harmonisierung des urheberrechtlichen Werkbegriffs? - Anmerkung zu EuGH “Infopaq/DDF”, GRUR 1019 ff
Schuster F, Reichl W (2010) Cloud Computing & SaaS: Was sind die wirklich neuen Fragen?, CR 38 ff
Söbbing T (2015) SaaS-Verträge, ITRB 172 ff
Spindler G (2016) Text und data mining – urheber- und datenschutzrechtliche Fragen, GRUR 1112ff
Spindler G (2018) Text- und Datamining im neuen Urheberrecht und in der europäischen Diskussion, ZGE 273 ff

Spindler G (2019) Die neue Urheberrechts-Richtlinie der EU, insbesondere “Upload-Filter” - Bittersweet?, CR 277 ff
Spindler G (2020) Urhebrrecht 4.0. In: Frenz W (ed) Handbuch Industrie 4.0, Berlin, pp 297 ff
Spindler G, Schuster F (eds) (2019) Recht der elektronischen Medien, 4th edn. Munich
Splittgerber A, Rockstroh S (2011) Sicher durch die Cloud navigieren - Vertragsgestaltung beim Cloud Computing, BB 2179 ff
Staudinger A (2018) In: Ferrari, F Kieninger E-V, Mankowski P, Otte K, Saenger I, Schulze G J, Staudinger A (eds) Internationales Vertragsrecht Rom I-VO, CISG, CMR, FactÜ Kommentar, 3rd edition
Stieper M (2019) Die Richtlinie über das Urheberrecht im digitalen Binnenmarkt, ZUM 211 ff
Stieper M (2020) Das Verhältnis der verpflichtenden Schranken der DSM-RL zu den optionalen Schranken der InfoSoc-RL, GRUR 1 ff
Sujecki B (2012) Internationales Privatrecht und Cloud Computing aus europäischer Perspektive K&R 312 ff
Triaille J-P, Francquen A, Meeûs d’Argenteuil J (Hsrg) (2014) Study on the legal framework of text and data mining (TDM), Brussels
Truyens M, van Eecke P (2014) Legal aspects of text mining. CLSR 30:153 ff
Ulmer P (2001) Die höchstrichterlich “enträtselte” Gesellschaft bürgerlichen Rechts - Zugleich Besprechung zu BGH ZIP 2001, 330, ZIP 858 ff
von Becker B (2002) Rechtsprobleme bei Mehr-Autoren-Werkverbindungen, ZUM 581 ff
von dem Bussche A, Schelinski T (2013) In: Leupold A, Glossner S (eds) Münchner Anwalts Handbuch IT-Recht, 3rd. edition, Munich
Waldenberger A (1991) Die Miturheberschaft im Rechtsvergleich, Munich
Walter M, von Lewinski S (eds) (2010) European copyright law a commentary. Oxford University Press, Oxford
Wandtke A-A, Bullinger W (eds) (2019) Praxiskommentar zum Urheberrecht, 5th edn. Munich
Weiss S, Indurkhya N, Zhan T (2015) Fundamentals of predictive text mining, 2nd edn. Heidelberg
Wertenbruch J (2002) Die Parteifähigkeit der GbR - die Änderungen für die Gerichts- und Vollstreckungspraxis, NJW 324 ff
Wiebe A (2013) In: Leupold A, Glossner S (eds) Münchner Anwalts Handbuch IT-Recht, 3rd edn. Munich
Wiebe A (2014) Der Schutz von Datenbanken - ungeliebtes Stiefkind des Immaterialgüterrechts, CR 1 ff
Wiebe A (2016) Protection of industrial data – a new property right for the digital economy?, GRUR Int. 2016, 877 ff
Wiebe A (2017) Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken, GRUR 338 ff.

Industry 4.0 and Competition Law

Walter Frenz

Contents

1 Possible Starting Points for Antitrust Law Violations

Industry 4.0 can lead to significant competition law problems, but it does not have to. Basic starting points arise on one side out of the assignment of data, as previously mentioned in the set of problems regarding the involvement of several enterprises (this complex was concerned with the disputable question of to what extent there are access and exchange claims between the parties involved) as well as the issue which

W. Frenz (✉)
Unit for Mining, Environmental and European Law, RWTH Aachen University, Aachen, Germany
e-mail: frenz@bur.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_14

cooperations are allowed by competition law. On the other hand, there is the increasing transparency of marketplaces, which softens the actual competition securing isolation of individual economic actors (Spindler 2018, p. 164) through the informal exchange of information or by business partnering.

1.1 Dominant Market Position of Corporations

1.1.1 Software Solutions and the Cloud

The evolving 'Industry 4.0' can cause major economic advances and thus lead to dominant market positions of individual corporations. They are the result of the desired performance in line with the market: progress through success in competition. The **formation** of a dominant market position itself with **means in line with the market** does **not cause a violation of the competitive prohibition of abuse** (ECJ, Case C-52/09, ECLI: EU: C:2011: 83 (paragraph 24)—TeliaSonera; Case 209/10, ECLI:EU:C:2012:172 (paragraph 21)—Post Danmark).

It is therefore the conduct of a dominant undertaking, which is crucial. The assessment can be different if Industry 4.0 requires special expertise, which for example can only be obtained through other corporate entities. Such an indispensable basis is formed above all by certain devices and developments that can and should be built on. This is especially true for **software** that is used to digitise business operations as well as for the **cloud** to securely store the growing amount of data. If providers attain a **dominant market position** by controlling these central instruments of digitisation, the competitive prohibition of abuse in line with Art. 102 TFEU becomes relevant in terms of competition law, but it must not be misused to impose high prices or complicated regulations on other companies or to exclude them from access to maintain their own technical advantage or to remain or become a monopoly provider.

The case **Amazon** stands for the first constellation. This corporation wanted to maintain the monopoly through its electronic book sales platform and forced booksellers to offer generous discounts, thus imposing unreasonable conditions on them (see Frenz 2015, p. 1206). The second constellation is exemplified by the Microsoft rulings. **Microsoft** had to ensure the compatibility of its platform for individual components or operating systems from other providers by disclosing the interface information for the server-client interaction to its competitors so that they can advance their developments.[1]

[1] CJEU, Case T-201/04, ECLI:EU:T:2007:289—Microsoft I; Case T-167/08, ECLI:EU:T:2012:323—Microsoft II; Frenz (2016b), 671 also for following with further examples.

1.1.2 Merging of User Data

The prohibition of abusive conduct further limits the possibilities of comprehensive **processing** and use of **personal data** based on the terms of service ('**abuse of conditions**') as the Federal Cartel Office decided on 7 February 2019 at the expense of Facebook. This is primarily a burden on private individuals and less on companies. The Federal Cartel Office refers to the basic right of informational self-determination. However, corporate data can also be affected by a merger, for example if a software company requires merging data as a condition for using its software for digitisation purposes. The property of the company may then be violated. This is another reason why it is important to clarify who owns the data (see Frenz, Who owns the data and who protects them?).

The reference point for a violation of the prohibition of abuse is a dominant position of such a system, through which users communicate on a larger scale and which can therefore merge large amounts of data. Companies that exclusively offer these kinds of services or dominate the marketplace are usually in such a position. Therefore, they are not allowed to abuse their power. The conditions required for its use must not violate **fundamental and data protection law evaluations**. According to the Federal Cartel Office, it is not necessary for data to be processed to use Facebook. This also contradicts evaluations under data protection law because there is no justification against the background of the **EU General Data Protection Regulation.**

The European Court of Justice is also very strict about data protection law. It is therefore to be expected that the ECJ will adopt the BGH's approach chosen by the Federal Cartel Office. In addition, this approach should be expanded to include **corporate data**, as its processing is not necessary either when using a communication platform, or a search engine. To require acceptance for this in the terms of use is therefore abusive.

1.2 Business Partnering

As previously unrelated industries cooperate in the area of Industry 4.0 and thus come up with new products, this will primarily be the starting point for competition violations. In this way, initial **technological innovations** that are basically justifiable within the framework of Article 101 (3) TFEU can be achieved. However, this requires necessity. Moreover, care must be taken to ensure that the competition is still maintained. This applies, for example, to the diversity of the production systems developed through collaboration in the context of Industry 4.0 respectively the resulting products that are still in competition.

Thereby it is less problematic if smaller businesses without significant market shares collaborate. If, on the other hand, it is a question of large companies, especially if they are competing with one another, such partnerships can have a greater impact on competition, particularly if this creates a dominant market position through several corporate entities. For this, a cooperation itself is sufficient, as long as there is an external appearance as a collective unit.[2]

On the side of the **customer**, a **cartel** can arise in such a way that the customers join forces to dictate the competitive conditions for the provider. However, the cooperation between companies at the same level tends to be more problematic than between companies at different levels and thus between providers and customers. A violation of competition law can only exist if one of the involved parties has sufficient market power (European Commission, Guidelines on Vertical Restraints, OJ. 2010 C 130, p. 1, paragraph 6).

1.3 Exchange of Information

Industry 4.0 is about the networking of vertical processes, i.e. within the scope of the supply chain, as well as horizontal cooperation, which can result in completely new products. In both variants, Industry 4.0 is essentially based on the **exchange of information**. It is a part of the development or production of a product and therefore an **integral part of the manufacturing process**. It can therefore hardly be a measure that is prohibited under commercial law, otherwise the product development and manufacturing itself would be made impossible. This regularly applies to supply chains. Another level is reached when competing companies cooperate (see Sect. 1.2). The exchange of market-relevant information between competitors, for example about **corporate strategies and prices**, can already be anti-competitive. Incidentally, the exchange of information in the context of Industry 4.0 is naturally required and must not be refused. If a dominant position is exploited as a result, there may be a violation of the prohibition of abuse under competition law (see Sect. 1.1 and in this context Spindler 2018, p. 164).

It is also problematic to define standards that guarantee the exchange of information more closely or relate to the handling and business assignment of the data. Agreements can come about here, too, for example, which standards should be used between different companies that have parallel interests and thus want to practically exclude competitors from the exchange of information through cooperation. This can for instance be done by defining certain access codes, which are closed to the non-participating companies. Such an approach practically excludes their access and hinders the development in competition. It also applies when companies cooperate to the detriment of competitors by setting standards within the framework of company associations (see the following paragraph).

[2]ECJ, Cases C-395 and 396/96 P, ECLI:EU:C:2000:132 (paragraph 44)—Compagnie maritime belge transports and others.; Schröter and Bartl (2015), Art. 102 TFEU (paragraph 83 et seq.).

2 Necessary Access to Software Solutions and the Cloud

2.1 Conditions

If market participants have to fall back on dominant corporate entities for the required software or access to a cloud to implement Industry 4.0 in their operations, **a denial of access** raises the question of a violation of the **prohibition of abuse** in line with Article 102 TFEU. This was evident in the case of *IMS Health*, where a modular system protected by immaterial law was developed, which emerged as the standard for market reports on regional sales of pharmaceuticals and health products and was therefore absolutely necessary. Market entry depended on obtaining a licence to use this system, which is why it had to be granted. There was no justification for a refusal.[3] Similarly, in the area of Industry 4.0, a software solution can establish itself as a standard to develop individual solutions for the individual actors on this basis. Access to it is then essential to develop solutions that are tailored to the individual case. Especially the refusal of **access to product platforms** is a constellation of unjustified refusal to deliver.[4] A current example is the use of the electrification kit (MEB) from Volkswagen for E-mobility. Volkswagen opened it to the Aachen company E.Go (see E-mobility in the contribution by Kampker). This access can particularly be made more difficult or impossible if the standards to be complied with are unnecessarily complicated, which results in barriers to entry (Spindler 2018, p. 164). This also applies to the setting of standards by business associations (see the following article).

As a starting point, companies that dominate the market can also legally exploit their success (ECJ, Case C-52/09, ECLI:EU:C:2011:83 (paragraph 39)—TeliaSonera) and, in principle, freely choose their business partners. The Court of Justice of the European Union also recognised this in principle in the conflict between antitrust law and intellectual property law in the Microsoft case (CJEU, Case T-167/08, ECLI:EU:T:2012:323 (paragraph 119)—Microsoft II). Even monopoly-like intellectual property rights do not have to be disclosed without restriction, and neither do other trade secrets (see Spindler 2018, p. 165).

Extraordinary circumstances must therefore arise so that a refusal of access can constitute a violation of the prohibition of abuse.

At first, the refusal must concern products or services that are essential for the pursuit of a specific activity in a neighbouring market. It furthermore must be able to preclude any effective competition in that market. Eventually, it must prevent the emergence of a new product for which there is potential consumer demand (ECJ, Case—C-418/01, ECLI:EU:C:2004:257 (paragraph 38)—IMS Health).

---

[3] ECJ, Case C-418/01, ECLI:EU:C:2004:257 (paragraphs 28 et seq.)—IMS Health; thereto Frenz (2016a), paragraphs 2068 et seq.

[4] See Weidenbach et al. (2012), pp. 66 et seq. As well as specifically for a software CJEU, Case T-201/04, ECLI:EU:T:2007:289, (paragraphs 374 et seq.)—Microsoft I.

Specifically for the exercise of an intellectual property right, which can also include the use, dissemination and opening of a platform for Industry 4.0, the case law requires that the refusal to perform **prevents the emerging of a new product** for which there is a potential consumer demand. Under Article 102 sentence 2 letter b) TFEU, the restriction of a technical development is sufficient.[5]

Industry 4.0 is primarily about the networking, connection and digitisation of manufacturing processes, which paves the way for new production methods and products. This technical development is excluded if there is no access to a software or cloud service through which the company concerned can establish and develop an Industry 4.0 implementation. A platform is required for this, on the basis of which an individual adjustment to the operational conditions is possible to produce new products or to produce products that have already been developed more cheaply or in a way that is adapted to customer requirements. This also fulfils the further condition for an improper use of a dominant position because the **denial of access excludes subsequent developments** that are only possible with access to a certain operating system. This means that any competition in a neighbouring market that depends on the refused delivery is excluded (CJEU, Case T-201/04, ECLI:EU: T:2007:289 (paragraph 593)—Microsoft I).

During the latest GWB reform, this national prohibition of abuse was specifically extended to the **withholding of data required for economic development**. Under paragraph 19 Section 2 No. 4 GWB, abuse particularly occurs if a dominant company, as a supplier or customer of a certain type of goods or commercial services, refuses to supply another company with such goods or commercial services for a reasonable fee. This applies especially to the granting of access to data, networks or other infrastructure facilities, because the delivery or the granting of access is objectively necessary to be active on an upstream or downstream market, so that the refusal threatens to eliminate effective competition on this market, unless it is objectively justified. In this way, the judgments detailed below are codified and at the same time **data** is clearly defined as '**essential facility**'. Thus, paragraph 20 Section 1a GWB extends the prohibition to companies whose market power arises from the control of data, even if there is no obvious dominant market position. It is rather sufficient if a potential competitor is dependent on access to data, which is controlled by another company, for its economic activities. The refusal of access to such data in exchange for reasonable compensation can then constitute an unreasonable hindrance under paragraph 20 Section 1 in conjunction with paragraph 19 Section 2 No. 1 GWB, **even if business transactions for this data have not yet been opened** respectively when there is a first delivery by a third party.[6]

---

[5]ECJ, Case C-418/01, ECLI:EU:C:2004:257 (paragraph 44)—IMS Health; CJEU, Case T-201/04, ECLI:EU:T:2007:289 (paragraphs 647 et seq.)—Microsoft I.

[6]RegE BT-Drs. 12/23492, p. 81.

2.2 Justified Denial of Access?

Furthermore, an objective justification for a refusal to deliver is to be asked, which does not make the behaviour appear abusive. It is important to weigh the positive effects of a mandatory disclosure against the potential brakes on innovation (CJEU, Case T-167/08, ECLI:EU:T:2012:323 (paragraph 139)—Microsoft II). The resulting disadvantages for competition must go hand in hand with equally strong or higher efficiency advantages for the market and necessarily also for the consumer. Briefly, they must be necessary to achieve a market equilibrium. If they go beyond this, they are abusive (ECJ, Case C-95/04 P, ECLI:EU:C:2007:166 (paragraph 86)—British Airways).

Since the foundations for the development of Industry 4.0 practically form the basis of a necessary progress for the industry to survive in international competition and thus also questions competition in the affected industry, the disadvantages are dominant. Efficiency advantages for the consumer are not evident from a refusal to deliver, nor are there any indispensable aspects for the provider, which could consist of a necessary holistic design (see CJEU president, Case T-201/04 R 2, ECLI:EU:T:2004:372 (paragraph 44)—Microsoft I). However, it is relevant if an economic operator does not have the **necessary reliability**, for example not to pass on any data (Spindler 2018, p. 166).

Incidentally, it is clear from the start that the **software** or **access to a cloud** for Industry 4.0 is only a **basic element** that must be adapted to the individual needs of each corporate entity. This means that every company must have access and must be included into the problem-solving process, which excludes uniform and schematic models from the outset. This requires access and individual coordination. The platform provider, however, keeps his gains and thus it is only a matter of company-specific adaptation, not of losing one's own market position (this could lead to a justification for denying access, Spindler 2018, p. 165). It is necessary to pay appropriately for access so that it does not lead to economic losses.

2.3 Obligation to Pay Appropriate Remuneration

2.3.1 Measurement Based on the Innovative Character of the Service

If the providers with a dominant market position in software and access to a cloud for Industry 4.0 and the other corporate entities are obliged to cooperate, the former must make their offers available—but not free of charge. They are allowed to charge licence fees resp. user fees. However, this does not apply indefinitely. Otherwise, the owner of a dominant position would have the opportunity to use high rates of remuneration to discourage the companies dependent on access from using it. The owner of a dominant market position must therefore not be paid for the mere power of disposal over a product or a service on which other economic operators are indispensably dependent. On the other hand, rates are **appropriate** that are based on the **extent to which the service offered is new or innovative** in character and thus embodies progress that the owner can get paid for with corresponding remuneration. However, the required rates must also be common in other business transactions and thus for comparable technologies (CJEU, Case T-167/08, ECLI:EU:T:2012:323 (paragraphs 143 et seq.)—Microsoft II).

2.3.2 Prohibition of Discount Systems

Conversely, the required remuneration must not be too low. It is a recurring phenomenon that dominant companies want to consolidate their position through discount systems. By demanding low remuneration, they try to crowd out other competitors to be without significant competition afterwards and thus be able to increase prices. Such an opportunity exists above all at the beginning of a development that originally began with Industry 4.0. As long as there are few providers on the market or in some fields, it is particularly attractive to secure the exclusive position as a provider by initially attracting customers through low prices—eventually at the expense of the other providers.

This **displacement effect** results from the examination of all circumstances, especially the criteria and modalities used for granting discounts, the extent of the dominant position and the special competitive conditions relevant to the market. An indication of the likelihood of such an effect is when a discount system captures the majority of customers in the market (ECJ, Case C-23/14, ECLI:EU:C:2015:651, 1st headnote—Post Danmark/Bring Citymail Danmark). This would be the case if, for example, a provider of software for Industry 4.0 grants discounts for a whole branch. Conversely, however, because of the need to tailor Industry 4.0 to specific branches of the economy and thus industries, it is also possible to argue that the discounts are relevant for all customers in a specific industry. Therefore, the acquisition of a large number of customers can be due to the orientation towards Industry 4.0. This means that misuse cannot necessarily be deduced from the size and breadth of a discount system (generally Seitz 2015, p. 963, in her comments on the ECJ ruling on market size). However, the likelihood of an anti-competitive effect is sufficient without having to prove its severity or its appreciable nature at all (ECJ, Case C-23/14, ECLI:EU:C:2015:651, 3rd headnote—Post Danmark/Bring Citymail Danmark). Such proof would also be difficult, especially for Industry 4.0, as it is a strongly emerging and rapidly developing phenomenon. In doing so, it is not necessary to depend on whether there is an equally capable competitor. Such a criterion can and must be omitted if the market-dominating company has a very large market share and structural advantages. This is also imperative, as the structure of the market practically rules out the entry of an equally capable competitor (ECJ; Case C-23/14, ECLI:EU:C:2015:651 (paragraph 59)—Post Danmark/Bring Citymail Danmark). If there is still no competitor available in the area of Industry 4.0, the criterion of ‘equally efficient competitor’ is also not appropriate.

2.4 Injunctive Relief After the Huawei Judgment

But what if a software is used without paying the patentee the appropriate remuneration? Particularly with Industry 4.0, expensive know-how can be copied by collecting a lot of data. It is also possible to adopt standards that have been set by a standardisation organisation. As in the Huawei case ruled by the ECJ, this can then be accompanied by an **irrevocable commitment** to the organisation concerned to grant licences to third parties on predetermined **FRAND terms**. In this case, the patent holder can hardly deviate from granting a licence under these conditions if he does not want to act abusively (ECJ, Case C-170/13, ECLI:EU:C:2015:477 (paragraph 53)—Huawei Technologies). If he sues for an omission or a recall, the refusal to grant a licence can be held against him.

A lawsuit for **injunctive relief** or for recall can on its part be abusive because of the **compulsory licence objection**. This might be the case if the patent user has not been confronted and heard about the alleged infringement. Moreover, if he was not threatened with legal action (ECJ, Case C-170/13, ECLI:EU:C:2015:477 (paragraphs 61 f.)—Huawei Technologies). In addition, the patent user must then not have expressed his will to conclude a licence agreement on FRAND terms. And even then, the patent owner must first submit a specific written **licence offer** and, above all, state the licence fee, including how it is calculated (ECJ, Case C-170/13, ECLI:EU:C:2015:477 (paragraphs 63 f.)—Huawei Technologies). The company reliant on the patent must then react immediately, be it by accepting the offer under the FRAND conditions or immediately making a specific counteroffer which, in its view, corresponds to these conditions (ECJ, Case C-170/13, ECLI:EU:C:2015:477 (paragraph 67)—Huawei Technologies).

However, the assessment of whether the respective **conditions offered** are **appropriate** can diverge. The ECJ does not specify whose ideas should prevail, even when it comes to the question of adequate security, provided the patent is already in use. The fact that it is used before the contract is concluded speaks in favour of aligning with the ideas of the patent holder (Palzer 2015, p. 705). If, however, the needs of the patent using company are not taken as a basis, the patent holder can block and thus prevent the development of effective competition.

It is therefore problematic that the Federal Court of Justice even wanted to impose a specific, acceptable offer on the licence seeker (see BGH Kartellsenat, Case—KZR 39/06, BGHZ 180, 312—Orange-Book-Standard, WRP 2009, 858). The BGH generally assumed the admissibility of an injunction if the licence seeker had not made an unconditional offer to the patent holder or did not use the subject matter of the patent prior to the conclusion of the contract as this corresponds to the usual conditions according to the licence agreement to be concluded and the associated obligations (BGH Kartellsenat, Case—KZR 39/06, BGHZ 180, 312—Orange-Book-Standard, WRP 2009, p. 858).

However, this decision was already classified as contrary to Union law (de Bronnet 2009, p. 899, 902; Picht 2014, p. 1, 16; different contrary Wirtz 2011, pp. 1392, 1403 f.). In any case, this also means that a medium-sized company, which is likely to be less familiar with antitrust law, is obliged to act correctly in terms of procedures. This can result in considerable difficulties in using a patent. It also speaks in favour of the already decisive solution of the ECJ, which sees the patent holder in the procedural obligation to confront and hear the patent-using company about the alleged infringement.

Overall, however, questions remain open, such as the specific amount of the fee. The companies concerned are even more in demand (ct. Palzer 2015, p. 706). An individually adjusted, successful implementation of Industry 4.0 in a company will be only possible through a narrow and trusting cooperation and the support of a corresponding developer with the appropriate software or cloud.

3 Anti-Competitive Information Exchange

The exchange of information for the realisation of Industry 4.0 is sensitive. Even the deliberate exchange of information can lead to a contact between two companies, which trade with each other without any close contact. Therefore, the transparency in the market increases (see Heyers 2013, p. 99, 101 about OLG Düsseldorf, Case V-1 Kart 1-6/12 (OWi) inter alia (paragraphs 33, 37, 41, 44 et seq.)—Silostellgebühren I) and further coordination may arise. No formal agreements have to be made. A **coordinated behaviour** is already given with a conscious practical cooperation. This does not yet have to correspond to a specific plan to replace risky competition (ECJ, Case C-8/08, ECLI:EU:C:2009:343 (paragraph 26)—T-Mobile Netherlands; Case C-7/95 P, ECLI:EU:C:1998:256 (paragraph 86)—John Deere). Just **exchanging strategic data** can be enough. Mutual contacts can result from a unilateral disclosure of strategic data that the recipient accepts (CJEU, Case-25/95 inter alia, ECLI:EU:T:2000:77 (paragraph 1849)—Cimenteries CBR; Heyers 2013, p. 99, 102).[7] In any case, this means that a coordinated behaviour is quickly established.

The competitive relevance for entrepreneurial agreements, for example on research and development, which includes the exchange of made progress, is even more relevant. Direct or indirect influencing of competitors in their market behaviour or information about their own future demeanour is also sufficient (already ECJ, Case 40/73 inter alia, ECLI:EU:C:1975:174 (paragraphs 173 f.)—Suiker Unie). **Discussing confidential business information** is particularly questionable in terms of competition law in the context of frequent contacts (European Commission, Case K (2013) 8286, OJ 2014 C 453, p. 16—Nordseekrabben). This particularly concerns information about orders, prices, sales, investments and current business policy.[8] This can also include information about Industry 4.0 if it results in **future**

[7] See European Commission, notices from 14.1.2011 Guidelines on the applicability of Article 101 of the Treaty on the Functioning of the European Union to horizontal co-operation agreements, OJ 2011 C 11, p. 1, paragraph 61.

[8] See European Commission, Case KOME 92/157/EWG, OJ 1992 L 68, p. 19, paragraphs 17 ff.

investments and **the direction of the company**. This is to be distinguished from the purely technical solution of problems and the professional exchange about arising difficulties. Problems of a competition law nature only arise when market-relevant company data is exchanged, which the company would otherwise not be able to obtain. This applies above all to confidential company data.[9]

In contrast, public data, which can be obtained easily and cheaply, usually does not have an impact on competition (CJEU, Case T-191/98 inter alia, ECLI:EU: T:2003:245—Atlantic Container Line) unless such an exchange serves to operate or maintain a cartel and is therefore based on cooperation (OLG Düsseldorf, Case V-1 Kart 1-6/12 (OWi) inter alia (paragraphs 33, 37, 41, 44 et seq.)—Silostellgebühren I). In such a case, even publicly available information can be the subject of an antitrust violation (Heyers 2013, pp. 99, 102). However, this is more of an exception. The decisive factor is whether it is information that eliminates the uncertainty that companies face about their business and pricing policy, which is typical for competition. If this does not follow from the nature of the information, other cooperative or anti-competitive elements must be added.

The question of anti-competitive factors also depends on the structure of the **market**. If it is very **differentiated**, even exchanged sensitive information and thus business secrets cannot be felt, provided that the anti-competitive effects are completely lost (Frenz 2014, pp. 193, 199 f.). An exchange of information can even stimulate lively competition (CJEU, Case T-35/92, ECLI:EU:T:1994:259 (paragraph 51)—John Deere).

4 Pro-Competitive Behaviour With Definitional Exclusion

In many cases, it is only cooperation between companies that enables Industry 4.0 to be implemented effectively. The vertical or horizontal cooperation then enables new products to come about and to survive in competition. Small companies are particularly often not in a position to create new products completely independently using Industry 4.0. Therefore, **cooperation** is required. This is **necessary** to create **effective competition**, thus, also to enable small and medium-sized companies to participate in developments in the context of Industry 4.0. In such cases of competitive necessity, competition is not impaired, but only created. Accordingly, such behaviour is not a violation of the ban on cartels (already ECJ 56/65, ECLI:EU:C:1966: 38—Maschinenbau Ulm; also Case 258/78, ECLI:EU:C:1982:211 (paragraphs 56 et seq.)—Nungesser). This can even be the case with large, dominating companies cooperating, if they only come up with a new solution for the market by combining their specific orientations.

[9] See European Commission, Case KOME 1999/60/EG, OJ 1999 L 24, p. 1 (paragraph 138)—Fernwärmetechnik-Kartell; Roth and Ackermann (2018), Art. 81 paragraph 1 EG—Grundfragen, paragraph 117.

5 Exemption

If the ban on cartels does apply nonetheless, considerable efficiency gains can occur through technical or economic progress, so that an exemption in line with Article 101 Paragraph 3 TFEU can be considered. With Industry 4.0 the necessary advantages for the consumer result from new products and cost savings. Cooperation is indispensable if the companies concerned are not in a position to achieve the desired innovations in the context of Industry 4.0 on their own.

The **cooperation** however, must be limited in such a way that it **only relates to technical development** and specifically not to later production, so that there are still different products in competition and this form of competition is not impaired. Thus, part of the way is walked together, but afterwards acted separately again on the market.

These principles have been specified in Article 4 Paragraph 1 Regulation (EU) No. 1217/2010.[10] This limits customer and sales restrictions to seven years, with the market share not exceeding 25 %. In any case, there is a further extension according to Art. 4 Paragraph 3 for this period. This means that an exclusive licence can exist for the inventor of a technical development, which excludes other providers from the market, until the development costs have been amortised and adequate utilisation has taken place. With this approach, collaborations in research and development can be exempted. For consistency and in the interest of the security of the participants, an explicit standardisation for typical 4.0 platforms would be very desirable (Spindler 2018, p. 166). The exchange of information may also be part of this if the experience of several companies is necessary to get a sufficient picture of how to proceed further with regard to Industry 4.0. Therefore, the **exchange of information can exceptionally be exempt**.[11]

6 Conclusion

Industry 4.0 is an innovative process that requires the greatest possible openness. It is a particular challenge for all companies that can hardly ignore this process if they want to remain competitive. It is even more important to ensure that they have access to technologies and devices to implement Industry 4.0 in their business. This assumes that the providers of the corresponding software and cloud services guarantee this access. If there are difficulties with dominant companies, the prohibition of

[10] Regulation (EU) No. 1217/2010 by the commission from 14th December 2010 on the application of Article 101(3) of the Treaty on the Functioning of the European Union to certain categories of research and development agreements OJ 2010 L 335, p. 36.

[11] See European Commission, Guidelines on the applicability of Article 101 of the Treaty on the Functioning of the European Union to horizontal co-operation agreements, OJ 2011 C 11, p. 1, paragraph 74.

abuse in line with Art. 102 TFEU can help. This also prohibits an abuse of conditions, for example through an acceptance for further data processing required in the terms of use.

However, an appropriate remuneration can be charged for access to software or cloud services. If companies simply use protected patents without remunerating or the patent holder's consent, according to the ECJ, even an action for an injunction against it can be abusive if there is no informative advance warning or if the user timely offers an appropriate remuneration.

Cooperation between companies will be indispensable and can be exempted with regard to the technological and economic achieved progress by this, if the cooperation is limited to the indispensable areas, is temporary and does not result in uniform products, so that competition is largely eliminated. The exchange of information is also elementary in this area, often related to specific orders and therefore tends to be more permitted than according to the general standards of the cartel ban. However, information about future business policy and prices and strategies must not be passed on to the competition in the context of Industry 4.0.

Because of Industry 4.0, many competition law problems arise but they can be managed well within the framework of the applicable competition rules. There is therefore no need for Antitrust Law 4.0, but rather the application of competition rules in relation to Industry 4.0. Together with a definition of platform standards, these are considered to be more flexible than an assignment of data based on ownership (Spindler 2018, p. 170). Certainly, competition law cannot deal with all constellations. However, a property assignment is possible in every situation. The overwhelming opinion recommends recognising a right to one's own database (see action recommendations by Hornung/Hofmann, Hornung 2018, p. 216). There is also flexibility in this regard, as the *sampling* decision of the Federal Constitutional Court and now the ECJ (Case C-476/17, ECLI:EU:C:2019:624) shows from the point of view of artistic freedom (see contribution 6 by Frenz, Who Owns the Data?).

References

de Bronnet G-K (2009) Gemeinschaftsrechtliche Anmerkungen zum "Orange-Book-Standard"-Urteil des BGH, WuW 2009: 899 et seq

Frenz W (2014) Datenschutz durch Kartellrecht. EWS 2014:193 et seq

Frenz W (2015) Das EU-Missbrauchsverbot. JURA 2015:1206 et seq

Frenz W (2016a) Handbook of EU Competition Law, Berlin

Frenz W (2016b) Industrie 4.0 und Wettbewerbsrecht. WRP 2016:671 et seq

Heyers J (2013) Systemkonformität der kartellrechtlichen Beurteilung sog. Informationsaustauschs. NZKart 2013:99 et seq

Hornung G, Hofmann K (2018) Industrie 4.0 und das Recht. Drei zentrale Herausforderungen. In: Hornung G (ed) Rechtsfragen der Industrie 4.0, Datenhoheit – Verantwortlichkeit – rechtliche Grenzen der Vernetzung, Baden-Baden, 9 et seq

Palzer C (2015) Patentrechtsdurchsetzung als Machtmissbrauch - der Zwangslizenzeinwand aus unionsrechtlicher Sicht. EuZW 2015:702 et seq

Picht P (2014) Standardsetzung und Patentmissbrauch – Schlagkraft und Entwicklungsbedarf des europäischen Kartellrechts. GRUR Int. 2014:1 et seq

Roth W-H, Ackermann T (2018) In: Jaeger W, Kokott J, Pohlmann P, Schroeder D, Frankfurter Kommentar Kartellrecht, Effective: 5/2018, Otto Schmidt, Köln, art. 81 paragraph 1 EG

Schröter, H/Bartl, U (2015) In: von der Groeben/Schwarze/Hatje, Europäisches Unionsrecht, 7th edn. Nomos, Baden-Baden, Art. 102 TFEU

Seitz C (2015) Anmerkung zu einem Urteil des EuGH vom 06.10.2015 (C-23/14) - Zur Frage des Missbrauches einer marktbeherrschenden Stellung durch Rabattsysteme, EuZW 2015:959 et seq

Spindler G (2018) Plattformen und Plattformregulierungen als Alternative zu Dateneigentumsrechten. In: Hornung G (ed) Rechtsfragen der Industrie 4.0, Datenhoheit – Verantwortlichkeit – rechtliche Grenzen der Vernetzung, Baden-Baden, p 151 et seq

Weidenbach G, Vogt B, Hauser SM (2012) Zugang zu Produktplattformen im europäischen und deutschen Kartellrecht. WRP 2012:66 et seq

Wirtz M (2011) Der kartellrechtliche Zwangslizenzeinwand nach der Orange-Book-Standard-Entscheidung des BGH, WRP 2011:1392 et seq

Standard Setting by Associations

Walter Frenz

Contents

1 Essential Open Concept

Not only companies but also **corporate associations** are legally obliged by competition law, according to the ECJ decision of 14 November 2017 (ruling C-671/15, ECLI:EU:C:2017:860 – APVE et al.) regarding the agricultural sector. When it comes to **regulations concerning technical developments,** problems arise regularly in the sense that large corporations come together in working groups and force

The elaborate summary is a written version of the discourse held by the author at the Deutscher Anwaltstag in Mannheim on 8 June 2018, published in NJOZ (2018), p. 1321. See also energy related articles in N&R (2017), p. 258 and N&R (2018), p. 139.

W. Frenz (✉)
Unit for Mining, Environmental and European Law, RWTH Aachen University, Aachen, Germany
e-mail: frenz@bur.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_15

their rules through, while smaller companies have great difficulties in getting their ideas to take effect. Business organisations are also not allowed to cooperate indefinitely with one another. Agreements on prices and quantities made between several agricultural producer organisations or associations can, according to the ECJ, constitute a cartel within the meaning of competition law. If such agreements are made within the same organisation, this is, at best, in conformity with the competition if they serve the goals with which the organisation or association is entrusted and if they are proportionate. However, this concerned the special objectives of agricultural policy, for the implementation of which associations can be duly recognised by the state.

The energy industry, for example, is shaped by numerous regulations, which are harmonised under Union law. However, it is not 'communitised' like agricultural policy, not even under the workings of an energy union (see Frenz 2015b, p. 481). Therefore, the energy industry cannot legitimise the exchange of strategic information, agreements on the generated electricity quantities placed on the market or the coordination of the price policy of the individual producers. According to the ECJ (C-671/15, ECLI:EU:C:2017:860 – APVE et al.), this is also not possible between different producer organisations but only within their internal realm and thus within its framework between the members.

Because of this, corporate associations must also exercise caution not to define standards that only favour certain companies and put others at a disadvantage or completely exclude them from competing.[1] In addition, such standards must also be open to competitors from other EU countries so that **no mere national platforms** arise. If this were the case, market access across borders would be more difficult or even impossible. Two corporate platforms want to set global standards in the field of Industry 4.0: The Industry 4.0 platform, located in Germany, focuses on social aspects of digitised production and legal issues, such as ownership of data, while the American platform 'Industrial Internet Consortium IIC' focuses on communication and thus the interface problem. Some companies are members of both platforms, which fans fears of smaller companies that they will not be able to survive in the face of the emerging Internet giants (FAZ from 2 March 2016, No. 52 p. 18). This shows the need for global openness and accessibility so that all companies can participate in Industry 4.0 and are not practically excluded by set standards that are only tailored to the needs and wants of certain corporate entities. This also applies to area-related standard setting.

[1] Factual measures can also be recorded, ECJ, ruling C-209/78 et al., ECLI:EU:C:1980:248, Paragraph 88 – van Landewyck; on corporate agreements, e.g. on institutional standardisation, differentiating from Jakobs (2012), p. 58.

2 Competitive Claim for Consideration of New Developments

If a company's sales are dependent on their **product being included in usual standards**, the question arises whether there is a claim for inclusion. This applies especially to innovative developments that deviate from the established standards or require new rules so that they can be offered on the market with legal certainty. They must then be represented by name in the DIN standards to receive the required approval for marketability.

The prohibition of cartels in line with Article 101 TFEU also includes decisions by business associations, which often concern the setting of standards. This is the case, for example, with **branch-related standardisation by private associations**, to which this task has been assigned by the state. The latter must, however, ensure that it is filled in compliance with competition (ECJ, ruling C-184/13 et al., ECLI: EU:C:2014:2147 – API. Described in more detail under Sect. 5). These industry-specific standards bridge the gap to the legal sources in the form of the generally recognised rules of technology. They must be scientifically correct in theory, reflect the majority opinion of technical experts and have proven themselves in practice. The Federal Constitutional Court also focuses on the application. The authorities and courts can limit themselves to ascertaining the prevailing opinion among the technical practitioners to determine whether the respective technical equipment may or may not be brought into circulation (BVerfGE 49, 89, 135 f. – Kalkar I; already Breuer 1976, p. 67). **DIN standards** carry a rebuttable presumption that they correspond to the recognised rules of technology (BGH, ruling of May 24th, 2013 – V ZR 182/12, Paragraph 25; Seibel 2013, p. 3001). This also shows the elementary importance for placing products on the market. DIN pre-standards according to the rules of procedure of DIN with only a provisional character are not yet recognised.

To be measured against competition law, decisions by business associations must also be capable of affecting trade between Member States and of preventing, restricting or distorting competition within the internal EU market. That already happens with a **market foreclosure**. Setting standards tailored to domestic companies, for example, can affect companies from other EU countries in their entry into the market.

For such a market foreclosure, mere speculative considerations are not sufficient (ECJ, ruling 22/78, ECLI:EU:C:1979:138, Paragraph 18 et seq. – Hugin). In any case, it must be possible to foresee indirect and potential influences with sufficient probability based on the overall objective circumstances (ECJ, ruling 31/80, ECLI: EU:C:1980:289, Paragraph 18 – L'Oréal). In particular, these have not occurred yet. The level of the requirements to be made must not be excessive, so that planned market entries of new economic participants who only need to be potential competitors remain possible (more in Communication from the Commission – Guidelines on the applicability of Article 101 of the Treaty on the Functioning of the European Union to horizontal cooperation agreements, OJ 2011 C 11, p. 1, Paragraph 10;

Schröter and Voet van Vormizeele 2014, Art. 101 TFEU Paragraph 81; Frenz 2015a, Paragraph 1095, 1106).

Accordingly, **relevance for the domestic market is given due to potential transnational impacts** (see the generous assessment by the ECJ, C-518/13, ECLI: EU:C:2015:9 – Eventech). Apart from that, § 1 GWB (German Act against Restraints of Competition) contains the same prohibition as Article 101 Section 1 TFEU, which applies regardless of cross-border effects.

3 Competition Law Design

In terms of competition law, it is initially problematic to define standards that then define rules for the admissibility and handling of technical developments. Even here, contrary to Article 101 TFEU, **agreements** can already be reached (KOME 93/174/ EEC, OJ 1993 L 73, p. 38, Paragraph 20 et seq. – regarding tariff structures in combined freight transport). These concern, for example, the question of which standards should be used between different companies with parallel market interests. The problem is that the standard-setting actors sometimes have different interests than the other companies and therefore tend to exclude competitors by their cooperation.

Resolutions by corporate associations are equivalent in line with Article 101 TFEU. Particularly in this context, dominant actors can assert themselves or coordinate with other corporate representatives. Another problem is the adoption of complicated rules and standards that small and medium-sized companies (hereinafter SMCs) are unable to meet and therefore keep them away from the market (Spindler 2018, p. 164; see already Frenz 2016b, p. 673 et seq., p. 678) – for example, as a result of the necessary cost-intensive certifications. Consequently, the development of other companies in competition can be significantly hindered, if not excluded at all, without large standard-setting companies having to appear themselves externally. This shows the function of preventing circumvention (Stockenhuber 2017, Article 101 TFEU Paragraph 86).

As far as agreements between companies are concerned, resolutions by corporate associations are also to be made to avoid gaps in protection caused by relocating entrepreneurial behaviour to associations (Emmerich 2012, Article 101 Section TFEU Paragraph 74 with further references). They are traditionally defined by the Commission as any **expression of will** provided for in the **articles of association** and in accordance with the provisions laid down there (KOME 85/75/ EEC, OJ 1985 L 35, p. 20, Paragraph 23 – Fire insurance). Against this background, even non-formal declarations by business associations, even **association recommendations**, can be viewed as resolutions within the meaning of Article

101 Section 1 TFEU.[2] Such recommendations are technical specifications drawn up by associations or standardisation organisations (see Jakobs 2012, p. 58).

For this, however, it is important that the companies feel bound by the requirements of the corporate associations, regardless of whether the resolution itself is binding. In accordance with the broad conception of preventive and general consent in agreements (see for example ECJ, ruling C-2/01 P et al., ECLI:EU:C:2004:2, Paragraph 141 – Bundesverband der Arzneimittel Importeure eV; C-74/04 P, ECLI:EU:C:2006:460, Paragraph 46 – Volkswagen), this can also be done by the fact that the membership statutes provide for association resolutions that are mandatory. However, the sectors and objectives must be named to be able to delimit the affected area in a sufficiently concrete manner.

In addition, **factual measures** are also recorded because targeting the consequences combated by Article 101 TFEU is decisive (with an explicit example, ECJ, 209 et al. / 78, ECLI:EU:C:1980:248, Paragraph 88 – van Landewyck). This means that it is also relevant if a technical specification is not included, even if it has not been formally agreed. In fact, such an approach is usually based on a resolution, be it unilaterally by an acting body or factually by simply not voting on a motion. That is sufficient (Schröter and Voet van Vormizeele 2014, Article 101 TFEU Paragraph 50 with reference to EGC, ruling T-25/95 et al., ECLI:EU:T:2000:77, Paragraph 928 – Cimenteries CBR et al.), as otherwise the official form of resolution could be deliberately circumvented to avoid the intervention of the cartel ban.

4 Consequences Under Competition Law

4.1 Nullity Under Article 101 Section 2 TFEU

The resolutions of the committee of a corporate association for technical regulation violate Article 101 TFEU if the proposals of unconsidered companies are left out of the picture (except for technical reasons and circumstances) and it is thus no longer possible or significantly more difficult to bring their products to the market. This creates a **closed shop** that also blocks potential competitors from other EU countries from entering the market and therefore has cross-border effects.

It is true that a company cannot force an association of companies to behave in a certain way based on Article 101 TFEU. A decision based on the facts and circumstances outlined above is, however, nugatory under Article 101 Section 2 TFEU if this **affects a company's competitiveness for no objective reasons** (especially technical reasons for excluding the product concerned). This does not require a resolution (already ECJ, 48/72, ECLI:EU:C:1973:11, Paragraphs 25, 26 – Haecht

[2] See for example KOME 89/512/EEC, OJ 1989 L 253, p. 1, Paragraph 46 – Dutch banks; 90/25/EEC, OJ 1990 L 15, p. 25, Paragraph 16 et seq. – Concordato Incendio; 93/3/EEC, OJ 1993 L 4, p. 26, Paragraph 16 et seq. – Lloyd's Underwriters.

II), but this legal consequence takes effect immediately. It is pronounced by the national courts, but their decision is not constitutive: it only establishes nullity, and it is mandatory. Due to the direct application of the EU cartel ban, the courts have no scope of interpretation but strictly must enforce Union law.

If the **consequence of nullity** is given, it is **absolute**. The anti-competitive decision must therefore be treated as if it had not come about. It has no legally binding effect, neither between the parties involved nor to third parties (Frenz 2015a, Paragraph 1775). An appeal to this is therefore not possible with regard to authorities and other bodies. The measures concerned are completely ineffective. Everyone can invoke this legal consequence (ECJ, ruling 22/71, ECLI:EU:C:1971:113, Paragraph 29 – Béguelin).

If the standard-setting of the corporate association or its subdivision is nugatory, **certification** may not be refused because the corresponding product was not included. Rather, the responsible body must then ignore this omission and continue with the system or product certification based on the previous unit certificate. Otherwise, it is behaving illegally.

4.2 Claim to Inclusion

To avoid the automatically occurring consequences of nullity, the association or its subdivisions are **in fact** forced to pass their resolution on the setting of standards in such a way that the interests of all companies are adequately safeguarded. This orientation is also supported by the fact that all economic operators depend on the standards being set in such a way that they can sell their products. In this respect, there is a dependency that has resulted in access claims to platforms within the framework of Article 102 TFEU.[3] In parallel, claims to be considered in standards as the basis for certification and thus in fact also for market access are to be advocated.

A mere formal participation by reserving a seat in a standardisation committee for a representative of small and medium-sized companies is not enough. In this way, their interests might be brought in securely at a first glance, but not necessarily considered in such a way that the competitive interests of **SMC**s are preserved. If the majority of the votes alone counts, dominant actors can easily prevail. Therefore, a material and not a purely formal approach is required. This approach depends less on the procedure than on the content that has been approved. It is important to preserve the effect of the competition. Observing the principle of democracy is not enough for this either. However, the procedure and thus the **representation of the SMCs in a standardisation committee** can guarantee that the relevant issues are brought up to

[3]Regarding intellectual property cases ECJ, T-201/04, ECLI:EU:T:2007:289, Paragraph 334 – Microsoft I as well as Case T-167/08, ECLI:EU:T:2012:323, Paragraph 139 – Microsoft II and in this the citated ECJ ruling Case C-418/01, ECLI:EU:C:2004:257, Paragraph 28 et seq. – IMS Health; See more in the previous Paragraph 26 by Frenz.

date. Ultimately, however, they must assert themselves in terms of content insofar as this is necessary to maintain effective competition.

4.3 Defence, Injunctive Relief and Claims for Damages

Resolutions of an association to set rules that do not consider the submissions of companies in an anti-competitive manner are already nugatory. If other economic operators do not adhere to it, for example by trying to reject the affected products with reference to the setting of standards, the question of **injunctive relief** arises. These are not regulated by Union law and therefore result only from national law. However, national law must always be seen in connection with EU law, such as § 823 Section 2 BGB (German Civil Code) in conjunction with Article 101 TFEU. However, all the prerequisites of national law must then also be met, for example fault under § 823 Section 2 sentence 2 BGB, even if the violated law such as Article 101 TFEU does not contain such a requirement. This applies specially to claims for damages. § 33 Section 3 sentence 1 GWB also requires an intentional or negligent inspection.

In the case of **preventive injunctive relief**, traditionally only an imminent danger of an objectively unlawful encroachment on a legal asset protected by §§ 823 et seq. BGB is required. It does also not depend on the fault or the awareness of the illegality (Palandt 2019, introduction to § 823 Paragraph 27 et seq. Also on the following). Under § 33 Section 1 sentence 2 GWB, a claim for injunctive relief already exists if there is a threat of a violation. This is the case in the event of a violation of Article 101 et seq. TFEU. As a preventive measure, an action for injunctive relief is generally possible if someone asserts the right to be allowed to carry out the infringing act or has already made the decision to violate so that it is entirely up to the regarding person (BGHZ 117, 264, 271). Accordingly, it comes into consideration if the competent authority claims that it will no longer certify the products in question, although they have not been considered in an anti-competitive manner in a technical set of rules of a corporate association. It would perpetuate this violation of Article 101 TFEU and thus possibly violate it itself. However, the company concerned must then prove that the non-consideration is anti-competitive. There must also be no comprehensible technical reasons that legitimise the exclusion of the affected product. However, the assertion of a continuing certification claim will correspond to the need for legal protection.

Enforcing an injunction against the corporate associations would prevent the product in question from being excluded from the outset. The technical standards would be far enough not to jeopardise certification. Such a risk would arise primarily from the publication of technical standard-setting based on anti-competitive decisions. Therefore, a claim against the corporate associations **not to accept the technical rules in the intended form** comes into consideration. The prerequisite, however, is that the company concerned can prove that such a decision will impair competition. The decision must not be the result of technical reasons. Apart from that, however, the corporate association has a secondary burden of disclosure as for large companies under the Product Safety Directive (Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety, OJ 2002 L 11, p. 4) when it comes to product approval.

There are considerable additions to **claims for damages** through Directive 2014/104/EU.[4] Under Article 3 of this directive, the Member States must ensure full compensation. This includes all damage caused by the anti-competitive act (see Frenz 2016a, Paragraph 1790 et seq.). High claims for damages cannot be ruled out because, according to the will of the European legislator, these are intended to ensure the enforcement of antitrust law and thus deter violations. These threatened claims for damages may induce corporate associations not to violate the ban on cartels when setting standards or to leave them unconsidered in the following.

5 Involvement of State Institutions

Furthermore, **claims under competition law against state institutions**, which approve or even make binding anti-competitive standard-setting by corporate associations, come into consideration. **Generally recognised rules of technology** must be observed. This must also be checked by state institutions. Here, too, care must be taken to ensure that both neutrality and general comprehensibility as well as the resulting accessibility are maintained. There must be no protection standards that are almost impossible to meet or can only be met by certain companies. As a result, the rules must not have any prohibitive effect on competition. In the area of the prohibition of abuse, therefore, a high degree of structural responsibility can be assumed (Frenz 2016b, p. 678; Frenz 2015c, p. 428).

In some cases, EU law must be drawn up and its implementation must be ensured. Regulation (EU) 2016/631 (Commission Regulation (EU) 2016/631 of 14 April 2016 establishing a network code on requirements for grid connection of generators, OJ 2016 L 112, p. 1) defines a network code with network connection provisions for electricity generators. This is relevant for the developments in the field of **Energy Industry 4.0** on power generation itself or for feeding it into the grids. Under recital 2 of the regulation, the Member States or their regulatory authorities must ensure in line with Article 5 of Directive 2009/72/EC (of the European Parliament and of the Council of 13 July 2009 concerning common rules for the internal market in electricity and repealing Directive 2003/54/EC, OJ 2009 L 211, p. 55 – No longer in force, Date of end of validity: 31/12/2020; Repealed by Directive (EU) 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27/EU), among other

[4] Directive 2014/104/EU of the European Parliament and of the Council of 26 November 2014 on certain rules governing actions for damages under national law for infringements of the competition law provisions of the Member States and of the European Union, OJ 2014 L 349, p. 1.

things, that objective and non-discriminatory technical regulations with minimum requirements for design and operation are drawn up for the grid connection. If requirements constitute conditions for the connection to national networks, under Article 37 Section 6 of the aforementioned directive, the regulatory authorities are responsible for at least determining or approving the methods for calculating or defining these requirements if they have been developed by private standardisation bodies.

In the API case ruled by the ECJ (ruling C-184/13 et al., ECLI:EU:C:2014:2147 – API; more on this and also for the following Frenz 2015c, p. 421 et seq.) it was also about **setting standards**, albeit related to minimum operating costs in commercial road freight transport. These were determined by representatives of the association and announced as mandatory by a state authority, but they did not adequately protect the interests of competitors. The ECJ has, forasmuch, consistently ruled on state measures that form the background for competition violations by companies[5] and advocates a violation in line with Article 101, 102 TFEU via Article 4 Section 3 TEU if EU countries prescribe or facilitate cartel agreements contrary to Art. 101 TFEU or intensify their effects, for example by adopting or **making them mandatory** as a state measure (ECJ, ruling C-311/85, ECLI:EU:C:1987:418, Paragraph 9 et seq., 22 et seq. – Vlaamse Reisbureaus). It is also possible that the state removes the state character of its own regulation (virtually the other way around) and **transfers responsibility** for decisions affecting the economy to private economic operators (ECJ, riling C-198/01, ECLI:EU:C:2003:430, Paragraph 46 – CIF with further references). These can be, for example, **associations** or their sub-committees. However, this must actually be done (ECJ, ruling C-35/99, ECLI:EU:C:2002:97, Paragraph 43 – Arduino), so that the state activity itself only increases and is no longer constitutive. The two groups of cases mentioned above are therefore the main categories.[6] Art. 102 TFEU is affected if the state induces a company to behave in a way that abuses a dominant position (ECJ, ruling C-13/77, ECLI:EU:C:1977:185, Paragraph 30, 35 – Inno/ATAB).

The generally recognised rules of technology are binding for all economic operators in business dealings and set the relevant standards (see Sect. 1 above). This makes them look like they have been set by the state. The latter leaves the standardisation to the industry-specific representatives within the framework of DIN etc., so that the second case group is present. If these are the rules to be drawn up by the Member States according to Regulation (EU) 2016/631, they act as if the Member State had drawn them up itself. Ultimately, it is up to the Member States how they organise the development of the rules that must be adopted as such. In addition, under Article 37 Paragraph 6 of the Directive 2009/72/EC, the definition of

---

[5]ECJ, ruling C-184/13 and others, ECLI:EU:C:2014:214, Paragraph 28 - API; already ECJ, ruling 13/77, ECLI:EU:C:1977:185, Paragraph 30 and 35 – Inno/ATAB; for example, also ECJ, ruling C-94 and 202/04, ECLI:EU:C:2006:758, Paragraph 46 – Cipolla.

[6]ECJ, ruling C-184/13 et al., ECLI:EU:C:2014:214, Paragraph 29 – API; Rose and Bailey (2013), Paragraph 11.031; with three categories GA Léger, Opinion v. 07/10/2001 – C-35/99, ECLI:EU:C:2002:97, Paragraph 37 – Arduino.

these requirements must be approved by the state (namely by the regulatory authority) in good time before they come into force.

If the state has left an area to private-sector standardisation and association work in this way, it must **ensure compliance with the competition rules** and act accordingly on the corporate associations that set technical standards so that the interests of all companies are considered in a way that does not suggest anti-competitive effects – or ignore the rules that have been worked out in this way. In any case, they cannot be approved.

6 Prevention of Competition Through Close EU Standardisation

6.1 Effects

The extent to which national law comes into play depends on the intensity of the regulation in the EU regulations. When applying Regulation (EU) 2016/631, the technical guidelines of the recently developed EN 50549 on the requirements for generating plants that are operated in parallel with a distribution network must be observed. This determines the material standards at EU level. At national level, on the other hand, procedures continue to be determined, as evidenced by the continuing procedural autonomy of the Member States under Article 197 TFEU: The procedural modalities are based on national law, unless there is a relevant EU regulation (ECJ, ruling C-392 and 422/04, ECLI:EU:C:2006:586, Paragraph 57 – i 21 Germany and Arcor).

This includes measurement and testing regulations, while the EU level defines the certification requirements. Against the background of the national procedural design, these must be both enforceable and achievable so that the EU requirements can be handled in accordance with the internal market and in accordance with the fundamental liberties. The Member States are therefore responsible for success. Any inadequate or incorrect implementation of Union law constitutes a breach of contract and can trigger state liability – provided the legal situation is clear and therefore a sufficiently qualified legal violation has occurred (see BGH, ruling of January 22, 2009 - III ZR 233/07, NVwZ 2009, p. 795; Frenz 2013, p. 582).

Regulations under Union law, be it enactments, network codes or EN standards, leave the **Member States latitudes**, which are then filled by national company associations. In this process, there is a risk that large companies will prevail, because they tend to be more represented in these associations. Only these are personally able to support all the drafting of standards. If, on the other hand, Union bodies meet regulations that leave the national committees for regulation with a potentially anti-competitive orientation no room for configuration, anti-competitive behaviour is avoided from the outset. That speaks in favour of the most **comprehensive Union**

regulation possible for the effective realisation of competition. This is then an expression of the **effet utile** in relation to **EU competition law**.

6.2 The Union Organs As the Primary Guardian of Competition

Meanwhile, competition law is directed at companies. It is true that Article 106 Section 1 TFEU in conjunction with Article 4 Section 3 TEU include the Member States. However, the standardisation itself is carried out by the Union bodies so that these must also be bound. This applies to all Union bodies and thus not only to the European Commission but also to the **European Committee for Electrotechnical Standardization (CENELEC)** as an association of national regulatory authorities – insofar as it is responsible for the adoption of technical regulations and EU standards at Union level. Otherwise, the regulation could be partially exempted from EU law.

The European Court of Justice (ECJ) has repeatedly affirmed the commitment to fundamental liberties when issuing directives and regulations.[7] This is particularly true in the field of **legal harmonisation**. These measures must improve the conditions for the establishment and functioning of the internal EU market. Under Article 26 Section 2 TFEU, the latter is essentially constituted by the fundamental liberties (Frenz 2012, Paragraph 22 et seq., 335 et seq. Also for the following). The ECJ must assess whether this purpose is actually fulfilled (ECJ, ruling C-376/98, ECLI:EU:C:2000:544, Paragraph 84 – Germany v Parliament and Council – Advertising and sponsorship of tobacco products related to the free movement of goods and services; ECJ, ruling C-465/00, ECLI:EU:C:2003:294, Paragraph 41 – ORF).

The freedom of competition is not mentioned in Article 26 Section 2 TFEU. However, the competition rules themselves refer to the internal market by examining the compatibility of conduct with it. This is done by the European Commission. It would be therefore a **venire contra factum proprium** if the Commission would enforce competition rules against companies and also the states or their subdivisions including the standardisation committees (see in this respect ECJ, ruling C-184/13 et al., ECLI:EU:C:2014:2147 – API) but would not stick to it itself (Ehlers 2001, p. 274).

The central target provision of Article 3 Paragraph 3 Sentence 1 TEU names a highly **competitive market economy** after the establishment of the internal market by the Union (see Article 3 Paragraph 3 Sentence 2 TEU). Protocol declaration No. 27 also refers to Article 3 TEU. This still requires a system that protects competition from distortion and postulates that the Union should take action. The

[7] ECJ, ruling C-51/93, ECLI:EU:C:1994:312, recital 11 – Meyhui; already ECJ, ruling 41/84, ECLI:EU:C:1986:1, Paragraph 24 – Pinna; ECJ, ruling 20/85, ECLI:EU:C:1988:283 Paragraph 17 – Roviello.

most effective way of protecting competition is to use a standardisation system that prevents the distortion of competition.

This means that the Union must comply with and promote all competition rules. This applies to all bodies and organisations that issue EU standards, even if there are only technical regulations concerned that are also made mandatory. Accordingly, all these organs and organisations at EU level must pursue legal harmonisation so intensively that compliance with the competition rules is guaranteed. The Union continues to aim for an open, free and undistorted competition regime (Behrens 2008, p. 193; Dietrich 2012, p. 41; Nowak 2009, pp. 184, 190).

Internal market-related legal approximation intervenes, among other things, if the conditions of competition on the internal market are distorted and thereby distorted by existing differences in the legal and administrative provisions of the Member States (Article 116 TFEU; see Frenz 2015a, Paragraph 45). This approach shows the connection between the internal EU market and competition and, furthermore, that the approximation of laws should neutralise any distortions of competition that arise due to divergent state regulation. Article 116 TFEU presupposes an ongoing monitoring process with a primacy of consultations with the Member States. If this is not successful, however, the Commission must take action – with the aim of eliminating the distortions and occurring distortions of competition.

6.3 Extension Towards a Structure of Equal Opportunities

In the meantime, the ECJ attaches particular importance to maintaining **equal opportunities**. According to the ECJ ruling 'Dimosia', it may not be postponed by state measures or even just recognition. A state-maintained, reinforced market-dominant position, which distributes the opportunities in the market unevenly, is sufficient (ECJ, ruling C-553/12 P, ECLI:EU:C:2014:2083, Paragraph 47 – Dimosia. In explicit deviation from the CFI, ruling T-169/02, ECLI:EU:T:2005:46, Paragraph 105, 118 – Cerveceria Modelo). It does not even have to be shown what concrete entrepreneurial behaviour can result from the framework conditions created by the state. The **abstract existence of threats from unequal treatment of competitors** is sufficient (ECJ, ruling C-553/12 P, ECLI:EU:C:2014:2083, Paragraph 47 – Dimosia, according to Triantafyllou 2014, p. 737). Article 106 Paragraph 1 TFEU thus constitutes an abstract endangerment fact (more in Frenz 2014, p. 1455 f.).

In the area of standardisation, such a dominant position is held by the corporate associations and committees that develop technical rules or elaborate requirements under Union law. As their standards are recognised by the state and are binding for certification, they experience state reinforcement, which further consolidates their existing market positions. This is done by securing market opportunities in favour of the products of the large companies that are strongly represented in the standard-setting committees. In the API ruling, the ECJ therefore also measured such standards against competition law standards (ECJ, ruling C-184/13 et al., ECLI:EU: C:2014:2147 – API).

If, however, the Member States are subject to competition law in their activities and have to maintain equal opportunities, the **Union** can guarantee this **particularly effectively**. It can do so provided that it harmonises as much as possible, especially matters as are prone to the fact that companies do not maintain equal opportunities, from the outset. In this way, the Union can take countermeasures in advance and limit national structures, which, through private regulation, make it more difficult for competitors to operate on the internal EU market with equal opportunities. It would therefore be contradictory to assume that standardisation would most certainly reduce the competitive chances of some economic operators at Union level. In addition, this development would then have to be laboriously corrected by the competition oversight afterwards.

7 Conclusion

As standardisation bodies for technical rules such as DIN regulations, corporate associations are not allowed to block technical developments in the context of Industry 4.0. This is particularly prohibited by competition law. The prohibition of cartels under Article 101 TFEU also applies to decisions of associations, in which large companies often dominate. It prohibits an unfounded exclusion or even failure to consider the offers of competitors of these large companies – their mere formal participation does not change anything. In such a case, the setting of standards is nugatory and may not be considered for certifications or product approvals. They can also trigger claims for damages.

References

Behrens P (2008) Der Wettbewerb im Vertrag von Lissabon, Editorial, EuZW, 193

Breuer R (1976) Direkte und indirekte Rezeption technischer Regeln durch die Rechtsordnung AöR, 46 et seq

Dietrich J (2012) Dienstleistungen von allgemeinem wirtschaftlichem Interesse, Mitgliedstaatlich differenziertes Unionsrecht in Deutschland und Polen?, Baden-Baden

Ehlers D (2001) Die Grundfreiheiten des europäischen Gemeinschaftsrechts (Teil I), Jura, 266 et seq

Emmerich V (2012) In: Immenga U, Mestmäcker E-J (pub.) Wettbewerbsrecht, tome 1/1, 5th edn. Munich

Frenz W (2012) Handbuch Europarecht, tome 1, Europäische Grundfreiheiten, 2nd edn. Berlin

Frenz W (2013) Fehlende Anerkennung der umsatzsteuerlichen Unternehmereigenschaft und unionsrechtliche Staatshaftung, DStZ, 582 et seq

Frenz W (2014) Anmerkung zu einem Urteil des EuGH vom 17.07.2014 (C-553/12 P), DVBl, 1455 et seq

Frenz W (2015a) Handbuch Europarecht, tome 2, Europäisches Kartellrecht, 2nd edn. Berlin

Frenz W (2015b) Energieunion ultra vires?, EnWZ, 481 et seq
Frenz W (2015c) Staatliche Wettbewerbseinwirkungen nach dem EuGH-Urteil API, WRP, 421 et seq
Frenz W (2016a) Handbook of EU competition law. Berlin
Frenz W (2016b) Industrie 4.0 und Wettbewerbsrecht, WRP, 671 et seq
Jakobs M (2012) Standardsetzung im Lichte der europäischen Wettbewerbsregeln, Baden-Baden
Nowak C (2009) Binnenmarktziel und Wirtschaftsverfassung der Europäischen Union vor und nach dem Reformvertrag von Lissabon, EuR, Beiheft 1, 129 et seq
Palandt O (Begr.) (2019) Bürgerliches Gesetzbuch: BGB, 78th edn. Munich
Rose V, Bailey D (2013) Article 101(1). In: Bellamy C, Child GD (Begr.), European Union law of competition, 7th edn. Oxford, para 11.031
Schröter H, Voet van Vormizeele P (2014) In: Schröter H, Jakob T, Klotz R, Mederer W (pub.), Europäisches Wettbewerbsrecht, 2nd edn. Baden-Baden
Seibel M (2013) Abgrenzung der “allgemein anerkannten Regeln der Technik” vom “Stand der Technik”, NJW, 3000 et seq
Spindler G (2018) Plattformen und Plattformregulierungen als Alternative zu Dateneigentumsrechten. In: Hornung G (pub.): Rechtsfragen der Industrie 4.0, Datenhoheit – Verantwortlichkeit – rechtliche Grenzen der Vernetzung, Baden-Baden, p 151 et seq
Stockenhuber P (2017) In: Grabitz E, Hilf M, Nettesheim M (pub.) Das Recht der Europäischen Union: EUV/AEUV, Loseblatt, Effective: 05/2018
Triantafyllou DN (2014) Der Staat als Garant der Wettbewerbsstruktur, EuZW, 734 et seq

Administration in the Age '4.0'

Annette Guckelberger

Contents

I would like to thank Bianca Böhme, LL.M., Europa-Institut Saarland University, for her valuable support with the translation.

A. Guckelberger (✉)
Chair of Public Law, Saarland University, Saarbrücken, Germany
e-mail: a.guckelberger@mx.uni-saarland.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_16

1 Introduction

The formulation '**Industry 4.0**', which is linked to the usual version designation for software products, was coined by *Henning Kagermann*, *Wolf-Dieter Lukas* and *Wolfgang Wahlster*. After it was used in public for the first time in 2011 on the occasion of the Hanover Fair (Wikipedia article 'Industrie 4.0'; Kagermann et al. 2011), this term is now considered established. In contrast, **the term '4.0' is rarely used to refer to administration**. If this happens nevertheless, there are different perceptions about its meaning. This can be explained by the fact that unlike industry, administration generally does not produce goods, but decisions based on information (Schuppan and Köhl 2016, p. 28). Other causes are likely to include the administration's lagging behind in digitalisation and the different legal framework conditions in the state-citizen relationship. Instead of simply transferring the technology-related abbreviation '4.0' one-to-one to the administration, it makes sense to reflect on which elements of Industry 4.0 could be further developed in the area of administration, and how a good administration should position itself with regard to Industry 4.0. Based on this, an overview of the digitalisation of the administration will be given below. It will become apparent that individual principles that shape Industry 4.0, such as the support of public authority employees through assistance systems, or even the issuing of fully automated administrative acts, are gradually making their way into the everyday life of public authorities.

2 Electronic Government As a Guideline and Project

Not only the development of society, but also of the administration, is influenced and shaped by technological progress (Denkhaus 2019, ch. 1 para. 9; Richter 2019, ch. 10 para. 2). As early as the late 1950s/early 1960s, large-scale computer systems were used, especially in 'number-heavy' mass procedures, such as tax or social administration (Denkhaus 2019, ch. 1 para. 9; Guckelberger 2019a, p. 237; Guckelberger 2019b, para. 4; Kaiser 2009, p. 233 ff.; von Lucke 2019, ch. 2 para. 1).

At a time when increasingly efficient computers and the use of the internet were gradually gaining ground, the era of e-government began (Britz § 26 para. 18; Eifert 2006, p. 20; Guckelberger 2019a, p. 237; Guckelberger 2019b, para. 9). **In 1993**, for example, US Vice President ***Al Gore*** used the term 'electronic government' to reduce the backlog of the US administration by using information and communication technologies (hereafter: ICT) (Gore 1993, p. 112 f.). Although it was initially a costly undertaking, he claimed that 'we can design a customer-driven *electronic government* that operates in ways that 10 years ago, the most visionary planner could not have imagined' (Gore 1993, p. 112, italics by the author).

Despite its vagueness (see for example Eifert 2006, p. 20 f.; Skrobotz 2005, p. 21), since **the beginning of the first decade of the 21st century**, the Anglicism 'e-government' has increasingly become established **in German administrative**

language (on the charm of this abbreviation, Voßkuhle 2012, § 1 para. 65; on its use in administrative language, Duden online), and functions today as a **model** according to which **all administrative action makes comprehensive use of ICT**, i.e. the use of ICT is no longer limited to specific, selective auxiliary functions (Denkhaus 2019, ch. 1 para. 14).

2.1 Concept of E-Government

According to the well-known ***Speyer* conceptualisation** of the term by *Heinrich Reinermann* and *Jörn von Lucke*, e-government includes 'the management of business processes in connection with government and administration with the help of information and communication technologies via electronic media' (Reinermann and von Lucke 2002, p. 1). E-government is said to encompass processes within the public sector (Government to Government = G2G) as well as those between the public sector and the citizens (Government to Citizen = G2C) or the business community (Government to Business = G2B) and vice versa (Reinermann and von Lucke 2002, p. 1). According to *Reinermann/von Lucke*, the benefits of e-government include acceleration of administrative proceedings, improvements in service, quality and organisation, better performance and transparency, as well as cost savings (Reinermann and von Lucke 2002, p. 6). Although the *Speyer* conceptualisation of e-government does not express enough that e-government goes beyond the mere electronic replication of existing proceedings (Britz 2012, § 26 para. 2; Schliesky 2009, p. 15; in this sense Siegel 2014, p. 243; Scheer et al. 2003, p. 3 f.), *Reinermann* and *von Lucke* were also aware of these potentials.[1]

Due to the specifics of ICT, it is also possible to put **administrative organisation and procedure to the test** and to include considerations on the **user-friendliness** of administrative services in their design.[2] After holding back for a long time with legal requirements on e-government, the legislature is increasingly resorting to **legal regulations to promote administration electronification more vigorously** (see Eifert 2014, p. 422 f.; Guckelberger 2019a, p. 279; Guckelberger 2019b, paras. 50, 307). Initially, the federal legislature limited itself to issuing regulations on the conditions under which written form can be replaced by electronic form (§ 3a Administrative Procedure Act, VwVfG), but since 1 July 2014, all authorities executing federal law have been obliged under § 2 (1) Federal E-Government Act (EGovG Bund) to open up a point of access for the transfer of electronic documents, including such documents provided with a qualified electronic signature (Gesetz zur

[1] On the fact that e-government is more than the mere electronification of outdated and inefficient administrative processes, Scheer et al. (2003), p. 3 f.; Reinermann and von Lucke (2002), p. 2 ff.

[2] Britz (2012), § 26 para. 2 ff.; Guckelberger (2019a), p. 238 f.; Heckmann (2018), ch. 5 para. 3 f.; Siegel (2014), p. 243. Eifert (2006), p. 21 therefore understands e-government as the increased use of ICT to improve administration.

Förderung der elektronischen Verwaltung sowie zur Änderung weiterer Vorschriften, Federal Law Gazette 2013, p. 2749).

The **optimising approach of e-government** is expressed in § 9 (1) and (3) EGovG Bund. Accordingly, the federal authorities shall first document, then analyse and optimise administrative procedures which are to become largely electronically-based for the first time, as well as in the event of a substantial change to such procedures or the IT systems used. In the interest of the parties involved, the necessary workflows should be designed so that information on progress status and on the further course of the process can be retrieved by electronic means, together with contact information regarding the competent point of contact at the time of the inquiry concerned. Overall, the use of ICT can lead to changes in the administration's work processes in terms of time, content, organisation and/or even location (Löbel et al. 2018, p. 299). This change from regulations that merely enable electronic administrative action to mandatory requirements[3] already points to a need for improvement in Germany's existing e-government.

2.2 Realisation of E-Government

In **European and international digital transformation and administrative service rankings, Germany's scores are usually only mediocre**.[4] In its 2016 annual report, the National Regulatory Control Council (NKR) came to the following disillusioning conclusion regarding the realisation of e-government in the previous year: '[T]here is no efficient eGovernment in Germany, on the contrary, there is a heterogeneous and fissured IT and eGovernment environment with many isolated solutions and individual lighthouses' (NKR 2016, p. 71). According to the 2018 annual report, awareness of the problems in Germany with regard to a structural deficit in e-government has increased, and efforts are gradually being made to implement it successfully. However, 'a major breakthrough and a full-coverage user-friendly online offer for all important administrative services still remain to be implemented' (NKR 2018, p. 35). In its annual report 2020, the NKR emphasised that only the systematic implementation of digital public services enables efficient crisis management in the context of the Covid-19 pandemic (NKR 2020, p. 8).

---

[3] Denkhaus (2019), ch. 1 para. 42 speaks in this respect of the transition 'from the mere enabling function to an "activating function" of the law of electronic administration' (translation by the author).

[4] In Capgemini, eGovernment Benchmark (2018), p. 48, Germany is assigned to countries in the unconsolidated e-government category due to its medium level of digitalization and is criticised for the low use of online services by the private sector. In the Digital Economy and Society Index Report 2018 – Country Reporting, Germany, p. 10, Germany ranks 21st out of 28 Member States for digital public services.

The main barriers to e-government in Germany are the lack of awareness, availability or consistency of online administrative services, including their difficult-to-navigate structures (eGovernment Monitor 2018, p. 18). Compared to the neighbouring countries of Switzerland and Austria, the majority of contacts with public authorities in Germany still takes place in person. This is primarily attributed to the need for advice (eGovernment Monitor 2018, p. 33) and to the fact that the respondents prefer to complete processes with a high level of commitment, such as submitting a building application, in person (eGovernment Monitor 2018, p. 36).

Nevertheless, **the future of e-government** is viewed optimistically, as the barriers to use as well as the data protection concerns among the population have declined and many respondents have expressed the desire to handle administrative matters electronically (eGovernment Monitor 2018, p. 4). According to the eGovernment Monitor 2020, the use of digital administrative services exceeded the 50% mark for the first time in the relevant reporting period (eGovernment Monitor 2020, p. 14). As a side-effect of the Corona pandemic, the population's openness to e-government services is growing (eGovernment Monitor 2020, p. 19). While in 2018 77% of respondents could still well imagine digital government advisors, often referred to as GovBots, guiding them through application forms or providing other assistance (in South Korea the image of the 'personal secretary' is used for this, see Han 2019, p. 546), this figure has fallen by 10% in 2020 (eGovernment Monitor 2020, p. 44). Respondents' scepticism also grows as the autonomy of digital assistants increases.[5] Overall, it can be assumed that the Corona pandemic will lead to a digitalisation push in the administration, especially with regard to the authorities' IT equipment as well as the nationwide introduction of electronic record keeping.

As a result, it is currently difficult to make generally valid statements about online government services in Germany (Groß 2004, p. 402; Guckelberger 2019a, p. 250; Guckelberger 2019b, para. 155 ff.). This is because the administration is a heterogeneous entity, the financial strength of the Federal states (*Länder)* and municipalities varies, and broadband expansion and mobile phone coverage are known for not being equally good everywhere.[6] E-government is very advanced in **tax administration**. This is because tax administration is continuously advancing the digitalisation of its internal procedures and external offers, as shown by ELSTER electronic tax return or the pre-completed tax return (NKR 2018, p. 36).

[5] eGovernment Monitoring (2018), p. 43 ff.; on the handling of administrative matters by bots of private law entities communicating with GovBots, Guckelberger (2019a), p. 247; Martini and Nink (2018), p. 1129.

[6] Guckelberger (2019a), p. 250. According to the study of the German County Association (2018), p. 8, 34% of the counties (*Landkreise*) see inadequate broadband expansion as an obstacle to digitalization, 29% complain about the lack of seamless mobile network.

3 Administration 4.0 and Smart Government

To express that the **guiding principle of electronic administration has reformed and climbed to a new level** (Djeffal 2017b, p. 86), formulations such as 'Administration 4.0' or 'Smart Government' are occasionally found in the German-speaking area. For example, a study by the Bertelsmann Foundation from 2017 claims that an Administration 4.0 is needed to shape the framework conditions for Industry 4.0 and for follow-up reasons (Bertelsmann Foundation 2017, p. 11). A study by McKinsey&Company together with bitkom from November 2018 emphasises the need for smart governance as an intelligent use of data to improve services, processes and decisions in public administration (McKinsey&Company and bitkom 2018, p. 3).

One encounters similar developments abroad. For example, the South Korean government has decided to use e-government to achieve W.I.S.E.-government (Wonderful Mind-Caring, Innovative Problem-Solving, Sustainable Value-sharing and Enhanced Safety-Keeping Government) in the face of technological progress (Han 2019, p. 544).

3.1 New Guiding Principles for the Administration?

The problem with these new terms in the German-language literature is that **the ideas about their meaning diverge** and these concepts are less consensual than e-government (see also Schuppan and Köhl 2016, p. 27). In some cases, the term Administration 4.0 refers to traditional e-government and its modernisation impulses,[7] so that it is merely a new label for what already exists (Schuppan and Köhl 2016, p. 27, speak in this respect of old wine in new bottles). In view of the reference to Industry 4.0, however, this concept could also be narrowed down to how the administration should best be positioned with regard to its needs (the statement by the Bertelsmann Foundation 2017, p. 11, can be understood in this direction). Other considerations go in the direction of transferring the regulatory principles of Industry 4.0 to the administration.[8] However, caution is urged here because the administration generally makes decisions based on information (Schuppan and Köhl 2016, p. 26).

Still others reject the use of the German neologism '4.0' for administration due to its lack of international applicability or its over-emphasis on technology (Schuppan and Köhl 2016, p. 28) and prefer other formulations, such as 'Smart Government' (von Lucke 2019, ch. 2 para. 3; von Lucke 2016; Hölzel 2017, p. 1017), in reference

[7] See also von Lucke (2019), ch. 2 para. 14. The paper by Hogrebe and Kruse (2014) does include the Internet of Things in the considerations in some place, see for example p. 50.

[8] A connection to the Internet of Things and the new technological developments is made, for example, by Djeffal (2017a), p. 808 ff. See also Richter (2019), ch. 10 para. 4.

to the objects described as smart (von Lucke 2019, ch. 2 para. 3) or the concept of the ‘Smart City’ (see von Lucke 2019, ch. 2 para. 7, 16; Richter 2019, ch. 10 para. 1). According to *von Lucke*, the Anglicism ‘smart’ has meanwhile become a globally recognised term for the next level of digital development, which is essentially about intelligent networking. In the Internet of Things, everyday objects are equipped with sensors, actuators and communication units and are embedded in cyber-physical systems so that the objects can interact with each other or with humans (von Lucke 2019, ch. 2 para. 4 f.).

Researchers at the Zeppelin University have therefore updated the Speyer definition of e-government to **smart government** as ‘the handling of business processes related to governing and administration (government) with the help of *intelligently networked* information and communication technologies’ (so-called **Häfler definition**; von Lucke 2019, ch. 2 para. 18, translation and italics by the author). In smart government, the state and administration use intelligently networked objects and cyber-physical systems to fulfil their tasks efficiently and effectively, whereby the service portfolio of e-government and open government, including big data and open data, is to be included (von Lucke 2019, ch. 2 para. 18).

Demaj wants to use the term smart government for the governmental use of technologies such as the blockchain, artificial intelligence and big data, where, unlike traditional e-government, digital coordination and transaction structures no longer have an analogue counterpart (Demaj 2018, p. 128).

Without taking a position here on whether such conceptual neologisms are not rather confusing in view of their need for explanation and the different meanings (on this, for example, Guckelberger 2019a, p. 243; Guckelberger 2019b, para. 72), they at least draw attention to the fact that **digitalisation consists of a process of continuous learning and change** (Böhmann 2018, thesis 1). The state and administration are therefore called upon to reflect on how they deal with these developments,[9] in particular whether they can also make use of technological innovations for their own purposes. This way, precisely those technical innovations that have received too little attention or been neglected in e-government to date come into focus (Schuppan and Köhl 2016, p. 32). *Von Lucke* therefore rightly demands that ‘in times of the Internet of Things and the Internet of Services, the Federation and the Federal states must think about the further development of state IT architecture models’ (von Lucke 2019, ch. 2 para. 53, translation by the author).

[9] On the fact that the state and administration should recognise and assess disruptive developments at an early stage and adapt to them with new approaches, von Lucke (2019), ch. 2 para. 12.

3.2 Examples of Applications for the Implementation of Technical Innovations

It is only gradually being recognised that the **Internet of Things can provide valuable assistance**, for example in the area of **public infrastructures**. An example of this is the equipment of transport infrastructures, which informs the competent authorities about the condition and use of bridges, motorways or seaports.[10]

By communicating with traffic lights, emergency vehicles can reach their destination as quickly as possible by adapting the traffic signals (Höhl 2018; on the implementation of intelligent traffic systems, Djeffal 2017a, p. 810). The deployment of emergency services can also be facilitated, for example, by providing them with additional augmented reality information using glasses (von Lucke 2019, ch. 2 para. 27 f.; on augmented reality without reference to administration, Boehme-Neßler 2018, p. 9). Furthermore, it is assumed that objects equipped with sensors can, for instance, facilitate food inspections (Schuppan and Köhl 2016, p. 29). In the smart government study by McKinsey&Company/bitkom, the data-based catch quota monitoring of fishing vessels in German waters is cited as a best practice example of intelligently networked administrative action (McKinsey&Company and bitkom 2018, p. 13).

While human officers face limitations when analysing large amounts of data, **IT-based big data analyses** are expected to recognise patterns or provide new insights (for a detailed discussion, see Hoffmann-Riem 2018, p. 20 ff.). For example, under § 4 (1) Passenger Name Record Act (FlugDaG), the Passenger Information Unit matches passenger data with data files and patterns to identify certain persons who have committed or will commit certain criminal offences in the near future (for more details, including legal concerns, Arzt 2017, p. 1023; Rademacher 2017, p. 370 f.).

Predictive policing uses software-based crime forecasting technologies to better control the deployment of police forces (*Bundestag*-Printed Matter 19/1513, p. 1; Seckelmann 2019, ch. 22 para. 51; Singelnstein 2018, p. 1; Zenner 2018, p. 117 ff.). At present, such forecasting technologies are being tested in individual federal states with regard to residential burglaries (for more details, see *Bundestag*-Printed Matter 19/1513, p. 2 f.).

AI systems that can learn from given data have been tested at Berlin-Südkreuz station for biometric facial recognition. Even if the number of false hits appears low at less than 0.1% according to the first evaluation report (BMI press release 2018), this would result in serious follow-up measures against a large number of falsely identified persons at highly frequented locations (analysis by Gigerenzer 2018). While according to a BMI press release from September 2019, such video surveillance is to become an important support tool for the Federal police to increase security at railway stations in the near future (BMI press release 2019), no regulation

[10] Schuppan and Köhl (2016), p. 29; von Lucke (2019), ch. 2 para. 31 f., see also Djeffal (2017a), p. 809; on the smart fire brigade, von Lucke (2019), ch. 2 para. 25 ff.

on the use of electronic facial recognition is planned in view of the 'false positive' cases according to the key issues paper on the amendment of the Federal Police Act (BPolG) of late November 2020 (lto of 1.12.2020). AI systems are currently already being used by the Federal Institute for Drugs and Medical Devices (*Bundesinstitut für Arzneimittel und Medizinprodukte*) for knowledge assurance, data processing and user support, or by the Robert Koch Institute for the detection of statistically conspicuous accumulations of reported infectious diseases (*Bundestag*-Printed Matter 19/1982, p. 12 f.). The Federal Office for Information Security (*Bundesamt für Sicherheit in der Informationstechnik*) uses machine learning to detect cyberattacks and the Federal Office for Economic Cooperation uses deep learning methods to recognise, index and allocate incoming mail (*Bundestag*-Printed Matter 19/1982, p. 14).

3.3 Benchmarks for Their Use

These examples, albeit few, prove that the state and administration are also concerned with the implementation of technical innovations. They are also involved in the digitalisation process of society as a whole (Denkhaus 2019, ch. 1 para. 60). The administration must position itself both internally and externally in terms of technology, personnel, organisation and procedures so that it can fulfil its tasks in the age of digital transformation in the best way possible (Windoffer 2018, p. 364). By also making use of technical innovations and possibilities, such as the Internet of Things, the public sector can fulfil **certain tasks better, more economically and more efficiently**.[11]

In the age of **big data**, information can be better structured and analysed (Schuppan and Köhl 2016, p. 30). Real-time data analysis leads to more correct and fairer decisions by the authorities (Schuppan and Köhl 2016, p. 30). If the state has a large amount of data at its disposal, it can be shared with private parties (on the accessibility of data, see also Djeffal 2017a, p. 810; on the exchange of information between authorities, see Schuppan and Köhl 2016, p. 30) and used for innovations. However, because nowadays certain private companies have particularly large amount of data at their disposal, the question increasingly arises as to whether and under what conditions the state may also access their data to fulfil its tasks (on this, for example, Barczak 2020, p. 997 ff.). In general, digitalisation is seen as a great opportunity to increase prosperity and quality of life, as well as to promote the future viability of the economy or more user orientation (see, for example, Denkhaus 2019, ch. 1 para. 1; von Lucke 2019, ch. 2 para. 60).

[11] Djeffal (2017a), p. 811; see also Denkhaus (2019), ch. 1 para. 8, according to which e-government contributes to the performance and efficiency of the administration, to administrative modernization and to the reduction of bureaucracy.

When considering the use of ICT by the state, the **various interests of the common good must be appropriately balanced**, and the state agencies must not allow themselves to be unilaterally guided by certain interests, such as the promotion of the business location or the simplification of administration, without considering the specific task.[12] According to surveys, there is no e-government hype among the population (Windoffer 2018, p. 366) and, furthermore, care must be taken to ensure that hardly any or non-tech-savvy sections of the population are not left behind in the handling of administrative matters compared to digital natives (keyword: **avoidance of a digital divide in society**). Therefore, a narrowing of the state-citizen relationship to exclusively digital public authority services is rarely considered at present.[13] Because digitalisation is more advanced in the business sector, e-government solutions in this area can be promoted and be more strongly aligned with their interests (Guckelberger 2019c, p. 457 ff.).

4 Legal Framework for the Digitalisation of the Administration

The digitalisation of the administration initially requires **considerable** financial, organisational, personnel and technical **efforts** (Denkhaus 2019, ch. 1 para. 46; see also Windoffer 2018, p. 367 f.). Since digitalisation has not only advantages but also disadvantages, it must be designed in such a way that the latter are reduced as much as possible, e.g. by ensuring **sufficient IT security and data protection** (Windoffer 2018, p. 367; see also Djeffal 2017a, p. 811 ff.). In this respect, to what extent the **legal framework** can or must be readjusted should also be examined.

Accordingly, the currently applicable **data protection law**, which is largely determined by EU law, sets narrow limits on big data analyses or blockchain applications.[14] Since this considers the protection of fundamental rights from Art. 7, 8 CFR or the 'right to informational self-determination' (*Recht auf informationelle Selbstbestimmung*), it should be carefully examined whether it is possible to agree on certain relaxations in this respect, in which personal data are still protected in a balanced manner.[15] In such considerations, it should be taken into account that even

---

[12] Windoffer (2018), p. 364, who emphasises on p. 366 that when considering citizens' expectations, these should actually be elicited and not be based on assessments from the administration's own point of view or the self-interests of consulting companies.

[13] See also Windoffer (2018), p. 366, according to which the administrative offerings are to be supplemented, but not substituted, by digital solutions, see also Denkhaus (2019), ch. 1 para. 62. For detailed information on the advantages of electronic promulgation of laws, see Guckelberger (2009), p. 69 ff.

[14] Martini and Weinzierl (2017), p. 1252 ff; on big data, Paal and Hennemann (2017), p. 1700 f.; on blockchain, Schrey and Thalhofer (2017), p. 1436.

[15] Denkhaus (2019), ch. 1 para. 55; Kühling and Sackmann (2018), p. 686. For a very far-reaching break-up of existing data protection law, Veil (2018), p. 686 ff.

the president of the Microsoft group considers its legal regulation indispensable in view of the broad social impact of increasingly intelligent facial recognition (Smith 2018).

The complete transfer of administrative decisions to learning **AI systems** is strictly limited by fundamental rights, the principle of democracy and the rule of law (on the latter, see Guggenberger 2019, p. 850).[16] Such use on the part of the authorities should only be considered after the systems have been trained and tested until they reach such a low error potential that this risk appears acceptable (Guckelberger 2019a, p. 276; Guckelberger 2019b, para. 586). Further requirements are that the AI systems are regularly monitored during the deployment phase and that there is always the possibility of intervention by human officials (Guckelberger 2019a, p. 276 f.; Guckelberger 2019b, paras. 248, 587; Guggenberger 2019, p. 850).

Particularly where the use of AI has an impact on fundamental rights, the democratically legitimised parliament must regulate all essential points (Guckelberger 2019a, p. 277; Guckelberger 2019b, para. 625; so-called *Wesentlichkeitstheorie* 'essential matters doctrine'). Legal requirements on the quality of learning data or on the non-discriminatory modelling of AI systems can make an important contribution to ensuring the legality of their output (Guckelberger 2019a, p. 276; Guckelberger 2019b, para. 578).

Law can influence not only digitalisation in general but also that of the administration (on the fact that law can be the reason for technological development, influence it in a formative way or limit it, Djeffal 2018, p. 503 ff.). In the meantime, there are **a large number of legal provisions** that affect or have an impact on e-government, from union law to constitutional law to ordinary law. At the end of 2014, a provision on basic digital services, access to public authorities and courts was added to the **constitution of Schleswig-Holstein** in Art. 14, and on the protection of citizens' digital privacy in Art. 15. At the end of October 2018, Art. 26d sentence 1 State Constitution (LVerf) was adopted in a referendum in **Hessen**, according to which the state, the municipalities and municipal associations promote the establishment and maintenance of technical, digital and social infrastructure. Even if this is only a state objective, it can be significant, for example, for the expansion of broadband or the closing of mobile phone holes. Such measures serve not only Industry 4.0, but also e-government. Since, for reasons of space, these regulations cannot all be dealt with comprehensively here, only three important aspects of e-government will be dealt with below, also with regard to Industry 4.0.

[16] Hill (2018), p. 289 f.; Martini and Nink (2018), p. 1134; generally on the limits of legitimacy and responsibility, data protection, data security and sufficient availability, Windoffer (2018), p. 370.

4.1 Comprehensive Information Technology Access to Administrative Services

As digitalisation progresses, businesses and citizens expect more and more to be able to use administrative services electronically as easily as they are used to from private sector
services (McKinsey&Company and bitkom 2018, p. 3). To overcome the limits resulting from the *Verbot der Mischverwaltung* (prohibition of shared administration), **Art. 91c (5) GG** was inserted in July 2017, according to which comprehensive access by means of information technology to the administrative services of the Federation and the Federal states (= Länder) shall be regulated by a federal law with the consent of the *Bundesrat* (= Federal council).

4.1.1 Portal Network

§ 1 (2) Online Access Act (OZG) obliges the federal, state and local levels of government to 'link' their administrative portals 'in a **portal network**' (NKR 2018, p. 36; see also § 2 (1), § 3 (1) OZG; *Bundestag*-Print Matter 18/11131, p. 16). Since 20 September 2018, the Federation has released **a beta version of its administrative portal**.[17] The linking of this portal with the administrative portals of the Federal states will in future result in the portal network which is invisible to external users. Its great advantage is that users will no longer have to worry about the competent authority for the respective service, because they will be guided to the desired service via each participating portal (*Bundestag*-Printed Matter 18/11131, p. 16, see also Berger 2018, p. 442).

Under **§ 1 (1) OZG**, the Federation and the Federal states must 'also offer their **public services online through administrative portals' by the end of 2022** at the latest (NKR 2018, p. 36). While the position is sometimes taken with regard to the **local governments** that this obligation in the context of Art. 91c (5) GG only concerns the consolidation of already existing online services (Martini and Wiesner 2017, pp. 200, 205; Ruge 2018, Art. 91c para. 45), such a restriction is to be rejected in view of the meaning and purpose of comprehensive access by means of informational technology as well as the legislative history (Siegel 2018, p. 187; see also Gröpl 2018, Art. 91c para. 58 ff.; Seckelmann 2018, Art. 91c para. 56; Stocksmeier and Hunnius 2018, p. 1 ff.). If one follows this reading, comprehensive electronic access to the approximately 600 public services must be possible by the end of 2022. Despite considerable efforts on the part of the stakeholders, the NKR is concerned

[17] Link to the beta version available at: https://www.beta.bund.de/DE/Navigation/Home/home_node.html;jsessionid=E91ED7226309C06BA1F42F71302F79E2.live3872; see also IT Planning Council on the launch of the beta version, https://www.it-planungsrat.de/DE/ITPlanungsrat/OZG-Umsetzung/Portalverbund/02_VerwPortal_Bund/VerwPortal_Bund_node.html (both accessed on 30 January 2019).

about meeting this deadline, since according to statements by those with political responsibility, ‘the OZG would not be implemented by the announced date and what mattered was not the number of services as long as the services of importance were online’ (NKR 2020, p. 30).

Under **§ 3 (2) OZG**, **service accounts** (see § 2 (5) OZG) are provided in the portal network. Users can use these to identify themselves uniformly for the available administrative services, whereby the special requirements of individual administrative services for user identification must be considered. Since companies have different requirements for user accounts than citizens, the IT Planning Council is currently having them specially surveyed to continue existing authentication solutions convergently so that companies can log in with only one account (*Bundestag*-Printed Matter 19/12775, p. 10).

Because many people have not had the **eID function of the identity card** activated, it is now issued with the eID function already installed (on the deactivation of this function, see § 10 (2) Act on Identity Cards and Electronic Identification, PAuswG) in accordance with § 10 (1) PAuswG. Germans living abroad can now also use German e-government services by having their foreign address included in the passport and on the chip. Likewise, an eID card is envisaged on a voluntary basis for citizens of a Member State of the EU or a Contracting Party of the Agreement on the European Economic Area (Act on a Card with an Electronic Identification Function for Citizens of the European Union and the European Economic Area, eIDKG, Federal Law Gazette I 2019, 846 ff.).

To reconcile the user-friendly design so that citizens and companies do not have to provide the authorities with their data more than once (**‘once only’ principle**) with data protection law, the consent solution is used, i.e. when filling out a form or similar actions, the users themselves can decide whether to enter the data manually or whether to transfer it directly from the citizen or company account (Berger 2018, p. 443).

4.1.2 ‘Your Europe’ Portal

Even if, according to the legislative materials, the OZG does not imply any personal legal entitlement to online access to administrative services (NKR 2020, p. 30; *Bundestag*-Printed Matter 18/11135, p. 92; see also Herrmann and Stöber 2017, p. 1404; Schliesky and Hoffmann 2018, p. 196), this position will be difficult to maintain in the future within the scope of **Regulation (EU) 2018/1724 (hereinafter: SDG Regulation) establishing a single digital gateway to provide access to information, to procedures, and to assistance and problem-solving services**.[18] Art. 2 (1) SDG Regulation obliges the Commission and the Member States to

[18] Regulation (EU) 2018/1724 of the European Parliament and of the Council of 2 October 2018 on the establishment of a single digital gateway to information, procedures, assistance and resolution services and amending Regulation (EU) No. 1024/2021, OJ EU, L 295/1, p. 1 ff.

establish such a digital gateway, which consists of a common user interface managed by the Commission. This shall be integrated into the **Your Europe portal** and provide access to relevant union and national webpages.

Under Art. 1 (1) (a) SDG Regulation, the regulation lays down rules to provide citizens and businesses with easy access to high-quality information, to efficient procedures and to effective assistance and problem-solving services with regard to union and national rules for citizens and businesses to exercise their rights derived from union law in the field of the internal market. The regulation, which among other things contains very **detailed requirements, e.g. on the quality of the information to be made available, including translation into another official language** of the union (Art. 9, 12) and elaborates on the **'Once Only' principle** in Art. 14 (7), enters into force within the tight timeframe set out in Art. 39 (by the end of 2023). It contains important requirements for competent authorities (see Art. 3 (4) SDG Regulation) to successfully implement digital administrative access and is expected to have a spill-over effect (Siegel 2019, p. 908 f.).

To ensure that the single digital gateway **meets with the approval of users**, the regulation also prescribes, for example, information on the average, estimated or indicative time of the procedures (Art. 10 (1) (g)) as well as on any delays and extension of deadlines or any consequences thereof (Art. 10 (2)). Art. 25 SDG Regulation allows for user feedback on the services of the gateway. By providing information on their satisfaction with the services provided through the gateway and the information made available therein, it is possible to continuously reflect on how to improve the single digital gateway.

4.2 Digitalisation and Administrative Procedures

Due to the increase in electronic communication, the administration must also be able to communicate electronically with citizens and companies, but also with other state agencies. To enable **communication without media discontinuity** despite federalism and the department principle on the state side, it is necessary **to agree on certain uniform standards**. As early as 2009, Art. 91c (1)-(4) was inserted into the basic law with regard to the cooperation between the Federation and the Federal states. They may agree to specify the standards and security requirements for exchange between their information technologies, whereby the IT Planning Council, created based on a state treaty on IT, has a central role in standardisation (for more details, see Guckelberger 2019a, p. 256; Guckelberger 2019b, paras. 262 ff., 302 ff.).

4.2.1 E-Government Laws

In the meantime, individual federal states have regulated some mechanisms for ICT coordination with the municipalities (e.g. § 23 EGovG of Baden-Wuerttemberg (EGovG BW); § 17 EGovG of Mecklenburg-Western Pomerania (EGovG M-V);

§ 21 EGovG of North Rhine-Westphalia (EGovG NRW); § 19 EGovG of Saarland (EGovG SL)), but sometimes also for interdepartmental ICT coordination (e.g. § 19 (6) sentence 1 EGovG BW; § 22 EGovG of Berlin (EGovG Bln); § 17 (1) sentence 1 EGovG of Saxony (SächsEGovG)) in their **e-government acts (hereinafter: EGovG)**. After the Federation enacted an e-government act in 2013 (hereinafter: EGovG Bund), comparable acts are increasingly being found at the state level,[19] e.g. with **requirements for electronic record keeping**. For example, federal authorities should keep their records in electronic form since 1 January 2020. In the EGovG of the Federal states, a later date was chosen in some cases,[20] and in others the municipalities were exempted from these obligations (e.g. Art. 7 (1) sentence 1 EGovG of Bavaria (BayEGovG); § 6 (2) EGovG BW).

While the administrative procedure acts of the Federation and the Federal states are characterised by the model of *Simultangesetzgebung* (simultaneous legislation)—the states follow the regulations in the Federal Administrative Procedure Act by enacting regulations with the same wording—**the regulations in the EGovGs of the Federal states are quite different**. The BayEGovG is strongly characterised by 'individual public rights' (*subjektive öffentliche Rechte*), as can be seen in Art. 2 sentence 1 BayEGovG (under Art. 2 sentence 1 BayEGovG, everyone has the *right* to communicate electronically with the authorities via the internet in accordance with Art. 3–5). Under § 4 (1) sentence 2 EGovG NRW, if citizens or companies choose the electronic communication channel, the authorities should reply to them in the same way. Since the regulations in the **EGovGs are leges speciales in relation to the general administrative procedure acts** (e.g. § 1 (4) EGovG Bund; Art. 1 (1) BayEGovG), the latter provision, for example, displaces the freedom of choice of form in § 37 (2) sentence 1 VwVfG NRW.

It should be noted that under § 1 (2) EGovG Bund, this act shall further apply to the public law activities of the authorities of the federal states, local authorities and other legal entities under public law which are subject to state supervision in executing federal law (however, at the federal state level, this priority is partly attributed to the execution of federal law 'on federal commission', see e.g. Art. 1 (3) BayEGovG), insofar as the standards are not explicitly limited to the federal authorities. Important elements of electronic administration are shaped in the EGovGs. The EGovG of the Federation contains, for example, provisions on electronic access to government (§ 2), information on authorities and their procedures in publicly accessible networks (§ 3), on electronic means of payment (§ 4), electronic receipt of invoices (§ 4a), on required documentation (§ 5), on electronic record-keeping (§§ 6 ff.) or on electronic forms (§ 13). This creates the conditions for

[19] E.g. in Baden-Wuerttemberg: EGovG BW; in North Rine-Westphalia: EGovG NRW; in Mecklenburg-Western Pomerania: EGovG M-VM; in Saarland: E-GovG SL; in Bremen: BremEVerwG; in Berlin: EGovG Bln; in Bavaria: BayEGovG.

[20] E.g. from 2022, § 6 sentence 1 BremEVerwG; § 9 III 1 EGovG NRW; from 2023, § 7 (1) sentence 1 EGovG Bln and from 2025, § 5 (1) sentence 1 E-GovG SL.

handling procedures completely electronically, if necessary (Guckelberger 2019b, para. 713).

4.2.2 Benefits Without Application, Electronic Supports

The increasing regulation of e-government shows that administrative procedures are increasingly to be handled electronically. It should be considered whether more user-friendly e-government services can be created by reducing application requirements (Denkhaus 2019, ch. 1 paras. 15, 20). It is also conceivable that some projects could be made tangible by providing augmented reality information, so that those involved in the process can better comment on it (see Bertelsmann Stiftung 2017, p. 14 f., which cites land use plans as an example of application). Another example of improving administrative procedures is the **electronic advisory services** offered by the authorities, for example by using digital assistants.[21]

4.2.3 Fully Automated Administrative Acts

As a further innovation, it should be emphasised that the Federation and several Federal states have included a provision on **fully automated administrative acts** in their Administrative Procedure Acts (VwVfG).[22] Under **§ 35a VwVfG**, an administrative act may be adopted solely on automated processing provided that (1) this is authorised by law and (2) that the administration has no discretion or prerogative of assessment. These are administrative decisions in which all procedural steps within the administration do not require personnel to process the individual case (Braun Binder 2019, ch. 12 para. 5, although the notification of the administrative act is no longer counted as part of this), and in particular the factual data is also collected, evaluated and verified by machine (Stelkens 2018, § 35a para. 21).

§ 35a VwVfG is intended to clarify that regulations of the authorities in which human officials are only involved in the programming and/or release of the IT

[21] § 25 (1) VwVfG, for example, makes no statement on the form of consultation. For people with an affinity for technology, however, sufficient advisory services must be available in another form. The same applies to matters that are very specific to individual cases, where standardised advisory services are not helpful.

[22] Parallel to this, corresponding norms were also included in the tenth book of the Social Security Code (SGB X) and in the Fiscal Code (AO), which, however, were not worded in the same way. Under § 31a sentence 1 SGB X, an administrative act can be issued entirely by automatic institutions, provided there is no reason to process the individual case by public officials. Under § 155 (4) sentence 1 AO, the tax authorities may make, correct, revoke, cancel and amend tax assessments as well as credits of tax deductions and advance payments exclusively by automated means based on the information and details of the taxpayer available to them, insofar as there is no reason to process the individual case by public officials. Sentence 2 specifies further cases of application for such processing. On these two areas, see Braun Binder (2019), ch. 12 para. 8 ff. (AO) and para. 14 (SGB X).

systems and their possible activation are administrative acts despite the lack of the acting administrative organ's human will formation in individual cases (*Bundestag*-Printed Matter 18/8434, p. 122). The **advantages of such administrative acts** include acceleration, the reduction of enforcement deficits and uniform enforcement of the law.[23] As a result of the liberated resources, human officials can ideally invest more time in dealing with difficult or special cases (Braun Binder 2019, ch. 12 para. 1; *Bundestag*-Printed Matter 18/7457, p. 48).

4.2.3.1 No Discretion or Prerogative of Assessment

§ 35a VwVfG sends a **warning signal** (Braun Binder 2019, ch. 12 para. 12; Ziekow 2018, p. 1171): administrative decisions on norms with **discretionary powers and prerogative of assessment are not suitable for fully automated decisions** because they are regularly highly dependent on the situation and individual case (Groß 2004, p. 409) and currently only human officials can make evaluative decisions (Braun Binder 2019, ch. 12 para. 20; Martini and Nink 2017, p. 2; Siegel 2017, p. 26).

Even if it is speculated that learning **AI systems** could possibly become better than some officials in the long term, especially in this area, due to their flexibility (Thapa and Parycek 2018, p. 62), the use of such systems would only be possible for reasons of legal protection (Art. 19 (4) GG) if the courts can assess based on a statement of reasons whether the IT systems have made a 'discretionary error' (§ 114 sentence 1 Code of Administrative Court, VwGO). Since AI decisions are currently generally obscure to humans (keyword: black box), it remains to be seen whether such progress will be made with explainable AI to meet the legal requirements (Guckelberger 2019a, p. 277 f.; Guckelberger 2019b, para. 521; in detail Wischmeyer 2018, p. 61 ff.). Also, according to the Saarland Constitutional Court (SaarVerfGH), '[s]tate action [. . .], no matter how slightly burdensome it may be in individual cases and no matter how much there is a need for routinised decision-making processes, must not be impenetrable for citizens in a liberal constitutional state [and make them] the immature object of state availability' (SaarVerfGH NJW 2019, 2456, 2458, translation by the author).

4.2.3.2 Mandatory Decisions As a Field of Application

Because conditional programming corresponds to the system type of the machine (Luhmann 1966, p. 36; see also Groß 2004, p. 409 f.; Martini and Nink 2018, p. 1129), it is primarily **'mandatory decisions'** (*gebundene Entscheidungen)* that come into consideration for full automation (Siegel 2017, p. 26; Stelkens 2018, § 35a para. 1). The prerequisite for this is that the **content of the enforcing norms can be**

[23] Guckelberger (2016), p. 403 f., who also discusses the dangers of automated enforcement on the following pages; on acceleration, see Braun Binder (2019), ch. 12 para. 1.

translated into computer language (Groß 2004, p. 410 on IT translatability; see also Guckelberger 2019a, p. 264 with general information on the different views on IT translatability of law in fn. 162). This can be thought of in the case of certain or also undefined legal terms with broad agreement on their content, but not in the case of such legal concepts where, as e.g. in the examination of the requirement of consideration within the factual characteristic of 'blending in' in § 34 (1) sentence 1 Federal Building Code (BauGB), the interests of the party entitled to consideration and the party obligated to consideration must be weighed up.[24]

Another argument against a fully automated issuing of **building permits** is that in building law, it depends considerably on the circumstances of the individual case. More realistic here are partial automations, e.g. automatic distance area calculations (on this topic see also Edenharter 2020, p. 341 ff.). Because the (non-)existence of the constituent elements of a norm must be assessed with regard to concrete facts of life, matters in which the necessary **determination of the facts is amenable to standardisation** are primarily suitable for fully automated enforcement of the law (Stelkens 2018, § 35a para. 46).

Furthermore, **procedural law**, with its lawfulness function and the purpose of early protection of citizens' rights, sets limits to fully automated administrative decisions, for example if the holding of a hearing is prescribed (Braun Binder 2019, ch. 12 para. 23). As things stand, simple, bipolar administrative decisions are particularly suitable for fully automated administrative acts (Braun Binder 2019, ch. 12 para. 19; see also Siegel 2017, p. 26). Examples include the extension of parking permits, the granting of *BAföG* (training allowance) (see § 39 (4) Federal Training Assistance Act; Martini and Nink 2018, p. 1128 f.) or simple subsidy decisions. The field of application for fully automated administrative acts is therefore much narrower than the wording of § 35a VwVfG (without discretion or prerogative of assessment) would lead one to expect at first glance.

4.2.3.3 Reservation of Rights

Since fully automated administrative acts involving the processing of personal data under **Art. 22 (2) (b) General Data Protection Regulation (GDPR)** may only be issued on the basis of legal provisions with appropriate measures to safeguard the data subject's rights and freedoms as well as the legitimate interests of affected persons, and since administrative action under Art. 20 (3) GG must also be in accordance with the law, **§ 24 (1) sentence 3 VwVfG** obliges the authority to consider the actual information of the party involved that is relevant to the individual case and that would not be determined in an automated procedure.[25] By allowing a

[24] On § 34 BauGB only BVerwG, BRS 84 no. 123; on the problems of IT mappability in the case of indeterminate legal terms, Bull (1964), p. 70; Guckelberger (2019a), p. 265 f.

[25] *Bundestag*-Printed Matter 18/8434, p. 122; Guckelberger (2019a), p. 266 f. According to its wording, this standard is not only applicable to fully automated administrative acts.

party to describe specific features of his or her case, e.g. by information in a section or data field that gives citizens the opportunity to provide information that in their view establishes cause for processing by officials, the conditions are created for such matters that do not fit under the schematising IT to be processed with the involvement of an official (Guckelberger 2019a, p. 267; Guckelberger 2019b, para. 686 ff.; *Bundestag*-Printed Matter 18/8434, p. 122; Maurer and Waldhoff 2017, § 18 para. 15, see on the need for improvement of this norm, Berger 2019, p. 1236).

4.2.3.4 Internet-Based Vehicle Registration As an Application Example

Currently, full or partial automation is being advanced in the area of **internet-based vehicle registration**. Under **§ 6g (2) sentence 1 of the Road Traffic Act (StVG)**, an administrative act may be issued entirely by automated means, subject to the more detailed provisions of a statutory instrument pursuant to para. 4, sentence 1, no. 1, if (1) the automated examination of the prerequisites for a decision is carried out based on an automated examination programme set up at the competent authority and operated exclusively by it, and (2) it is ensured that the result of the examination can only be the decision in favour of the application or its rejection. The Fourth Ordinance Amending the Vehicle Registration Ordinance and Other Road Traffic Regulations of 22 March 2019 (Federal Law Gazette I 2019, 382) (hereinafter: ordinance) will expand internet-based vehicle registration and is intended to relieve citizens of a total of 3 million hours of time expenditure and around 20 million euros of material expenditure annually when this procedure is established, as well as commercial owners of around 4.9 million euros (*Bundesrat*-Printed Matter 18/19, p. 2 f.). Although these regulations came into force in October 2019, it was still not possible to carry out such approval procedures in many places in the summer of 2020. According to the order of competences of the basic law (Art. 84 GG), the enforcement of the law is the responsibility of the Federal states, which explain such enforcement deficits with short-term increased security requirements to protect against unauthorised intrusion into the systems as well as still necessary coordination discussions with the Federation regarding individual points.[26] To avoid disappointed expectations, a sufficient period should be planned for the provision of the necessary IT solutions by the state authorities when determining the entry into force of regulations.

§ 15a of the ordinance regulates the admissibility of internet-based registration procedures. Under § 15b (1) sentence 1 of the ordinance, an application must be submitted by the keeper via the information technology system (portal) set up by the registration authority for this purpose. If this is done, the data entered into the portal and generated by the portal in accordance with § 15b (1) sentence 1 of the ordinance are either (1) transferred to the manual processing and decision of the licensing

[26] *Shabaviz*, ZDF.de, 6 August 2020: Online-Kfz-Zulassung, at: https://www.zdf.de/nachrichten/wirtschaft/kfz-neuzulassungen-corona-100.html, last accessed on 24 August 2020.

authority without the latter being bound to the result of the automatic preliminary check, or (2) transferred to the internal information technology procedures of the regulatory authority after automatic checking in the portal together with the decision issued in full by an automatic facility of the licensing authority's portal after its retrieval or at the latest after the end of its provision period.

According to the explanatory memorandum of the draft regulation, automated application processing should be the norm (*Bundesrat*-Printed Matter 18/19, p. 41). At the moment, however, it is limited to three registration processes: the application for a vehicle to be taken out of service, the application for a vehicle to be transferred to another keeper with a transfer of registration plates, and the application for a change of residence of the keeper without a change of registration plates. 'The comprehensive application of the automated decision is hindered by the fact that at the present time, certain initial registration requirements cannot be checked automatically in the case of initial registrations' (see *Bundesrat*-Printed Matter 18/19, p. 41, translation by the author). Elsewhere, the fully automated procedures are described to the effect that 'all prerequisites can be verified electronically [. . .] and [. . .] no discretionary leeway remains' (*Bundesrat*-Printed Matter 18/19, p. 76, translation by the author). § 15f of the ordinance contains a provision on notification in paragraph 1 and on the effectiveness of the administrative acts in paragraph 2. Paragraph 3, sentence 1 of this section makes fully automated decisions subject to review, withdrawal, revocation and new decisions by the regulatory authority for a period of one month after their effectiveness. § 15f of the ordinance takes a special path, as it modifies the general provisions of the VwVfG on the effectiveness, notification and validity of administrative acts.

4.2.3.5 Reasonableness of Fully Automated Decisions

In view of the unease of large parts of the population about fully automated decisions by the authorities (Fischer and Petersen 2018, p. 25), the existing trust in the administration must not be put at risk prematurely (on trust in the administration, Berger 2017, p. 804 ff.). Therefore, the advantages and disadvantages of fully automated law enforcement, on the one hand, and of such enforcement involving officials, on the other, must be put in relation to each other (Trute 2018, p. 323). Procedural efficiency is only one of several procedural purposes (without reference to full automation, Langenbach 2017, p. 25 f.). Art. 1 (1) GG [human dignity] prohibits to turn a person into a 'mere object' of state action (BVerfGE 57, p. 275; 122, p. 271; BVerfG NJW 2017, p. 619). Under Art. 12 (1) LVerf Bremen, '[the human being] stands higher than technology and machines' (on the uniqueness of this constitutional provision, Djeffal 2018, p. 511).

It is true that the (non-learning) IT systems currently used for fully automated administrative acts are of human origin and are based on a release decision by human officials (see also Djeffal 2017a, p. 815 f.). However, this does not change the fact that only those decisions may be provided for full automation that are **reasonable for their addressees without the involvement of a human official in the**

individual case.[27] This is usually not the case in matters with a special need for discussion (for example, to clear up misunderstandings, for emotional but also for symbolic reasons) or with a high intensity of fundamental rights (Guckelberger 2019a, p. 272; Guckelberger 2019b, para. 546; see also Ehlers 1991, p. 340).

In view of the functions of the intra-administrative **preliminary proceedings** (§§ 68 ff. VwGO), decisions on objections against fully automated administrative acts are not to be admitted for full automation (Guckelberger 2019a, p. 272 f.; Guckelberger 2019b, para. 553 ff.; in detail, Martini and Nink 2018, p. 1133 ff.). The more experience is gained with fully automated administrative acts, the more likely it is that considerations can be made over time to expand their scope of application.

To ensure that citizens in particular do not feel helplessly exposed to a mechanised administration, they must be left with **sufficient communication options with officials as a counterpoint** to the depersonalisation of the procedure (Guckelberger 2019a, p. 273 f.; Guckelberger 2019b, para. 677 ff.; Stelkens 2018, § 35a para. 53; see also Hufen and Siegel 2018, para. 349).

4.2.3.6 Right To Be Heard

Because fully automated IT systems are not capable, at least not according to their current state of development, of hearing the participant before issuing the administrative act encroaching on his rights under § 28 (1) VwVfG, such decisions can only be issued if a **hearing** under § 28 (2) VwVfG may be omitted. This is the case, for example, when the intent is not to diverge, to his disadvantage, from the actual statements made by a participant in an application or statement (no. 3) or if the authority wishes to issue administrative acts using automatic equipment (no. 4 var. 3).

Since at the time of the enactment of this regulation in the 1970s, partially automated administrative decisions were envisaged (Stelkens 2018, § 35a para. 19 f.; also Lazaratos 1990, p. 192 f.) and ICT advances could hardly be foreseen, the last-mentioned possibility of exception is to be handled narrowly and the dispensability of the hearing must not become the rule in the age of digitalisation (Guckelberger 2019a, p. 274; Guckelberger 2019b, para. 695 ff.). Already when deciding on (allowing) full automation, it must be carefully examined whether the motives in favour of hearing the participant do not outweigh those in favour of full automation, so that the latter should be omitted (Guckelberger 2019b, para. 696; see also Roth-Isigkeit DÖV 2020, p. 1018 ff.).

[27] Guckelberger (2019a), p. 272; see also Müller-Franken (2018), p. 121. On the need to decide how far technology should advance with people and society and what is desirable, Schuppan and Köhl (2016), p. 32.

4.2.3.7 Grounds for an Administrative Act

Because today's IT systems can do much more than those from the 1970s, the use of § 39 (2) no. 3 VwVfG, according to which written or electronic administrative acts do not require a **statement of grounds** when the authority issues administrative acts with the help of automatic equipment and individual cases do not merit a statement of individual grounds, should be restrained (Hufen and Siegel 2018, para. 480; Maurer and Waldhoff 2017, § 18, para. 480; Maurer and Waldhoff 2017, § 18 para. 10) and this exception could be deleted *de lege ferenda* (Guckelberger 2019a, p. 274; Guckelberger 2019b, para. 696; Polomski 1993, p. 164 ff.; Stelkens 2018, § 39 para. 97).

4.2.3.8 Inspection of Documents by Participants

Since the participants cannot effectively exercise their right to inspect files under § 29 VwVfG after the start of the IT system until the fully automated administrative acts are issued, this deficit must be compensated for by granting subsequent access to documents after the notification of the act (in detail, Guckelberger 2019b, para. 664 f.). If the participants are not very familiar with the documentation of the IT systems, they must be enabled to understand its content, for example by having it explained to them in an understandable way upon request.[28]

4.2.3.9 Notification by Retrieval of the Administrative Act

Whereas the administration previously had to ensure the transmission of the administrative act to the participant, **electronic administrative acts may now be disclosed with the consent of the participant** or his authorised representative by making them available for electronic retrieval **via publicly accessible networks** (§ 41 (2a) sentence 1 VwVfG). In this context, it must be ensured that the retrieval only takes place after authentication of the authorised person and that the latter can also store the administrative act (§ 41 (2a) sentence 2 VwVfG). However, under § 41 (2a) sentence 4 VwVfG, if an administrative act provided in this way is not retrieved within ten days after notification of the act's availability has been sent to the person authorised to retrieve the relevant data, it is deemed not to have been effected, so that it must be notified again in accordance with sentence 5. Since the

[28] Guckelberger (2019a), p. 275; on the fact that in fine proceedings, measurement data must be made accessible for defence purposes for reasons of fair trial and the right to be heard, SaarVerfGH, NJW 2019, 2456 ff. as well as Wendt (2018), p. 441 ff.

ten-day period is too long,[29] a more practical solution *de lege ferenda* should be found in this regard.

4.3 Opening Up the Data Pool for Innovative Uses

The **public sector has huge amounts of data at its disposal**, such as meteorological data, which can also be used by the business sector to promote innovation, e.g. by using it as raw material for the provision of data-based services and applications (COM(2018) 234 final, p. 1). Data in digital form is therefore also figuratively referred to in some places as the 'fuel of the future' or 'the new oil' (see for example *Bundestag*-Printed Matter 18/11614, p. 1). The EU Commission sees access to public sector data and its further use as an 'important driver in the area of big data analytics and Artificial Intelligence' (COM(2018) 234 final, p. 2).

While the **principle of limited public access to files** has applied in Germany for a long time (on this principle, e.g. BVerwG NVwZ 2011, p. 236; Guckelberger 2014, p. 411 with further references), there is an **increasing focus on a cultural change towards more public access and further use of official data by everyone** (*Bundestag*-Printed Matter 18/11614, p. 11). Since 7 December 2016, Germany has been participating in the **Open Government Partnership** (https://www.opengovpartnership.org/countries/germany. See also *Bundestag*-Printed Matter 19/4026, p. 8), whereby this young model, which focuses on opening up the state, aims to promote collaboration, participation, accountability and innovation through more transparency (*Bundestag*-Printed Matter 18/11614, p. 11).

On 13 July 2017, the inserted **§ 12a (1) sentence 1 EGovG Bund** came into force, according to which the authorities of the federal government shall make available for data retrieval via publicly accessible networks unprocessed data that they have collected for the fulfilment of their administrative tasks under public law or have been collected by third parties commissioned. The fact that the topic is approached cautiously at first becomes apparent in sentence 2, according to which no entitlement to the provision of such data is established. Insofar as the prerequisites for making the data available are met, this must be done in a machine-readable form and with metadata (§ 12a (5) sentence 1, 2 EGovG Bund). The metadata are posted on the **nation-wide metadata portal GovData** in accordance with § 12a (5) sentence 3 EGovG.

At union level, a **recast of the directive on the re-use of public sector information** has been adopted to fully exploit the potential of public sector information

[29] See also Guckelberger (2018), p. 361. For this reason, the Bavarian legislature has given preference to a different arrangement, critically Braun Binder (2016), p. 898. Under Art. 6 (4) sentence 3 BayEGovG, the administrative act is deemed to be notified on the third day after the electronic notification of the provision for retrieval has been sent to the person authorised to retrieve it. On the announcement of an examination assessment on the internet portal of a university without additional notification, BVerwG, NVwZ 2018, p. 498.

for the European economy and society (OJ L 172, 26 June 2019, p. 56 ff.). Significant improvements are expected to result, *inter alia*, from the provision of real-time access to dynamic data by adequate technical means (OJ L 172, 26 June 2019, p. 72), or the increased provision of high-value datasets free of charge (OJ L 172, 26 June 2019, pp. 67, 75 f.).

References

Arzt C (2017) Das neue Gesetz zur Fluggastdatenspeicherung – Einladung zur anlasslosen Rasterfahndung durch das BKA. DÖV:1023–1030

Barczak T (2020) Algorithmus als Arkanum. DÖV:997–1007

Berger A (2017) Digitales Vertrauen. Eine verfassungs- und verwaltungsrechtliche Perspektive. DVBl:804–808

Berger A (2018) Onlinezugangsgesetz und Digitalisierungsprogramm. Auf die Kommunen kommt es an! KommJur:441–445

Berger A (2019) Brauchen wir ein neues Verwaltungsrecht? – Funktionalität, Grundrechtsschutz, Teilhabe und Rechtsschutz im vollautomatisierten Verfahren. DVBl:1234–1237

Bertelsmann Stiftung (2017) Digitale Transformation der Verwaltung. https://www.bertelsmannstiftung.de/fileadmin/files/Projekte/Smart_Country/DigiTransVerw_2017_final.pdf. Last accessed on 25.01.2019

BMI (2018) Projekt zur Gesichtserkennung erfolgreich. https://www.bmi.bund.de/SharedDocs/pressemitteilungen/DE/2018/10/gesichtserkennung-suedkreuz.html. Last accessed on 25.01.2019

BMI (2019) Erhöhung der Sicherheit auf Bahnhöfen.https://www.bmi.bund.de/SharedDocs/pressemitteilungen/DE/2019/09/sicherheitauf-bahnhoefen.html. Last accessed on 07.10.2019

Boehme-Neßler V (2018) Das Ende der Demokratie?: Effekte der Digitalisierung aus rechtlicher, politologischer und psychologischer Sicht. Springer, Heidelberg/Berlin

Böhmann T (2018) Thesenblatt zum Workshop "Algorithmen und Verwaltungs(gerichtliche) Kontrolle" im Rahmen der Veranstaltung "Algorithmen und künstliche Intelligenz in der Verwaltung" in Hamburg

Braun Binder N (2016) Vollständig automatisierter Erlass eines Verwaltungsaktes und Bekanntgabe über Behördenportale – Anmerkung zu den §§ 24 Abs. 1 Satz 3, 35a und 41 Abs. 2a VwVfG. DÖV:891–898

Braun Binder N (2019) Vollautomatisierte Verwaltungsverfahren, vollautomatisiert erlassene Verwaltungsakte und elektronische Aktenführung. In: Seckelmann M (ed) Digitalisierte Verwaltung. Vernetztes E-Government, ch. 12, 2nd edn. Erich Schmidt, Berlin, pp 309–324

Britz G (2012) § 26 Elektronische Verwaltung. In: Hoffmann-Riem W, Schmidt-Aßmann E, Voßkuhle A (eds) Grundlagen des Verwaltungsrechts. Informationsordnung, Verwaltungsverfahren, Handlungsformen, vol II, 2nd edn. C.H. Beck, München

Bull HP (1964) Verwaltung durch Maschinen, 2nd edn. Grote, Köln/Berlin

Demaj L (2018) Smart Government: Die Verwaltung und der Staat der Zukunft denken. Informatik Spektrum:123–137

Denkhaus W (2019) Vom E-Government zur Digitalisierung. In: Seckelmann M (ed) Digitalisierte Verwaltung. Vernetztes E-Government, ch. 1, 2nd edn. Erich Schmidt, Berlin, pp 51–78

Deutscher Landkreistag (2018) Der digitale Landkreis – Herausforderungen, Strategien, Gute Beispiele, Schriften des Deutschen Landkreistages, vol 136. http://www.kreise.de/cms1/images/stories/publikationen/bd-136.pdf. Last accessed on 30.01.2019

Djeffal C (2017a) Das Internet der Dinge und die öffentliche Verwaltung – Auf dem Weg zum digitalisierten Smart Government. DVBl:808–816

Djeffal C (2017b) Leitlinien der Verwaltungsnovation und das Internet der Dinge. In: Klafki A, Würkert F, Winter T (eds) Digitalisierung und Recht – Tagung des eingetragenen Vereins Junge Wissenschaft im Öffentlichen Recht an der Bucerius Law School, 26. November 2016. Schriften der Bucerius Law School, vol 2/31:83–112
Djeffal C (2018) Normative Leitlinien für künstliche Intelligenz in Regierung und öffentlicher Verwaltung. In: Mohabbat Kar R, Thapa B, Parycek P (eds) (Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft. ÖFIT, Berlin, pp 493–515
Edenharter A (2020) Die Digitalisierung als Herausforderung für das Baurecht. VerwArch 111: 341–371
Ehlers D (1991) Die öffentliche Verwaltung im technischen Zeitalter. JURA:337–342
Eifert M (2006) Electronic Government. Nomos, Baden-Baden
Eifert M (2014) Elektronische Verwaltung – von der Verwaltungsreform zum Verwaltungsreformrecht. In: Bultmann PF, Grigoleit KJ, Gusy C, Kersten J, Otto CW, Preschel C (eds) Allgemeines Verwaltungsrecht – Institute, Kontexte, System, Festschrift für Ulrich Battis zum 70. Geburtstag. C.H.Beck, München, pp 421–435
European Commission (2018) Digital Economy and Society Index 2018 Report Country Reporting Germany, 14. May 2018. http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=52214. Last accessed on 25.01.2019
European Commission (2018) eGovernment Benchmark Securing eGovernment for all. Insight Report athttps://digitalstrategy.ec.europa.eu/en/news/egovernment-benchmark-2018-digital-efforts-european-countries-are-visibly-paying. Last accessed on 04.01.2022
Fischer S, Petersen T (2018) Was Deutschland über Algorithmen weiß und denkt. Ergebnisse einer repräsentativen Bevölkerungsumfrage. Bertelsmann Stiftung, Gütersloh
Fortiss, Initiative D21 (2018) eGovernment Monitor 2018. https://www.egovernment-monitor.de/fileadmin/uploads/user_upload/studien/PDFs/191029_eGovMon2018_Final_WEB.pdf. Last accessed on 30.01.2019
Gigerenzer G (2018) Millionenfach unnötige Personenkontrollen durch Gesichtserkennung?, Kontraste/ARD. http://mediathek.daserste.de/Kontraste/Millionenfach-unn%C3%B6tige-Personenkontrolle/Video?bcastId=431796&documentId=57987284. Last accessed on 25.01.2019
Gore A (1993) From red tape to results, creating a government that works better & costs less. Report of the National Performance Review, September 7, Times Books, New York
Gröpl C (2018) Art. 91c. In: Maunz T, Dürig G (eds) Kommentar zum Grundgesetz 84. EL, 6. vol, C.H.Beck, München
Groß T (2004) Die Informatisierung der Verwaltung – Eine Zwischenbilanz auf dem Weg von der Verwaltungsautomation zum E-Government. VerwArch 95:400–417
Guckelberger A (2009) Der Übergang zur (ausschließlich) elektronischen Gesetzesverkündung. Nomos, Baden-Baden
Guckelberger A (2014) Umweltinformationen unter europäischem Einfluss. VerwArch:411–440
Guckelberger A (2016) Vollständig automatisierte Verwaltungsakte. In: van Oostrom S, Weth S (eds) Festschrift für Maximilian Herberger zum 70. Geburtstag. juris, Saarbrücken, pp 397–414
Guckelberger A (2018) Die Digitalisierung und ihre Folgen für die postalische Bekanntgabe von Verwaltungsakten. NVwZ:359–363
Guckelberger A (2019a) E-Government: Ein Paradigmenwechsel in Verwaltung und Verwaltungsrecht? VVDStRL:235–288. de Gruyter, Berlin
Guckelberger A (2019b) Öffentliche Verwaltung im Zeitalter der Digitalisierung. Nomos, Baden-Baden
Guckelberger A (2019c) Digitale Verwaltungsdienste für die Wirtschaft. GewArch:457–464
Guggenberger L (2019) Einsatz künstlicher Intelligenz in der Verwaltung. NVwZ:844–850
Han M (2019) Intelligentes E-Government in Südkorea – zur öffentlichen Dienstleistungsinnovation. DÖV:543–551
Heckmann D (2018) jurisPK-Internetrecht: Telemediengesetz, E-Commerce, E-Government, 5th edn. juris, Saarbrücken

Herrmann M, Stöber K (2017) Das Onlinezugangsgesetz des Bundes. Wie der Gang zum Amt überflüssig werden soll. NVwZ:1401–1407
Hill H (2018) Was bedeutet Künstliche Intelligenz (KI) für die Öffentliche Verwaltung? VM:287–294
Hoffmann-Riem W (2018) Rechtliche Rahmenbedingungen für und regulative Herausforderungen durch Big Data. In: Hoffmann-Riem W (ed) Big Data – Regulative Herausforderungen. Nomos, Baden-Baden, pp 11–78
Hogrebe F, Kruse W (2014) Deutschland 4.0: Industrie, Verwaltung, Standort, Wohlstand. Verlag für Verwaltungswissenschaft, Frankfurt am Main
Höhl S (2018) Freiburger Feuerwehr testet ein System für die grüne Welle, Badische Zeitung. http://www.badische-zeitung.de/freiburg/freiburger-feuerwehr-testet-ein-system-fuer-die-gruene-welle%2D%2D149614211.html. Last accessed on 25.01.2019
Hölzel J (2017) Vom E-Government zum Smart Government. DVBl:1015–1018
Hufen F, Siegel T (2018) Fehler im Verwaltungsverfahren, 6th edn. Nomos, Baden-Baden
Initiative D21, TU München (2020) eGovernment Monitor 2020.https://initiatived21.de/app/uploads/2020/10/egovernment_monitor_2020_onlineausgabe.pdf. Last accessed on 04.01.2022
Kagermann H, Lukas WD, Wahlster W (2011) Industrie 4.0: Mit dem Internet der Dinge auf den Weg zur 4. industriellen Revolution, VDI Nachrichten v. 1. April 2011, Nr. 12. https://www.vdi-nachrichten.com/Technik-Gesellschaft/Industrie-40-Mit-Internet-Dinge-Weg-4-industriellen-Revolution. Last accessed on 22.01.2019
Kaiser AB (2009) Intelligente Verwaltungsmaschine – Intelligente Maschinen in der Verwaltung: Die Diskussion über Verwaltungsautomation in den 1950er- und 1960er-Jahren. In: Collin P, Lutterbeck KG (eds) Eine intelligente Maschine? Handlungsorientierungen moderner Verwaltungen (19./20. Jh.). Nomos, Baden-Baden, pp 233–243
Kühling J, Sackmann F (2018) Datenschutzordnung 2018 – nach der Reform ist vor der Reform?! NVwZ:681–686
Langenbach P (2017) Der Anhörungseffekt. Mohr Siebeck, Tübingen
Lazaratos P (1990) Rechtliche Auswirkungen der Verwaltungsautomation auf das Verwaltungsverfahren. Duncker & Humblot, Berlin
Löbel S, Schuppan T, Dozenko C (2018) Ein Blick in die Praxis: Akzeptanz der eAkte im Bereich SGB II. VM:299–306
Luhmann N (1966) Recht und Automation in der öffentlichen Verwaltung. Eine verwaltungswissenschaftliche Untersuchung, 2nd edn. Duncker & Humblot, Berlin
Martini M, Wiesner C (2017) Art. 91c Abs. 5 GG und das neue Zugangsregime zur digitalen Verwaltung – Quantensprung oder zu kurz gesprungen? ZG, pp 193–227
Martini M, Nink D (2018) Subsumtionsautomaten ante portas? – Zu den Grenzen der Automatisierung in verwaltungsrechtlichen (Rechtsbehelfs-)Verfahren. DVBl:1128–1138
Martini M, Wiesner C (2017) Art. 91c Abs. 5 GG und das neue Zugangsregime zur digitalen Verwaltung - Quantensprung oder zu kurz gesprungen? ZG:193-227
Martini M, Weinzierl Q (2017) Die Blockchain-Technologie und das Recht auf Vergessenwerden – Zum Dilemma zwischen Nicht-Vergessen-Können und Vergessen-Müssen. NVwZ:1251–1259
Maurer H, Waldhoff C (2017) Allgemeines Verwaltungsrecht, 19th edn. C.H.Beck, München
McKinsey&Company, bitkom (2018) Smart Government, Wie die öffentliche Verwaltung Daten intelligent nutzen kann. https://www.mckinsey.de/~/media/mckinsey/locations/europe%20and%20middle%20east/deutschland/news/presse/2018/2018-11-21-smart%20government/smart%20government_de.ashx. Last accessed on 24.01.2019
Müller-Franken S (2018) Grundlagen des Besteuerungsverfahrens. StuW:113–122
NKR (2016) 10 Jahre NKR – gute Bilanz bei Bürokratieabbau und Folgekostenbegrenzung – alarmierender Rückstand bei E Government.https://www.normenkontrollrat.bund.de/resource/blob/267760/444052/cdf2763c2be98d3fd61b79d744645144/2016-09-21-nkr-jahresbericht-2016-data.pdf?download=1. Last accessed on 04.01.2022
NKR (2018) Deutschland: weniger Bürokratie, mehr Digitalisierung, bessere Gesetze Einfach machen!https://www.normenkontrollrat.bund.de/resource/blob/267760/1536236/1bed933ea00

6098d6807ab48bd3c8574/2018-10-10-download-nkr-jahresbericht-2018-data.pdf?download=1. Last accessed on 04.01.2022

NKR (2020) Krise als Weckruf: Verwaltung modernisieren, Digitalisierungsschub nutzen, Gesetze praxistauglich machen.https://www.normenkontrollrat.bund.de/resource/blob/418738/1800428/44bc6f69bc0256967097282af768a05e/20201021-nkr-jahresbericht-2020-data.pdf?download=1. Last accessed on 04.01.2022

Paal BP, Hennemann M (2017) Big Data im Recht – Wettbewerbs- und daten(schutz)rechtliche Herausforderungen. NJW:1697–1701

Polomski R-M (1993) Der automatisierte Verwaltungsakt. Die Verwaltung an der Schwelle von der Automation zur Informations- und Kommunikationstechnik. Duncker & Humblot, Berlin

Rademacher T (2017) Predictive Policing im deutschen Polizeirecht. AöR 142:366–416

Reinermann H, von Lucke J (2002) Electronic Government in Deutschland – Ziele, Stand, Barrieren, Beispiele, Umsetzung. Forschungsinstitut für öffentliche Verwaltung, 2nd edn., Speyer

Richter E (2019) Die digitale Stadt – Chancen und Herausforderungen der Digitalisierung für die lokale Gemeinschaft. In: Seckelmann M (ed) Digitalisierte Verwaltung. Vernetztes E-Government, ch. 10, 2nd edn. Erich Schmidt, Berlin, pp 263–279

Roth-Isigkeit D (2020) Die Begründung des vollständig automatisierten Verwaltungsakts. DÖV:1018–1026

Ruge K (2018) Art. 91c GG. In: Schmidt-Bleibtreu B, Hofmann H, Henneke H-G (eds) Kommentar zum Grundgesetz, 14th edn. Wolters Kluwer, Köln

Scheer AW, Kruppke H, Heib R (2003) E-Government – Prozessoptimierung in der öffentlichen Verwaltung. Springer, Berlin/Heidelberg

Schliesky U (2009) Legitimation und Verantwortung im komplexen, arbeitsteiligen Staat – eine Einführung. In: Hill H, Schliesky U (eds) Herausforderung e-Government – E-Volution des Rechts- und Verwaltungssystems. Nomos, Baden-Baden, pp 11–38

Schliesky U, Hoffmann C (2018) Die Digitalisierung des Föderalismus. Der Portalverbund gemäß Art. 91c Abs. 5 GG als Rettung des E-Government? DÖV:193–198

Schrey J, Thalhofer T (2017) Rechtliche Aspekte der Blockchain. NJW:1431–1436

Schuppan T, Köhl S (2016) Verwaltung 4.0: Modernisierungsrelevant oder alter Wein in neuen Schläuchen? VM:27–33

Seckelmann M (2018) Art. 91c GG. In: Friauf KH, Höfling W (eds) Berliner Kommentar zum Grundgesetz, Lfg. 3/18. Erich Schmidt, Berlin

Seckelmann M (ed) (2019) Einsatz bei der Polizei: Twitter-Nutzung, Online-Streifen, Trojaner, Facebook-Fahndung, Biometriesoftware, (intelligente) Videoüberwachung, Predictive Policing, Body- Cams und Fotodrohnen. In: Digitalisierte Verwaltung. Vernetztes E-Government, ch. 22, 2nd edn. Erich Schmidt, Berlin, pp 483–506

Siegel T (2014) Der virtuelle Verwaltungsakt. VerwArch 105:241–261

Siegel T (2017) Automatisierung des Verwaltungsverfahren zugleich eine Anmerkung zu §§ 35a, 24 I 3, 41 IIa VwVfG. DVBl:24–28

Siegel T (2018) Auf dem Weg zum Portalverbund. Das neue Onlinezugangsgesetz (OZG). DÖV:185–192

Siegel T (2019) Der Europäische Portalverbund - Frischer Digitalisierungswind durch das einheitliche digitale Zugangstor ("Single Digital Gateway"). NVwZ:900–909

Singelnstein T (2018) Predictive Policing: Algorithmenbasierte Straftatprognosen zur vorausschauenden Kriminalintervention. NStZ:1–9

Skrobotz J (2005) Das elektronische Verwaltungsverfahren – Die elektronische Signatur im E-Government. Duncker und Humblot, Berlin

Smith B (2018) Facial recognition technology: the need for public regulation and corporate responsibility, Microsoft On the Issues. https://blogs.microsoft.com/on-the-issues/2018/07/13/facial-recognition-technology-the-need-for-public-regulation-and-corporate-responsibility/. Last accessed on 25.01.2019

Stelkens U (2018) § 35a VwVfG. In: Stelkens P, Bonk HJ, Sachs M (eds) Verwaltungsverfahrensgesetz-Kommentar, 9th edn. C.H.Beck, München
Stocksmeier D, Hunnius S (2018) OZG-Umsetzungskatalog. Digitale Verwaltungsleistungen im Sinne des Onlinezugangsgesetzes. BMI, Berlin
Thapa B, Parycek P (2018) Data Analytics in Politik und Verwaltung. In: Mohabbat Kar R, Thapa B, Parycek P (eds) (Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft. ÖFIT, Berlin, pp 40–75
Trute H-H (2018) Rechtliche Herausforderungen der Digitalisierung. In: Bär C, Grädler T, Mayr R (Hrsg) Digitalisierung im Spannungsfeld von Politik, Wirtschaft, Wissenschaft und Recht, 2. vol.: Wissenschaft und Recht. Springer, Heidelberg/Berlin, pp 313–330
Veil W (2018) Die Datenschutz-Grundverordnung: des Kaisers neue Kleider – Der gefährliche Irrweg des alten wie des neuen Datenschutzrechts. NVwZ:686–696
Von Lucke J (2016) Auf dem Weg zum Smart Government: Gestaltung eines intelligent vernetzten Regierungs- und Verwaltungshandelns, Society Byte. https://www.societybyte.swiss/2016/11/02/auf-dem-weg-zum-smart-government-gestaltung-eines-intelligent-vernetzten-regierungs-und-verwaltungshandelns/. Last accessed on 25.01.2019
Von Lucke J (2019) Smart government. In: Seckelmann M (Hrsg) Digitalisierte Verwaltung. Vernetztes E-Government, ch. 2, 2nd edn. Erich Schmidt, Berlin, pp 81–106
Voßkuhle A (ed) (2012) § 1 Neue Verwaltungsrechtswissenschaft. In: Hoffmann-Riem W, Schmidt-Aßmann E (Hrsg) Grundlagen des Verwaltungsrechts, Methoden, Maßstäbe, Aufgaben, Organisation, 1. vol, 2nd edn. C.H. Beck, München
Wendt R (2018) Das Recht auf Offenlegung der Messunterlagen im Bußgeldverfahren. NZV:441–446
Wikipedia, Artikel “Industrie 4.0” (2019). https://de.wikipedia.org/wiki/Industrie_4.0. Last accessed on 25.01.2019
Windoffer A (2018) Herausforderungen der Digitalisierung aus der Perspektive der öffentlichen Verwaltung. In: Bär C, Grädler T, Mayr R (eds) Digitalisierung im Spannungsfeld von Politik, Wirtschaft, Wissenschaft und Recht, 2. vol.: Wissenschaft und Recht. Springer, Heidelberg/Berlin, pp 363–376
Wischmeyer T (2018) Regulierung intelligenter Systeme. AöR 143:1–66
Zenner F (2018) Algorithmenbasierte Straftatprognosen in der Eingriffsverwaltung – Zu den verfassungsrechtlichen Grenzen und einfach-gesetzlichen Möglichkeiten von “Predicitve Policing”. In: Wieczorek MA (ed) Digitalisierung – Rechtsfragen rund um die digitale Transformation der Gesellschaft, Tagungsband Liberale Rechtstagung. Cuvillier, Göttingen, pp 117–133
Ziekow J (2018) Das Verwaltungsverfahrensrecht in der Digitalisierung der Verwaltung. NVwZ:1169–1172

Agricultural Law 4.0: Digital Revolution in Agriculture

Ines Härtel

Contents

1 Digitisation of Agriculture

The digital transformation of the economy and society is also changing agriculture at an above-average rate and intensity. It is true that agriculture has already been able to achieve considerable increases in yields in recent decades, also thanks to high levels of mechanisation, and thus guarantee food security. The high productivity is reflected in the fact that one farmer fed around 4 people in 1900, around 10 in 1949 and around 155 today (with less land)—with an increasing population. The digitisation of agriculture continues this growing productivity process on an

I. Härtel (✉)
European University Viadrina, Frankfurt (Oder), Germany
e-mail: IHaertel@europa-uni.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_17

expanded, innovative technological basis and at the same time is profoundly reshaping the agricultural and food sector compared to the past. More differentiated, effective and targeted agricultural production, increased quality and higher yields of food products in the animal and plant sector, reduced environmental impact and more economical use of resources in the sense of sustainable development, increased operational efficiency in the areas of time, costs and management, economic and social security for farmers, as well as improved transparency for trade and consumers are goals that are advanced in a differentiated manner through digitisation. The saturation of third-order information and communication technologies in the agricultural sector is successively leading to a new kind of combination of farming experience, agricultural entrepreneurship and digital technological competence in the comprehensive 'infosphere' as an advanced agricultural 'onlife'—the offline and online worlds that have grown together indistinguishably (see Floridi 2015, p. 67). Farming rules and algorithms, human intelligence and artificial intelligence, the eye of the knowledgeable farmer and the outsourced 'eye' via GPS, science-based and practical experience—all this will form a new synthesis, so to speak, in a future integrated digital food economy—encompassing the agricultural and food sectors. Agriculture is already one of the most comprehensively digitised sectors, although smart digital farming is still dominated by applications in specific areas and individual cases. The entire range of different digital technologies is used in the agricultural sector, which are in constant development (becaming, permanent updates). These include, above all, computing/PC, online technologies/Internet, satellite technology/GPS, data mining and Big Data, cloud technology, artificial intelligence (AI) approaches, robotics, sensor technologies, drones (unmanned aerial vehicles), augmented reality (AR) and virtual reality (VR), blockchain, mobile payment, mobile technologies such as smartphones and tablets,[1] but also apps, tracking systems, scoring, chatbots and social web connections. These digital technologies of various orientations, characteristics and degrees of maturity are used in different agricultural production segments, are being tested or are available as prototypes. The most important ones are currently:

Satellite technology: GPS data (geodata, weather data, etc.) are used for precise mapping, for the precise and (partially) autonomous driving of tractors (navigation) and for soil quality surveys, which enable the precise, pinpoint application of seeds, fertilisers or pesticides.

Sensor technologies: Sensor-based (real-time capable) digital systems record soil values, plant situation (nutrient supply, pest occurrence, disease infestation) and harvest maturity in the plant sector, and feed intake and animal diseases in the animal sector.

Artificial intelligence/robotics: AI-supported field robots for automated-mechanical control of weeds, robot use in grape harvesting or fruit harvesting

[1] These digital technologies are complemented and expanded by a range of technological innovations in other areas, such as biotechnologies (using various techniques such as gene editing, CRISPR/Cas), nanotechnologies, material sciences, but also 3-D printing.

(with various suction or gripping techniques in apple harvesting), robot systems in animal feeding and cow milking including milk quality recording.

Blockchain : This digital technology is used to secure transactions and contracts, enable traceability back to the farmer and provide trade finance for soybeans, for example (such as the US agri-food company Cargill).

Assisted/(semi-)autonomous driving: Tractors with digital assistance systems and semi-autonomous (GPS-assisted) driving capabilities are now widespread; the first prototypes of fully automated tractors are ready for testing. Here, both the problems of normal road traffic in rural regions and the special requirements in fields have to be managed digitally.

Drones: Agricultural drones can carry out mapping, apply pesticides and record plant populations. One in ten farms already uses agricultural drones.

Agribusiness: This includes agronomic-farm management systems, especially data management and process control, automated records (digital record systems) and evaluations, service-oriented apps, sharing systems and online exchanges, especially also in the context of the platform economy (in Germany alone, there are over 50 platforms in the agri-food sector), digitally controlled warehousing and logistics, market information systems, online marketing, digital stakeholder and customer/consumer integration—also in the sense of a comprehensive, transparent value chain ('digital product passport' in the agri-food sector).

Digital technologies are often used in combination, e.g. GPS data in conjunction with sensor data, Big Data applications and farm management controls (smart farming). Another example is digitally optimised and integrated potato production ('smart potato') using data analysis, sensor technologies, AI—from the field to factory production and the food trade. In the process, all data accumulating in the entire value chain is recorded and integrated: GPS, weather, soil and fertiliser data, planting data, farm machinery data/harvesters, harvest data, conveyor track data, logistics data, factory data, market, price forecast data and financial/investor data etc. All in all, the use of digital, algorithmic technologies always produces oversized amounts of different data that require new handling in analytics, evaluation innovation and data protection.

For the farmer, the use of digital technologies must be economically profitable overall. Therefore, an adequate digital infrastructure is also necessary—understood in the sense of a **digital provision of agricultural services**. The precondition for further successful digitisation is the expansion of fibre optic networks and LTE or 4G/5G mobile networks, especially in rural, agriculturally used regions.[2] This will not only advance precision agriculture, smart farming and digitised individual areas, but also enable greater automation and overarching production systems (so-called cyber-physical systems) in the sense of an agricultural IoT in the future. This could

[2] In addition to infrastructure problems, there are other problems such as lack of standards, lack of interfaces, incompatible devices, insufficient provision of public data/open data management, dealing with the effects of the platform economy, uncertainties regarding company and personal data protection, problems of cyber security.

be linked to both the platform economy and the branched value chains. Together with the upstream sectors (e.g. fertiliser, crop protection and feed businesses, conventional/genetic engineering plant breeding companies, agricultural machinery manufacturers, companies in the IT sector from software developers to platform operators) and the downstream sectors (e.g. processing and refining companies in the food sector, wholesale and retail trade, consumer sector), individual farmers and agricultural companies (partnerships and corporations) are forming an independent digital economic ecosystem through the use of various digital technologies, through digital networking (networks, platforms) and through integrated digital value chains: Agriculture 4.0 as a future comprehensive target system. At the same time, the training-related, political, socio-cultural and legal shaping forces must grow along with the pervasive digital-technological-economic developments. Above all, the law, with its regulatory competence (as hard law as well as soft law), has an enabling and at the same time protective task, which is complemented by legal ethics that stimulates public discourse. The presentation of this field of digital law conditions is in the foreground in the following. Based on the normative framework, the legal focus is primarily on data governance (data protection, data ownership, code of conduct, open data); in addition, aspects of artificial intelligence, agricultural drones/ air traffic law and cyber security are included.

2 The Normative Framework: Right to Food and Sustainable Development Goals

Digitisation as a technology is not an end in itself, but should serve the realisation of normative goals. Thus, the digitisation of the agri-food sector along its value chains contributes to the effectiveness of the right to food and sustainable development. The **human right to food** is enshrined in Article 25 (1) of the Universal Declaration of Human Rights (UDHR) of 10 December 1948 and Article 11 of the International Covenant on Economic, Social and Cultural Rights (UN Covenant) of 1966. It includes a right to food security, which is fulfilled 'when all people at all times have physical and economic access to adequate, safe and nutritious food to meet their dietary needs and food preferences for an active and healthy life'.[3] The right to food is implicitly protected by the fundamental right to life and physical integrity under Article 2 (2) sentence 1 of the Basic Law; the state's duty to protect is particularly relevant here (see Härtel 2015, p. 29 ff.). Likewise, it can be implicitly derived from the right to life under Article 2 (2) of the Charter of Fundamental Rights of the EU (CFR) and the right to physical integrity under Article 3 (1) CFR.

[3] World Food Summit 1996 - 'Rome Declaration on World Food Security', http://www.fao.org/docrep/003//w361e/w3613e00; European Parliament Resolution of 18 January 2011 on the recognition of agriculture as a sector of strategic importance for food security (2010/2112 (INI)), OJ 2012 C 136 E/8 (at R. 4.).

On food safety (Härtel and Yu 2018, p. 72 ff.), the EU and Germany have fulfilled their obligations to protect through detailed (composite) legislation.

In its '2030 Agenda for Sustainable Development', the United Nations has formulated Goal 2 of the **Sustainable Development Goals** (SDGs) as 'End hunger, achieve food security and improve nutrition and promote sustainable agriculture'. Further sub-goals give more concrete form to these goals. The communiqué of the G20 heads of state and government at the summit in Antalya (15–16 November 2015) also refers to the 2030 Agenda and contains an action plan on food security and sustainable food systems. The way in which food is produced, consumed and sold should be economically, socially and environmentally sustainable in the sense of the three dimensions (cf. paragraph 20 of the communiqué), with an increasingly integrated approach. On SDG 2, the G20 commits to reducing food loss and food waste worldwide. Further details on the G20 communiqué on food security can be found in the 'Communiqué of the G20 Agriculture Ministers' of 7/8 May 2015 and in the 'Implementation Plan of the G20 Food Security and Nutrition Framework'.

The communiqué 2019 '**Digital Agriculture - Smart Solutions for Future Agriculture**'[4] adopted by the Global Forum for Food and Agriculture (GFFA) is also worthy of special mention; agriculture ministers from 74 nations have signed this communiqué. The aim is to use digitisation to make agriculture even more efficient and sustainable and to improve life in the countryside. To this end, the development of appropriate and scalable digital solutions in agriculture that are appropriate to the location and situation is to be accelerated. Globally, the necessary '**digital infrastructure**' for farmers is to be created and its expansion accelerated. Support should be given to cooperative models and cooperatives in the implementation of digitisation in agriculture. Research and education in the field of Agriculture 4.0 will be promoted to make the value chain more efficient and sustainable. Digital solutions will also be used to provide farmers with adequate information and improved market access (including to e-markets in the agri-food sector). Furthermore, in improving data use and ensuring data security and **data sovereignty** in the interest of farmers, nine important goals are formulated in terms of a sustainable digital agricultural economy:

1. Develop international solutions with agricultural stakeholders to reduce global differences in regulations on data collection, data security and data use and to develop standards.
2. Enabling effective use of digitally captured data.
3. Increase the interoperability of digital systems to improve the possibilities for data exchange, data use and data analysis by farmers, science, business and politics.
4. Avoiding farmers' dependence on individual digital systems and protecting users' intellectual property and personal rights in digital innovations and information.

[4] The GFFA is an international conference on key issues in global agriculture and food that has been held annually since 2009 as part of the International Week in Berlin.

5. Develop trust and transparency in data governance principles, including rules for authorisation and supervision in data collection and use, and promote data use models that allow farmers to make their own decisions about the sharing of their farm, machinery and business data, taking into account national regulations.
6. Provision of public data using appropriate platforms.
7. Promote digital solutions to strengthen the transparency , efficiency and integrity of supply chains and effectively combat counterfeiting, fraud and smuggling.
8. Promote international digital data infrastructures to strengthen cross-border animal and plant disease control.
9. Establishment of digital methods at the World Organisation for Animal Health (OIE) as part of the modernisation of the OIE WAHIS (World Animal Health Information System) reporting system.

A useful institutional support at the global level would be an International Digital Council for Food and Agriculture, which would advise governments, promote the exchange of ideas and experiences and help to make the opportunities of digitisation more accessible to all.

Although these current declarations are merely **soft law** (on the legal and practical significance of soft law, see Monien 2014, pp. 790–792), they are groundbreaking for the design and implementation of the agricultural policies of the states and the EU in political, legal and practical terms; this applies to the responsibility of the sovereigns within the framework of their internal and external agricultural policies.

The EU secondary legislation on the Common Agricultural Policy (CAP) will also explicitly refer to digitisation. According to this, there will be a **new cross-cutting objective on the digitisation of agriculture**. The agricultural sector is to be modernised through financial support for digitisation in agriculture and rural areas.[5] This cross-cutting objective has an impact on all agricultural objective areas (food security, ecological resource protection, strengthening of rural areas). The application system for agricultural support will also be further digitalised, e.g. through a geodata-based application by the farmer and a geodata-based application system by the agricultural administrations.[6]

The financial support of digital agricultural technology can accelerate the progress of a sustainable agricultural and food economy, but requires the farmer to use the agricultural technology in a legally compliant manner.

[5] Article 5 lit. b (Regulation of the European Parliament and of the Council of 2 December 2021 establishing rules on support for strategic plans to be drawn up by Member States under the Common Agricultural Policy (CAP Strategic Plans) and financed by the European Agricultural Guarantee Fund (EAGF) and by the European Agricultural Fund for Rural Development (EAFRD), and repealing Regulations (EU) No. 1305/2013 and (EU) No. 1307/2013.

[6] Cf. Article 66 para. 1 lit. b Regulation of the European Parliament and of the Council of 2 December 2021 on the financing, management and monitoring of the Common Agricultural Policy and repealing Regulation (EU) No. 1306/2013 (‘Horizontal CAP Regulation’).

3 Artificial Intelligence: AI Liability Law

For AI systems with their advantages for sustainable agriculture to be increasingly used in the future, digital technology must be accepted by farmers in this area. For this, it is necessary that the law provides adequate solutions for AI-specific liability risks in agriculture (for agriculture-relevant liability risks with examples and for possible AI legal solutions, see Härtel 2020b, pp. 439, 441–453). The existing Product Safety Law and Product Liability Law must be further developed with regard to AI specifics. The specific characteristics of AI include: complexity (multiplicity of actors in digital ecosystems, multiplicity of digital components—hardware, software, services), opacity (opacity of processes, black box due to self-learning, more difficult to predict the behaviour of an AI-supported product), openness (updates or improvements/upgrades, for this interaction with other systems or data sources), autonomy (tasks with less human control, modification of algorithms), data dependency (dependence on external information that is not pre-installed but generated by built-in sensors or communicated externally), vulnerability (particular susceptibility to cybersecurity breaches due to the openness and complexity of digital ecosystems). Inherent in these AI specifics are also AI-specific liability risks in agriculture. If personal injury, property damage or environmental damage is caused by the use of AI systems, the question of who is liable for what arises for farmers but also for other AI actors or third parties involved. AI damage claims can be considered, for example, if an AI-controlled harvesting robot picks strawberries that are still unripe and not yet marketable, if farm animals such as chickens or pigs are injured during automated feeding, if an AI-controlled animal health management system fails to detect animal diseases or administers the wrong medication, if a robot spreads fertiliser or pesticides without adhering to the water body margin or exceeding the application limits, or if a cyberattack causes the entire (future) farm 4.0 collapses.

On the development of an AI legal framework, the European Commission's White Paper 'On Artificial Intelligence - A European Approach to Excellence and Trust' (COM(2020) 65 final), the Report of the Expert Group on Liability and New Technolgies 'Liability and Artificial Intelligence' (2019/2020), the Report on the Impact of Artificial Intelligence, the Internet of Things and Robotics on Security and Liability issued by the European Commission (COM(2020) 64 final) and the Resolution of the European Parliament of 20 October 2020 a regulation of civil liability in the use of artificial intelligence (2020/2014(INL)) as well as the 'Report of the Enquete Commission on Artificial Intelligence - Social Responsibility and Economic, Social and Ecological Potentials' (BTDrs. 19/2300, p. 74 ff.) can be considered as relevant. Meanwhile the EU drafted an Artificial Intelligence Act (COM 2021 (206) final) that intends to establish a differentiated regulatory regime for AI.

4 The Legal Framework for the Use of (Agricultural) Drones

Around one in ten farms in Germany already uses agricultural drones. The practical use of these digitalised (agricultural) drones is quite versatile. For example, drones are used for precise field measurement/mapping of agricultural areas or for surveying the soil condition or the condition of plant stands. Flying over meadows before the grass harvest with the help of a thermal imaging camera can be carried out to locate fawns and protect them from combine harvesters. Drones can also be used for the application of plant protection products.

However, the use of drones, with its benefits for sustainable agriculture, can also endanger the legal interests of the general public or individuals. For this reason, Union law lays down requirements for the authorisation and operation of drones. This is supplemented by legal provisions in German law.

A number of general EU regulations in this area also apply to agricultural drones. Secondary legislation for drones is set out in 'Regulation (EU) 2018/1139 of the European Parliament and of the Council of 4 July 2018 on common rules in the field of civil aviation and establishing a European Union Aviation Safety Agency (...)' and amending other acts (so-called Civil Aviation Regulation, OJ L 212/1 of 22.8.2018). The legal term for drones is '**unmanned aircraft**', which is legally defined as an aircraft that is operated or designed to be operated autonomously or remotely without a pilot on board (cf. Article 3 No. 3 Regulation 2018/1139). The Civil Aviation Regulation pursues the main objective of a high uniform level of aviation safety (cf. Article 1 para. 1). Among the central principles for measures under this EU regulation are the '**risk-based approach**' (the higher the risk posed by a particular type of operation to an asset to be protected, the higher the requirements for the certification and operation of drones) and the associated proportionality (Article 4 para. 2). Measures must also be based on the 'best available evidence and analysis' and must consider 'cyber security' (Article 4 para. lit. b, d).

The special regulations on unmanned aircraft are found in Articles 55 to 58 and in Annex IX of the Civil Aviation Ordinance. They represent a right to avert or prevent risks. The risks typically arising or to be considered with drone operations are to be averted. Risk prevention protects in particular the right to respect for private and family life under Article 7 CFR, the right to protection of personal data under Article 8 CFR, but also the environment under Article 37 CFR, aviation safety and general public safety (cf. recitals 28 and 31, Annex IX No. 1.3. Regulation 2018/113). In this respect, there is, among other things, a compulsory registration for unmanned aircraft (Annex IX No. 4 Regulation 2018/113). Further details are set out in Delegated Regulation (EU) 2019/945 (OJ 152/1 of 11.6.2019) and Implementing Regulation (EU) 2019/947 (OJ 152/45 of 11.6.2019).

With the Civil Aviation Regulation, the Union legislator has not established a fully exhaustive regulation in the field of drones. It clarifies that Member States may adopt national provisions 'to impose certain conditions on the operation of unmanned aircraft for reasons outside the scope of this Regulation, including public security or the protection of privacy and personal data under Union law' (Article 56(8) of Regulation 2018/1139).

In Germany, the legal standards relevant to (agricultural) drones are anchored in the Air Traffic Act (Luftverkehrsgesetz – LuftVG in the version of the announcement of 10.5.2007 (BGBl. I, p. 698), last amended by Article 2 Para. 11 of the Act of 20.7.2017 (BGBl. I, p. 2808)) and in the associated legal ordinances (Air Traffic Licensing Ordinance – Luftverkehrs-Zulassungs-Ordnung – LuftVZO (19.6.1964 (BGBl. I, p. 370), last amended by Article 1 of the Ordinance of 30.3.2017 (BGBl. I, p. 683)) and Air Traffic Regulations (v. 29.10.2015 (BGBl. I, p. 1894), last amended by Article 2 of the Ordinance of 11.6.2017 (BGBl. I, p. 1617)). The two legal ordinances were replaced by the 'Ordinance of the Federal Ministry of Transport, Building and Urban Affairs of 30.3.2017 regulating the operation of unmanned aerial vehicles' (BGBl. I, p. 683) drone-specific legal provisions.

As with the operation of a manned aircraft, strict **liability** applies to the operation of a drone. If someone is killed or injured or an object is damaged in an accident while operating an agricultural drone, the keeper of the drone is liable to pay compensation under Section 33, paragraph 1, sentence 1 LuftVG. Under Section 43 para. 2 p. 1 LuftVG, he must maintain appropriate liability insurance to cover his liability. Also, to be able to identify the party responsible in the event of accidental damage, the owner of the drone—e.g. the farmer—must (in the case of a take-off mass of more than 0.25 kg) affix his name and address visibly and in durable as well as fireproof lettering (Section 19 para. 4 LuftVZO, introduced by the Ordinance on the Regulation of the Operation of Unmanned Aerial Vehicles of 30.3.2017, BGBl. I, p. 683). The inconsistency with the insurance obligation, which does not apply to the owner but to the keeper, is to be criticised; however, both do not have to be identical persons (Kämper and Müller 2017, p. 405). A '**drone licence**' (proof of knowledge valid for 10 years) is required for the operation of a drone with a take-off mass of more than 2 kg (cf. Section 21a para. 4 LuftVZO). An 'operating permit' is required, among other things, for a drone with a take-off mass of more than 5 kg. To protect against hazards, there are several **operating prohibitions** (Section 21b para. 1 nos. 1–11 LuftVO). Among other things, the operation of drones is generally prohibited out of sight of the pilot, in and over sensitive areas (accident sites, hospitals, industrial plants, prisons, crowds of people, over and around airfields), over residential properties and at heights above 10 m. The general ban on flying over nature conservation areas (exception if permitted by state law) is controversial in its application (Schrader 2017, p. 381 f.).

A general ban on operation applies to drones with a take-off mass of more than 25 kg. The competent authority may grant an exemption for operation for agricultural or forestry purposes (Section 21b para. 2 sentence 2 LuftVO).

Another specific technical issue concerns the application of **plant protection products** with drones. For example, the use of GPS-controlled drones can be used to control the corn borer by dropping the 'natural enemy'—the ichneumon wasp Trichogramma—in capsule form with pinpoint accuracy and as needed. Furthermore, spraying drones for the application of plant protection products can facilitate management in steep slope viticulture. In principle, the application of plant protection products with (unmanned) aerial vehicles is currently prohibited without a permit (Section 18 para. 1 PflSchG). In vineyards on steep slopes and in the crown area of forests, the application of plant protection products by drones may be authorised under certain conditions (cf. Section 18 para. 2 Plant Protection Act—Pflanzeschutzgesetz—PflSchG in conjunction with the Ordinance on the Application of Plant Protection Products by Aircraft).

5 Data Governance: Data Protection Law and Data Law

5.1 Agricultural Data Governance and Agricultural Data Sovereignty

A competitive and sustainable European agriculture needs an adequate data legal framework for digital farming. Against this background, the goal of a 'Common European Agricultural Data Space' is being pursued (see 'European Data Strategy', COM(2020) 66 final, p. 37). An essential component of this is 'European Agricultural Data Governance'. This includes the legal rules relating to **data sovereignty/ data ownership, data safety**, **data security, data quality, data transparency, fairness, data portability and interoperability**. The need for protection of agricultural enterprises requires digital agricultural data sovereignty as a new legal instrument in the EU, which is an expression of agricultural entrepreneurial freedom/freedom to conduct a business (on this and the following see Härtel 2020a). Holders should have a right to availability of their own agricultural data (including the handling of metadata or, if applicable, consent to deletion), confidentiality in the sense of freedom to decide on the transfer of data to the provider they have chosen themselves, and authenticity as the knowledge about the provider and its business behaviour. Data sovereignty and overall digital sovereignty are expressions of fundamental rights. The reasons for the necessity of legally anchoring agricultural data sovereignty lie—in addition to the EU's food sovereignty—in the previous gaps in protection for farmers in current law and contractual practice. Information and negotiation asymmetries between farmers and providers are causes for a (so far) lack of balance of interests in the design of private data law. In practice, dependencies on a single digital provider also occur, making it difficult or impossible to switch to other providers (vendor lock-in effect).

In the context of data governance, the data generated and also the data resulting from the use of algorithmic, new agricultural digital technologies (including Big Data analytics, cloud computing, artificial intelligence, robotics) are to be recorded and classified in legal terms and oriented towards agricultural usability. A legally compliant, differentiated data law and data management concept records the **data types** and data processing procedures in a structured manner. In terms of data types, a distinction must be made at least between personal and non-personal data (e.g. pure company data, machine data) and open data (e.g. geodata, satellite data/GPS). In addition, '**operating data**' should be established as a new legal data category to promote agricultural data sovereignty (see Härtel 2020a).

In view of the multifaceted digital-based agribusiness models, various legal constellations arise in agricultural practice concerning the creation, processing and use of data, including hybrid data. In this context, the legal positions of the various actors/participants and data stakeholders in the respective agricultural technology or digital system/platform must be examined.

The most frequently asked question is: Who 'owns' the agricultural data. From a legal perspective, this not only refers to the question of possible data ownership, but also to the structuring data-related rights. Answering this question requires a differentiated view. On the one hand, because several rights holders come into consideration, and on the other hand, because the digital applications (smart products and smart services) are interconnected. Farmers and their contractual partners face a complex situation. There are no single owners and users of data; rather, multiple interests in data occur. In the field of Agar 4.0, the following can be considered as rights holders: the farmer, contractors, machinery rings, agricultural machinery manufacturers with telemetry (hardware and software combined in one machine), contract workshops/independent workshops, computer scientists/software developers, process data processors, agricultural machinery dealers, the apps providers, platform operators and cloud providers. But insurance companies (e.g. in the context of warranty insurance according to the use of the agricultural machinery used) may also have legitimate interests in agricultural data. For each case constellation, it is necessary to find a legal balance of interests between the various parties involved.

5.2 Open Data /Geodata (Spatial Data)

Open data are a separate category of data that should in principle be available to everyone. One area of open data that is particularly relevant for digital farming is geodata. Central legal requirements for this can be found in the 'Directive 2002/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)' (OJ L 108/1 of 25.4.2007). In federal Germany, this directive has been implemented by the Federal Act on Access to Digital Geodata (Geodata Access Act—Geodatenzugangsgesetz—GeoZG) and the geodata (infrastructure) laws of the 16 federal states.[7] Geodata are legally defined as all data with direct or indirect reference to a specific location or geographical area. German geodata law applies to topics significant to digital farming, such as parcels of land, land cover of

[7] The individual laws of the federal states have different titles. In terms of their legal form, they have a number of differences in terms of state-specific requirements, although they are fundamentally common; for practical implementation in Germany, see also Bundesministerium des Innern, 4. Geo-Fortschrittsbericht der Bundesregierung, June 2017.

agricultural land, geology, soil, land use, agricultural facilities, weather conditions, the water network and protected areas (cf. for the federal government Section 4 para. 1 GeoZG). The European earth observation programme Copernicus has created a significant basis for the acquisition of geodata. The legal basis for this is Regulation (EU) No. 377/2014 of the European Parliament and of the Council (OJ 122/44 of 24.4.2014). At the global level, there is the resolution 'A Global Geodetic Reference Frame Sustainable Development' adopted by the UN General Assembly on 26 February 2015.

The range and quality of geospatial data related to agriculture have evolved significantly in recent years in the context of satellite-based land monitoring. Accordingly, access to farm information by third parties has expanded. This information could be used as a basis for agricultural speculation or anti-competitive business practices. The extent to which data protection law can help farmers here is an open question. The classification of geodata in connection with information about the agricultural business as personal data in the sense of the GDPR is fraught with legal uncertainties. The new data category 'operating data' could provide information here in future (Härtel 2020a).

5.3 Privacy Under the EU General Data Protection Regulation

The outstanding set of rules for the current digital world in general and for the agricultural digital world in particular is the European Union's General Data Protection Regulation (GDPR), which came into force on 25 May 2016 and is mandatory in all Member States from 25 May 2018.[8] The GDPR requires entrepreneurs—such as agricultural machinery dealers, agricultural machinery manufacturers and app providers—to have a differentiated **data protection management.** The responsible data processing companies are directly bound by the legal provisions of the GDPR.

If farmers are affected by violations of the GDPR with regard to their personal data, they generally have **effective data protection law** at their disposal. Thus, data-processing entrepreneurs must prove to the farmer that they comply with data protection law, that they process data in accordance with the GDPR (accountability, Article 5(2), Article 24 GDPR). A reversal of the burden of proof applies at the expense of the (agricultural technology) company processing the data. For companies, this means that it is mandatory to fully document their data processing to be able to prove the lawfulness of their trade in case of conflicts.

[8] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data, on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119 p. 1, rev. No. L 314 p. 72).

The principle of transparency (Article 5(1)(a) GDPR) is particularly important for the farmer. This corresponds to a series of information obligations of the controller with regard to the framework conditions of the processing of personal data (cf. Article 13 and 14 GDPR). In addition, the farmer has various data subject rights to enforce his data protection rights under Article 15 ff. GDPR (including the right to information, the right to rectification of inaccurate data, the right to erasure/'right to be forgotten', the right to data portability).

The farmer can only rely on the rights under the GDPR if its scope of application is given. The GDPR only applies to **personal data of** natural persons (Article 1 (1), (2) and Article 2 (2) GDPR). Individual farmers as natural persons are protected, but not agricultural enterprises as legal persons in the form of a joint stock company or limited liability company. The case law of the ECJ in the Schecke case (ECJ Joined Cases C-92/09 and C-93/09, [2010] ECR I-11063) also confirms that the fundamental right to data protection under Article 8 Charter of Fundamental Rights in principle only applies to natural persons.

Furthermore, the question arises as to which of the many data that are generated in digital farming are personal and which are not. The term 'personal data' is to be understood broadly. Under Article 4 No. 1 GDPR, this includes 'any information relating to an identified or identifiable natural person'. This includes all information that can be attributed to a farmer. The deciding factor is therefore whether the information allows statements to be made about specific/determinable persons. Clearly personal data are name, date of birth, details of address, occupation, private activities, IP addresses, cookies. In addition, behavioural patterns usually have a personal reference. This concerns, for example, the way in which a contractor carries out his activity or also, if applicable, inferences from soil conditions to the skills of the farmer. Conclusions about the type of farming practised by individual farmers also justify the personal reference of data. Employee-related data/employees (own and external workers, their working hours and skills) are personal. Factual information such as property and ownership, communication and contractual relationships of the farmer with third parties and their environment may also qualify as personal data. In agricultural practice, the following data, among others, must be legally classified: basic farm data (e.g. location, size, land structure), production-related data (e.g. cultivation, yield planning, etc.), machinery-related data (e.g. machinery equipment, operating data), customer-related data (in the case of direct marketing; in the case of the farmer's own digital business model), key business figures (e.g. contribution margins, financing), strategy and market-relevant data (e.g. farm development and positioning) as well as other machine/system-related data (e.g. quality characteristics of the harvested crops, type/quantity of inputs; utilisation requirements on grassland or organic farming, business relationships, process knowledge).

No personal data exists if data is anonymised, pseudonomised or aggregated and therefore, as a rule, no conclusions can be drawn about individuals.

However, on effective data protection law in favour of the farmer, the GDPR has some problematic situations as well as protection gaps. First, there is a protection gap for legal persons. Second, the distinction between personal and non-personal data poses a number of difficulties. In view of new digital processing methods, data also have a hybrid character (between personal and non-personal). Such grey areas lead to legal uncertainties on the part of the controller (for the legal definition see Article 4 No. 7 GDPR) (the data processor as the provider of agriculture-related smart services/smart products) and the data subject (farmer). Third, some legal clarification is still needed regarding data protection in Big Data applications. In itself, anonymised data is not subject to the scope of the GDPR. However, due to the new technical possibilities of far-reaching Big Data analyses (de-anonymisation techniques), there is an increased risk of re-identification. Fourth, there are considerable doubts about the mass suitability of consent as a legal instrument effective for data protection. The ideal of fully informed and voluntary consent encounters various barriers in practice; there is also talk of loss of control by those affected. On the one hand, in practice it is hardly feasible for the data subject to accurately comprehend the extensive and complex data protection declarations. On the other hand, in practice there is often no parity between the contractual partners; the premise of 'take it or leave it' applies.

5.4 Data Usage Rights/Data Ownership

5.4.1 Private Law Claims of Agricultural Entrepreneurs on Data

For the farmer, the question arises as to what protection there is for his pure company data/farm data, which does not constitute personal data. The 'agricultural data sovereignty' sought by farmers can currently only be achieved selectively through private law (for an overview of the treatment of data under civil law, see Riehm 2018, p. 74 ff.). Thus, according to the prevailing view, no data ownership can be derived from the German Civil Code (Bürgerliches Gesetzbuch—BGB) with regard to the entire stock of business data. A right to the data stock as another right within the meaning of Section 823 (1) BGB is also not yet recognised. There is a neighbouring right to databases under Section 87a ff. of the Copyright Act (Urheberrechtsgesetz—UrhG) (see Wiebe 2018, p. 100 ff.). If data constitute trade or business secrets, they are generally subject to protection under Section 17, 18 of the Unfair Competition Act (Gesetz gegen den unlauteren Wettbewerb—UWG). Most agricultural business data is also not subject to the regime of the Act on the Protection of Business Secrets (Gesetz zum Schutz von Geschäftsgehemnissen).

As long as there are no mandatory legal provisions on data sovereignty and economic exploitation rights, the right to non-personal data is regulated by contract. Contractual law is flexible and can also be adapted according to the needs of the individual parties involved in digital farming. This makes it possible to specify precisely which contractual partner has which rights of use and which acts of exploitation should be excluded from the outset, especially on the participation of third parties not bound by the contractual agreement. It will not be possible to cover all conceivable scenarios at the time of the conclusion of the contract, but for open cases of doubt it should be determined which of the contracting parties ultimately has 'data sovereignty' in the internal relationship (cf. Zdanowiecki 2015, p. 24). With regard to a fair structuring of the contractual data provisions for the farmer, the 'EU Code of Conduct on Agricultural Data Sharing by Contractual Agreement' in particular offers useful guidance (see Sect. 5.5).

5.4.2 Data Ownership/The Right to Property

On the data sovereignty of the farmer, there are still considerable legal uncertainties. The regulations in data protection law, contract law, tort law and fair trading law (Lauterkeitsrecht) are not sufficient. Both at the European and national level, consideration is being given to changing the legal framework of data sovereignty. Thus, on data-based business models, it is being discussed whether data ownership or the right to property of data should be created. So far, the guarantee of ownership under Article 14 of the Basic Law and Article 17 of the Charter of Fundamental Rights does not establish a comprehensive positive power of disposal over data. There is no exclusive right to data. Business and trade secrets are protected by the property guarantee. However, this does not include most of the farmers' agricultural business data. The legal systems of the other EU Member States do not regulate the right to data ownership either. The starting point for the discussion is the economic interest of various economic sectors in the exclusive use and value creation of data (for the proposal of an exclusive right to data similar to ownership, see BMVI 2017). But also in the interest of consumers in the economic participation of the data-driven digital world, it has been proposed that behaviour-generated data of citizens within the market and economic relations as an economic good be given a legal form under property law (Fezer 2017, p. 99 ff., ibid. 2018; in contrast, Kühling and Sackmann 2018, pp. 19 ff., 43 f.). However, envisaged regulations of comprehensive data ownership encounter a number of concerns. For example, there are multidimensional interests in the handling of data, which lead to corresponding conflicts of objectives. There are problems in the data protection law, technical implementation and economic handling (MPI for Innovation and Competition 2016; Drexl 2017; Jentzsch 2018). Alternative options should therefore be examined in this context. In addition to suggestions for legal regulations, data trusts and the like, these include, above all, the proposal that every citizen set up their data silo on their own local data servers. This means that they are no longer in the hands of corporations. Citizens can thus decide with whom they share the data and to which companies they grant and withdraw access (Thus the new Internet concept of 'World Wide Web inventor' *Tim Berners-Lee*, based on the open source project 'Solid' developed together with MIT, www.solid/inrupt.de). A similar approach would be conceivable for farmers.

5.5 Self-Regulation and Codes of Conduct

5.5.1 Codes of Conduct

Gaps in protection in agricultural data law could be closed by contractual regulations. In practice, however, there is a lack of contractuality between farmers and providers. The EU Code of Conduct on Agricultural Data Sharing by Contractual Agreement, which was signed by nine organisations/associations from the agricultural sector in 2018, is intended to provide guidance for the fair structuring of agricultural data rights.[9] This Code of Conduct has come into being in the course of private self-regulation and, due to its voluntary nature, constitutes private soft law. The Code of Conduct contains the following aspects: Definitions; Attribution of rights to derive data; Data access, control and portability; Data protection and transparency; Data protection and security; liability and intellectual property rights; Agricultural data rights case studies; Checklist for a fair contract.

The Code of Conduct is a first good approach to balance the interests of market participants. However, its shortcomings lie in the fact that it is too vague in many formulations and does not make any concrete recommendations for resolving conflicts of interest. This particularly applies to the question of data ownership. The Code of Conduct does not provide an answer as to which specific rights farmers are entitled to in the event of multiple data authorship (several market participants are data authors) within the framework of the digital data value chain. The possible right of access to data is also insufficiently explained. Data security/cybersecurity are only rudimentarily addressed and need to be specified. Liability is also only superficially addressed. Data portability is oriented towards the General Data Protection Regulation. However, the concrete guarantee of data portability by the provider vis-à-vis the farmer is not clarified. Interoperability vis-à-vis the farmer is not demanded. In addition, there are no references to possible legal consequences and sanctions in the event of violations of contractual provisions.

These deficits should be eliminated in a further development of the Code of Conduct. A look at foreign codes of conduct is useful here. In the USA, the 'American Farm Bureau Federation', the largest American farmers' association, established the 'Privacy and Security Principles for Farm Data' in 2014. On 5 May 2015, a number of agricultural organisations and companies signed this Code of Conduct. The three-page document describes the 13 points listed as basic principles. The first basic principle is the 'education' of farmers. The industry should work to develop programmes that create educated customers who understand their rights and obligations. Furthermore, contracts should be written in simple, easy and understandable language. Under the heading 'ownership', it is stated that farmers have ownership of data generated in the course of their business operations. In contrast to the EU Code of Conduct, the US provides for certification as an instrument for enforcing the guidelines/principles for agricultural data contract

[9] Copa and Cogeca, CEMA, CEJA, ECPA, EFFAB, FEFAC and ESA.

design, whereby an agricultural data transparency seal (‘Ag Data Transparency Evaluator’) is awarded.

The New Zealand Farm Data Code of Practice (NZ Farm Data Code) was established in 2014 to provide guidelines for effective data sharing in the New Zealand agricultural industry. The code explicitly covers the data of primary producers (farmers). A certification system is used for implementation, whereby a seal is awarded. In the certification process, the AgTech company has to demonstrate that it complies with the Farm Data Code by submitting a self-declaration. The compliance checklist (list of questions) is more comprehensive than the one in the US certification system.

The ‘Australian Farm Data Code’ of 1 February 2020 has been drawn up by the National Farmers Federation and the agricultural industry. In contrast to the codes in the USA and New Zealand, no certification system is foreseen. After an introduction, the principles of agricultural data include transparency, comprehensibility, honesty, fair and equitable use of data, the farmer’s right to access data and portability.

In Germany, there is a joint industry recommendation of seven associations on ‘Data sovereignty of the farmer’ of 28 February 2018.[10] The recommendation refers to the ‘collection, use and exchange of digital farm data in agriculture and forestry’. One basic principle is the ‘ownership of data’.

From the farmers’ point of view, the problem with the existing Codes of Conduct is that agricultural data sovereignty is not sufficiently guaranteed. Certification systems seem to be able to provide a certain remedy. The certification system in New Zealand is more differentiated than in the USA.

5.5.2 Regulated Self-Regulation

The necessary impetus for the further development of codes of conduct in terms of an appropriate balance of interests between farmers and other market participants is provided by the EU legislation. This is where the Free Flow of Data Regulation (EU) 2018/1807 comes in by requiring in Article 6 that codes of conduct be based on the principles of transparency and interoperability and implement open standards. On a change of provider or a transfer of the data back into own IT systems, it should be ensured, among other things, that the data is made available in a structured, common, machine-readable format. Data dependency on one provider—data lock-in—should be reduced or avoided. Another important impulse is given by the Platform-to-Business Regulation/P2B Regulation (EU) 2019/1150. Under Article 9 Regulation 2019/1150, providers of tech products/services must explain in their general terms and conditions the technical and contractual access or lack of such access for users to data. Providers must adequately inform the user in particular

[10] German Farmers’ Association, Federal Association of Machinery Rings, Federal Association of Contractors, German Agricultural Society, German Raiffeisen Association, LandBau-Technik-Bundesverband, Verband Deutscher Maschinen- und Anlagenbau.

whether he has a right to access data, to which categories of data and under which conditions he has access (if applicable, against payment or surcharge/pricing in the contract), whether he also has access to aggregated data (relating to the agricultural business data), whether third parties are allowed to access the data, what options exist for the user to refuse the transfer of data.

6 Cybersecurity

In addition to data security, which is regulated for personal data in Article 32 ff. GDPR, cyber security is of considerable importance in the digitisation of the agricultural and food sector.[11] The Union regulatory framework in this regard can be found in Regulation (EU) 2019/881 of 17 April 2019 on ENISA (European Union Cyber Security Agency) and on the certification of cyber security of information and communication technology (OJ L 151/15 of 7.6.2019) and in Directive (EU) 2016/1148 of 6 July 2016 on measures to ensure a high common level of security of network and information systems in the Union (OJ L 194/1 of 19.7.2016). At national level, cyber security is regulated in particular for critical infrastructures. This explicitly includes the food sector. The provision of food supplies to the general public is classified as a critical service within the meaning of Section 10 (1) sentence 1 of the BSI Act.[12] The reporting obligations for IT security apply to operators of facilities for food production and processing as well as food trade above a certain size (Cf. Annex 3 Part 3 Column B Ordinance on the Designation of Critical Infrastructures under the BSI Act (BSI Critical Infrastructure Ordinance)). However, agriculture as a primary production has not yet been legally qualified as a critical infrastructure. Irrespective of this, digitalised agricultural enterprises require IT security measures. Relevant technical standards on IT security provide orientation for this.

7 Outlook

A developing Agricultural Digital Law or Agricultural Law 4.0 is a cross-sectional matter like the agricultural law on which it is based, which encompasses the areas of public law, private law and criminal law. At the same time, Agricultural Digital Law will be characterised by its multi-level reference in national, European and

[11] It refers in particular to technical and organisational measures taken by the controller to reduce the risks of destruction, loss, alteration or unauthorised disclosure (whether accidental or unlawful) of personal data.

[12] BSI: Bundesamt für Sicherheit in der Informationstechnik - Federal Office for Information Security.

international terms. Hard law and soft law as well as hybrid legal forms will be included. A constant dynamic development of the legal material is emerging, which accompanies the new technological developments of digital farming up to the comprehensive Agriculture 4.0 in the necessary differentiation. In the future, this will also include more advanced digital technology applications, such as fully automated agricultural machinery in the sense of autonomous driving. The first prototypes of autonomously driving agricultural machinery are ready for testing in the field. At the same time, unmanned (autonomous) driving on public roads is still prohibited in Germany.[13] For the development of a functioning digital farming, the creation of a high-performance 5G mobile as well as a comprehensive fibre optic network in rural areas is required within the framework of digital services of general interest. This applies in particular to the development of a digital ecosystem comprising value chains, value systems and value networks. This is linked to the formation of an agricultural Internet of Things (A-IoT). Other developments in the agricultural sector are also being advanced by digitisation technologies, e.g. vertical and horizontal urban farming. In addition, digital networking effects are emerging in the course of globalisation. One example is the control of irrigation of agricultural farms in Zambia from a corporate headquarters in Munich. The same applies to pest control, the use of agricultural drones (e.g. for fields in Liberia or rubber plantations in the Ivory Coast, combined with apps and AI), the use of agricultural apps such as Plantix or innovative applications from specialised companies such as The Yield. A further step means linking digitisation and the bioeconomy, which should lead to a biological transformation of industrial value creation in Germany under the aspects of sustainability and innovation.[14] However, the inclusion of agricultural value creation in the form of a digital ecosystem will be fundamental to the vision of a biointelligent total value creation. Biointelligence, digitisation and farmers' experiential knowledge are essential elements of a future comprehensive Agriculture 4.0.

References

Bundesministerium für Verkehr und digitalen Infrastruktur (2017) "Eigentumsordnung" für Mobilitätsdaten?

Drexl J (2017) Neue Regeln für die Europäische Datenwirtschaft. Ein Plädoyer für einen wettbewerbspolitischen Ansatz, NZKart, 339 ff

Fezer K-H (2017) Dateneigentum der Bürger, ZD, 99 ff

[13] Under Section 1a para. 2 no. 3 Road Traffic Act, manual override by the driver is required at all times. However, there are reference and test routes. On possible ethical problem situations, see Härtel (2019), p. 58.

[14] See the comprehensive preliminary study on biological transformation 'Biointelligence - a new perspective for sustainable industrial value creation', which brings together the various value creation areas from different bio-based sectors in a unified and perspective-based approach, Fraunhofer Gesellschaft 2019.

Fezer K-H (2018) Repräsentatives Dateneigentum. Ein zivilgesellschaftliches Bürgerrecht. In: Kuzev P, Wangermann T, Konrad Adenauer Stiftung
Floridi L (2015) Wie die Infosphäre unser Leben verändert, 2015
Härtel I (2015) Ein (Menschen)Recht auf Nahrung? In: Geis M-E u.a. (Hrsg.) FS Hufen, S. 23 ff
Härtel I (2019) Digitalisierung im Lichte des Verfassungsrechts – Algorithmen, Predictive Policing, autonomes Fahren, LKV, 49 ff
Härtel I (2020a) Report on the topic of "European Guidance and Rules for Agricultural Data" (European Agricultural Data Governance), 2020 /Gutachten zum Thema "Europäische Leitlinien bzw. Regeln für Agrardaten"
Härtel I (2020b) Künstliche Intelligenz in der nachhaltigen Landwirtschaft - Datenrechte und Haftungsregime, Natur und Recht 2020, 439 ff
Härtel I, Yu H (2018) Food security and food safety law. In: Härtel (Hrsg.) Handbook of agri-food law in China, Germany, European Union, S. 57 ff
Jentzsch N (2018) Dateneigentum – Eine gute Idee für die Datenökonomie?, Stiftung Neue Verantwortung
Kämper N, Müller I (2017) Rechtliche Rahmenbedingungen für den Betrieb von Drohnen, UPR 2017, 401 ff
Kühling J, Sackmann F (2018) Rechte an Daten. Regulierungsbedarf aus Sicht des Verbraucherschutzes?, Rechtsgutachten im Auftrag des VZBV
Max-Planck-Institut für Innovation und Wettbewerb (2016) Ausschließlichkeits- und Zugangsrechte an Daten, Positionspapier vom 16.8.2016
Monien J (2014) Die "Freiwilligen Leitlinien für die verantwortungsvolle Verwaltung von Boden- und Landnutzungsrechten, Fischgründen und Wäldern im Kontext nationaler Ernährungssicherhung" der Vereinten Nationen in rechtlicher Perspektive. In: Härtel I (Hrsg.) Nachhaltigkeit, Energiewende, Klimawandel, Welternährung, S. 789 ff
Riehm T (2018) Dateneigentum – Schutz nach allgemeinem Zivilrecht. In: Hornung G (Hrsg.) Rechtsfragen der Industrie 4.0, S. 73 ff
Schrader Chr (2017) Drohnen und Naturschutz(recht), NuR 2017, 378 ff
Wiebe A (2018) Schutz von Daten im Immaterialgüterrecht. In: Hornung G (Hrsg.) Rechtsfragen der Industrie 4.0, S. 97 ff
Zdanowiecki K (2015) Recht an den Daten. In: Bräutigam P, Klindt T (Hrsg.) Digitalisierte Wirtschaft/Industrie 4.0. Ein Gutachten der Noerr LLP im Auftrag des BDI zur rechtlichen Situation, zum Handlungsbedarf und zu ersten Lösungsansätzen, http://bdi.eu/media/themenfelder/digitalisierung/downloads/20151117_Digitalisierte_Wirtschaft_Industrie_40_Gutachten_der_Noerr_LLP.pdf

Law 4.0? Considerations on the Future of Law in the Digital Age

Volker Boehme-Neßler

Contents

V. Boehme-Neßler (✉)
Carl von Ossietzky University Oldenburg, Oldenburg, Germany
e-mail: volker.boehme-nessler@uni-oldenburg.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_18

1 Digitalisation of the World—and of the Law?

Digitalisation is ubiquitous and shapes the (post)modern world. It influences everything—the economy and politics, people's behaviour, thinking and psyche. The Internet and the phenomena associated with it will lead to a change of thinking habits, which in its consequences can hardly be assessed yet (Katzer 2016, p. 75).

This effect makes digitalisation not only a technological but also a cultural and highly political phenomenon (Boehme-Neßler 2008, p. 100 f). Of course, it has consequences for the law. Countless laws and individual standards have already been amended. The legislative work will—and must—continue undiminished. But the development goes much deeper: even basic ideas and basic structures of the law are under pressure for change. Maybe we even stand at the beginning of **the end of the law** (Kurer 2016; Susskind and Susskind 2015, p. 66 ff.)

2 Contrasts and Conflicts: Analogous Law vs. Digital World

Digitisation follows basic principles and an immanent logic which constitutes its character. This also applies to the law. It also has a typical logic and specific basic principles which it follows. Both logics are completely different. In other words, when digitalisation meets law, two completely different worlds collide.

2.1 The End of Borders?

The digital world is intangible or immaterial. Digitised content is potentially omnipresent or ubiquitous: it can be created anywhere, moved anywhere and accessed anywhere. As a result, borders—at least potentially—are neither here nor there when it comes to digitised content.

The concept of the dissolution of boundaries is carried to extremes by **ubiquitous computing** (UC)—the vision of the ubiquitous and invisible computer network. Computers are seamlessly and invisibly integrated into the environment and networked with each other (Weiser 1991, p. 66 ff). The ubiquitous but invisible computers create a **smart environment**.

Ultimately, the concept of the 'smart' environment boils down to a complete dissolution of boundaries: the real world is linked to the virtual world. The fundamental boundary between the real and virtual world is increasingly being dissolved. A particularly fascinating but at the same time irritating vision is **augmented reality** (Wellner 1993).

In this respect, law is the extreme counter model of digitalisation: it is not boundless, but inherently limiting and limited. Delimitation is a 'traditional strategy of law' (Hoffmann-Riem 2006, p. 202). Inspired by *systems theory*, law can be seen as the drawing of boundaries. The guiding distinction of the legal system is the distinction between right and wrong (Luhmann 1993, p. 60). Seen from this angle, the demarcation between right and wrong is constitutive for law.

Boundaries also play an important role within the legal system. The fundamental task in both legislation and application of law is to distinguish between relevant and irrelevant. Which (small) part of reality should be legally relevant is at first decided by the legislator, then by the person applying the law. From this point of view, the application of law in the form of subsumption is demarcation.

2.2 Algorithm or Parliamentary Act?

Since the Internet was primarily a technical rather than a political or social phenomenon in its early days, engineers, mathematicians and computer scientists dominated the technical and political development of the Internet. Technicians set the binding standards, not politicians and lawyers. As a result, the **digital code** has **no democratic legitimacy** (Boehme-Neßler 2020, p. 67 ff).

2.2.1 Internet Governance: Technology and Economy Instead of Politics

Technology and engineers are inevitably of great importance in the (further) development of technology law. This leads to a preponderance of legislation and is finally reflected in the content of technical standards. Non-technical thinking and interests which are **outside of the technical community** can only be integrated through state regulation.

Just as the **history of digitalisation** evidenced this old experience. Until the early nineties of the last century, the Internet community was primarily concerned with the development and adaption of technical standardisation, especially of network and application protocols (Hafner and Lyon 2008, p. 165 ff.). Political and ethical matters played, if at all, only a minor role. The basic technical standards and infrastructures can be described as **digital code** (the term was coined by Mitchell 1996, p. 111). Spoken in the categories of the sociology of power: the technical forums and programmers wield **data-setting power** (Popitz 1992, p. 31 and 180). They have the power of technical crafting: with their software and technical standards they shape a part of the world that is becoming increasingly important. Those 'persons affected by data' are confronted with this and have to adapt.

2.2.2 Democratic Deficit: The Digital Code

Just as legal regulations and social standards influence analogue life, the digital code shapes life in cyberspace. In other words, it is the law in the **City of Bits** (Mitchell 1996, p. 111).

However, the digital code has a specificity in principle: in contrast to laws or social norms, it is impossible to evade it in the digitalised space. In this respect, it corresponds more to a law of nature than to a set of norms created by man. Nevertheless, it is of course developed and put into force by humans. The digital code of cyberspace can—and must—therefore be politically controlled and changed—within the limits of what is technologically possible. In other words, in a **democracy** every **exercise of power is limited**. This is a fundamental idea of democratic theory. It goes without saying then that this must also apply to the power of the data-setters.

What is possible in cyberspace has so far not been defined by democratically legitimised legislators, but by technical standards and software. The **engineers** and programmers are becoming lawmakers. This is highly problematic, not least from the point of view of democratic theory (Boehme-Neßler 2020, p. 67 ff). Given the importance of the Internet, decisions on Internet architecture and software are highly political and of considerable explosiveness for politics, society and the economy. It is unacceptable that nerds and businessmen who have no democratic legitimacy shape the code—and thus make highly political decisions (Boehme-Neßler 2020, p. 69 ff).

2.3 On the Road to Pictorial Law?

Never has it been so easy to take pictures and distribute them. The digitalised world is therefore largely a world of pictures. This has consequences for the law, which has so far been more a world of texts (Boehme-Neßler 2010, p. 163 ff).

2.3.1 Pictorial Turn: From Writing to the Picture

The modern digitalised world is shaped **by the image, no longer by the word**. Communication on the Internet is visual communication to great extent. Cultural studies quite rightly diagnose a *pictorial turn*—not without scepticism or even fear (Mitchell 1994, p. 12 ff). The digitalisation of culture also means the visualisation of culture. This has a dramatic consequence: the significance of writing is relativised. This applies to the Internet, but also to culture in general. As part of culture, law is also affected by this development.

2.3.2 Law and Pictures

In any case, modern **law is very language-focused** and sceptical or even hostile towards pictures. Images lead a shadowy existence in law (Baer 2004, p. 240). Legal texts—be they laws, judgments or scientific texts—usually do not contain images. Textbooks have virtually become a symbol of law. However, here again, exceptions confirm the rule. The Road Traffic Act with its illustrations of traffic signs are the most prominent example. As well in copyright, patent and trademark law, images are not only commonplace but indispensable. Where does the picture scepticism of modern law come from?

Since the Reformation, pictures be a byword for **sensuousness, sin, emotionality** and **irrationality** (Goodrich 1995, p. 56). It that time are the roots of the **logocentric prejudice** in which the law is still trapped. Ever since especially language in written form has been considered rational. Images are dismissed as tending to be primitive or at best decorative. Modern legal thinking still stands in this highly problematic tradition. It postulates that law should not be emotional and irrational. The aim of law is to explore 'the truth' in a rational process by applying reasonable methods. According to the established tenet, images are not useful, but even counterproductive (Schuppert 2004, p. 75 ff).

Written language is an ideal tool for the law to implement **social control** (Rehbinder 2014, p. 100). Images, on the other hand, have—at least in part—an opposite function: they are supposed to give **freedom** to the individual for escaping from the social and legal pressure in society. Prima facie, pictures are in fact rather not an appropriate means for the law to fulfil its regulatory and control function. It is possible, however, that this mistrustful and negative relationship between law and pictures is currently undergoing profound changes. Cos within the law visualisation processes can be diagnosed.

2.3.3 Visualisation Tendencies Within the Law

Verbal communication—unlike written communication—is very much **visual communication**. In addition to verbal communication, non-verbal communication is essential. Facial expressions, gestures, postures, gaze behaviour and the use of space are communication signals that have a visual effect. It is therefore no wonder that it was a court case in which a film was used as procedural evidence for the first time. In the **Nuremberg War Crimes Trials**, a documentary about the Nazi concentration camps made by the American army was an important part of the indictment (Douglas 2000, p. 198 ff.).

Forensic work in litigation has always consisted less of abstract legal argumentation. The focus is—not only, but particularly clearly in the trial—on the development and construction of a concrete story on which the judgement can be based. In US-American courtrooms in particular, modern technologies are increasingly being used, which are also and especially intended **to visualise** the legal argumentation

(Katsh 1995, p. 159). Meanwhile, there are specialised companies which produce so-called **legal videos**. The videos are used—at least in larger trials—as part of the pleadings. Not surprisingly, the USA is taking a pioneering role in the visualisation of court proceedings. That depends not least on the American jury system, which accelerates the adoption of current cultural techniques and the 'spirit of the times' into the legal system.

Visualisation tendencies, which are promoted by modern media technologies, can also be observed in legal relations outside of court hearings. The media processing of general meetings of stock corporations is being tested in practice. In the area of administration, the first tentative signs of an opening to the culture of pictures can be found. In staged approval procedures, in which decisions are made on the admissibility of highly complex technical installations, administrative law, which is still **paper-based**, in practice is reaching the limits of its capacity. Pictures should suppose the administration for a better understanding of complex connections and causalities. This should have a positive effect on the legal quality of the administrative decision.

Visualisation can also be observed in quite unspectacular forms in everyday legal life. Information graphics are beginning to become an accepted means of legal communication—albeit for the time being mainly in educational literature (Röhl et al. 2005, p. 248). Tables, synopses, graphic representations, decision trees and flow charts are increasingly used in legal publications. Spatial relationships are represented graphically. Figures are represented, or at least clarified, by aesthetically designed diagrams.

Demographic developments will drive the visualisation of law. What can this—at first sight perhaps daring—forecast be based on? Younger generations are growing up in a world that is characterised, if not dominated, by pictures. They therefore develop communication skills and habits that are strongly coloured by visuals. Their communication habits will be more visual and less characterised by writing and text. With the generational change in legal personnel, the legal system will also gradually change its kind of communication. Visual communication will gain in importance.

2.3.4 Opportunities and Risks: Pictures Within the Law

Pictures are necessary in law because they can fulfil other functions important for communication than texts (Brunschwig 2001, p. 69 ff., 136 ff). **Visual communication** can be more **comprehensive** than linguistic communication and can cover considerably more aspects, information and content. Pictures, for example, can convey information that cannot be reflected at all in texts (Hasebrook 1995, p. 113 ff). This applies—not only, but especially clearly—to information about spatial arrangements or very complex issues.

If the law increasingly discovers **pictures as another instrument of communication** alongside language and writing, this will certainly have positive effects. Pictures can be a tool for a better understanding of increasingly complex processes and facts. In this respect, the efficiency of the law can increase. However, this optimistic view must not overlook the dangers associated with visual legal communication. This is because images have a few specific characteristics which at the same time significantly limit their capacity for legal communication. The limitations of visual communication are obvious when it comes to the representation of abstract facts and concepts. This is a decisive disadvantage, especially for continental European law, which operates at a high level of abstraction. In addition to its **weakness in abstraction**, visual communication also has other deficits that are not visible at first glance. One problem that should not be underestimated is that **visual communication is in principle more emotional** than communication with words and texts (Katzer 2016, p. 173 ff).

Images have specific communicative strengths and at the same time very marked weaknesses. The same applies to spoken and written language. However, the advantages and deficits of pictures and texts are not identical, but rather complementary. For the law it is therefore important to find an optimal word-image balance. In this way, the strengths of visual and linguistic communication can be exploited while avoiding the weaknesses. A successful **word-image balance** considerably enhances the quality of legal communication.

3 Indispensable? Law in the Digitalised World

Existing law is conflicting increasingly with the fundamental ideas, basic structures and requirements of the digitalised world. The digital code is in the process of supplanting the law in many areas. This raises a crucial question: **Does the digitalised world still need a legal system at all?** Or are new structures developing because of the Internet's triumphant advance, which (could) make law superfluous? In other words, can the digital code replace the law?

3.1 A Need for Law? Norms for the Digitalised World

Is law still necessary in the digitalised world? This depends, not least, which functions a legal system must fulfil for the society.

3.1.1 Public Good and Digital Code

In contrast to the law, the **public good** is not a relevant category in the digital code. The code is technically inspired and usually pursues economic and pragmatic goals. Thus, it can hardly prevent strong individual interests from asserting themselves against weaker individual interests or the general interest. It cannot—and does not want to—compensate unequal economic, social, cultural and political power relations. The digital code does not guarantee the necessary protection rights for weaker individual interests or minorities.

Guaranteeing the protection of minorities and enforcing the public good is a classic function of the state and state law. As long as the digital code does not consider the general interest, state law remains (still) necessary. However, it is not a law of nature that only state law can ensure the public good. A glance at history and a look at large parts of the world show that other structures and mechanisms may also be capable of doing this (Kurer 2016).

3.1.2 Law: Order Instead of Chaos

The question of what functions the law has will attract very different and differentiated answers. However, there is broad agreement on one function. The law is intended to **control** people's **behaviour**, attitudes and expectations in such a way that conflicts are avoided (Rehbinder 2014 marginal no. 100). It focuses on those areas where conflicts exist or are to be expected. Where no conflicts exist, law is not necessary.

In principle, conflicts arise when different actions clash interests, expectations, goals, characters or people who are—or appear to be—incompatible (see Myers 2008, p. 673). So, differences are not per se a cause of conflict. What is decisive is the—apparent or real—incompatibility of the differences. Conflict therefore occurs above all when differences are not (or cannot be) tolerated. When it comes to important, fundamental goals, characteristics, expectations, mutual tolerance is much more difficult. Differences always become particularly conflict-prone when they are emotionally charged. Against this socio-psychological backdrop, it is unlikely that conflicts will become less frequent in the digitalised world. Conditions will become more confusing, and boundaries will blur. This rather suggests that conflicts will increase.

3.1.3 Mutual Trust Through Law

Trust is indispensable (from a historical perspective Frevert (2003), p. 7 ff. with further references). It is of great importance for action in social situations. Those who trust can also act in situations whose complexity they cannot fully comprehend. Trust makes it possible to take calculated risks and thereby expand one's own options for action. Those who have trust can actively act in situations that are uncertain and risky. They do not have to wait passively and defensively. Thus, trust is an important basic prerequisite for cooperation and compromise; it is one of the most important social forces that hold societies together and make economic relations possible, or at least facilitate them (fundamentally Simmel 1968, p. 263).

Against this socio-psychological backdrop it is an important **task of the legal system to enable and protect trust**. Already as an institution—by its very existence—the law creates trust. When and because the law can be trusted, people's scope for action expand. Because in an intact legal system, mechanisms exist which sanction the disappointment of legitimate trust. If and because the law creates trust, one can live in more complex societies in which personal mechanisms for building and securing confidence are no longer sufficient (von Rohr 2001, p. 155 ff). Thus, the law fulfils an important function in the development of differentiated, complex societies.

Is the **digital code** capable of performing this function of law and creating the necessary trust? So far certainly not. As long as no other institution can effectively ensure confidence-building, state law remains necessary. But here too, it is not a law of nature that (systemic) trust can only be established by state law. Historical experience and global comparisons show that other mechanisms and structures may also be able to do this.

3.2 Democracy in Cyberspace?

There is also a democratic dimension to the question of the necessity of law. (State and supranational) Law has a great significance for democracy. The fundamental idea of democracy is popular sovereignty. Article 20 (2) sentence 2 of the German Basic Law sums this up succinctly: **all state power emanates from the people**. At least, in a representative democracy, legal regulations are the people's most important means of expressing their will and enforcing it with (constitutional) power (fundamentally BVerfGE 93, 37, 66 ff.) In democracy, the exercise of power is, in principle, only permissible if and insofar as it is democratically legitimised and controlled.

This is the case when policies are enforced through legal regulations made by the Parliament. However, a digital code enforced by Internet corporations with economic and social power contradicts the democratic idea. To put it bluntly: only parliamentary law is democratically legitimised, not Facebook law.

Of course, Facebook, Google, Apple, PayPal and other Internet companies are allowed to regulate how they want to sell their products and provide their services. This is a question of freedom of contract and the law on general terms and conditions. What is decisive, however, is that the binding framework and the scope for this, is set in place by supranational or state parliaments. In parliamentary democracy, the instrument of choice for this is the law. Anything else would contradict the current understanding of democracy in the German constitution and the European treaties (Boehme-Neßler 2020, p. 75 ff).

In any rate, at present **the digital code is not democratically legitimised**. For this reason alone, at present, it cannot be a constitutionally permissible substitute for state and supranational law. To the extent that it is in fact so in practice, it violates the German constitution and the European treaties.

4 The Blur: The Law of the Digitalised World

The digitalised world is blurred. It is doubtful whether the current law as we know it can cope with it. If law should be remaining relevant, it must—whether it wants to or not—accept the blur. It remains an open question.

4.1 Blurred World

A striking characteristic of (cultural) digitalisation is its **blurring** (Boehme-Neßler 2008, p. 662 ff). In the digitalised world, electronic impulses are crucial. The actual, material embodiment in the 'real' world is becoming less important. This leads to impermanence and blurring (Boehme-Neßler 2008, p. 378 ff). Materialisation facilitates demarcation and differentiation, and it can provide clarification. In the digitalised world of life, all boundaries blur or even dissolve. That is the big difference to the analogue world as we know.

The current law is adapted to the analogue world with its clear limits. However, its basic concepts and its regulations in detail increasingly no longer corresponds to the problems of the fuzzy, digital world. This is exemplified by problems of liability and responsibility within the digitalised world (Spiecker gen. Döhmann 2016, p. 700 ff). Ultimately, the question then arises about if and how the law can fulfil its classical functions in the fuzzy world. Unless it is not completely mistaken, the law will have to relinquish old functions and take on **new tasks and roles** simultaneously.

4.2 Fuzzy Law?

Digitalisation is changing the tasks of the law as well as its instruments. This is a threefold challenge—for legal theory, for legal dogmatic and for legal policy.

Legal thinking has so far been strongly characterised by sharp demarcation and precise control. The conditional logic, which—still–dominates law, is a visible expression of this. The **fuzzy logic**, which is spread by digitalisation, stands in marked contrast to this. This is an—important and difficult—task for legal theory and legal dogmatic: they must integrate the logic of uncertainty into law and translate it into new legal concepts, legal institutions, conceptions and ideas.

Fuzzy law can no longer fulfil the same functions for society in the same manner as before (Boehme-Neßler 2008, p. 665 et seq. Different interpretation for the German Civil Code Wendehorst 2016, p. 2609, which expects only 'selective changes' to the German Civil Code as a result of digitalisation). Some tasks will no longer be able to fulfil the fuzzy law at all. It will have to perform other functions in a different form than before. This applies, for example, to the control function. An example: Fuzzy law that is influenced by (digital) pictures and works with pictures can no longer provide a strict, precise control of individual behaviour and social processes (Boehme-Neßler 2010, p. 163 ff). Visualised law is more likely to trigger developments, mark a rough direction and set general goals.

This is an insight that has yet to be accepted in legal policy. Otherwise, false ideas about the design and efficiency of legal rules in the digitalised world will arise. A fatal consequence would be bad laws that want to operate in the blurred, digitalised world with the outdated sharp logic. A **legislative doctrine of fuzzy law** still needs to be developed.

5 Relativisation: Law and Algorithms

The digitalisation of the world puts the law under pressure: the more the society, the economy, politics and culture digitalise, the greater the appendant pressure for change on the legal system. If the legal system refuses to accept this pressure, it runs the risk of becoming alienated from the day-to-day world and losing its significance.

The significant increase in private, contract-based international commercial law and the importance of private arbitration are indications of such a **creeping loss of importance of state law** (Stein 1995, p. 35 ff). Another, practically very relevant example: In private electronic commerce, **primitive private rules** of law have now developed which take over the functions of state law (Fries 2016, p. 2861 et seq). There, in wide range, state law plays only a formal role anymore, but no longer de facto. An example: PayPal, for instance, decides according to its own rules. The sophisticated and differentiated rules for default and consumer protection rights of the German Civil Code have practically no effect.

Ultimately, this means that the really relevant rules are increasingly being set by Facebook, Amazon, Google or other Internet companies. This situation has no longer anything to do with the (German and European) constitution—such as the rule of law or the idea of democracy.

Even if the law continues to develop innovatively, there is probably no getting around by one finding: the importance of the **law will relativise in the long term** (Boehme-Neßler 2008, p. 635 ff). For a continuance, the legal system will be less and less able to fulfil its traditional functions in the digitalised world (that way Susskind and Susskind 2015, p. 66 ff. with further references. Kurer (2016) claims that in the digitalised world the law will dissolve ‘like sugar in a teacup’).

The resigned consequence would be that the law would accept its dwindling significance in the digitalised world and set out on a retreat. But that is risky. It is uncertain whether others will develop other institutions that can functionally replace the law. The combative alternative is that the law **must reinvent itself**. It must develop new instruments and seek for allies in other areas of society with whom it cooperates (Boehme-Neßler 2008, p. 641 ff). An important ally would be, for instance, computer science and software development (similar to Bräutigam and

Klindt 2015, p. 1142). Environmental law, which is just beginning to develop (Boehme-Neßler 2020, p. 70 f. with further references), would be particularly effective. Thereby, the law and its evaluations are 'inscribed' directly into the software, the code, and thus evolve control effects. Such **syntheses of software and law** would democratically legitimise the digital code. Then the law could continue to bring its strengths to fore, but at the same time compensate its deficits through intelligent cooperation with other areas of the world. The law would remain an influential—and democratically legitimised—part of the digitalised world. But where this development will take us is not certain. We live in an age of disruptive developments.

References

Baer S (2004) Schlüsselbegriffe, Typen und Leitbilder als Erkenntnismittel zur Rechtsdogmatik. In: Schmidt-Aßmann E, Hoffmann-Riem W (Hrsg.): Methoden der Verwaltungsrechtswissenschaft. Baden-Baden, S. 223 ff

Boehme-Neßler V (2008) Unscharfes Recht, Berlin

Boehme-Neßler V (2010) BilderRecht. Die Macht der Bilder und die Ohnmacht des Rechts, Berlin/ Heidelberg.

Boehme-Neßler V (2020) Digitising Democracy. On Reinventing Democracy in the Digital Era – A Legal, Political and Psychological Perspective. Cham/Switzerland

Bräutigam P, Klindt T (2015) Industrie 4.0, das Internet der Dinge und das Recht. NJW, S. 1137

Brunschwig C (2001) Visualisierung von Rechtsnormen. Zürich

Douglas L (2000) Der Film als Zeuge. Nazi Concentration Camps vor dem Nürnberger Gerichtshof, in: Baer, Ulrich (Hrsg.): "Niemand zeugt für den Zeugen." Erinnerungskultur nach der Shoah. Frankfurt am Main, S. 197 ff

Frevert U (2003) Vertrauen – eine historische Spurensuche. In: Frevert U (Hrsg.): Vertrauen. Historische Annäherungen. Göttingen, S. 7 ff

Fries M (2016) PayPal Law und Legal Tech - Was macht die Digitalisierung mit dem Privatrecht? In: NJW, S. 2860

Goodrich P (1995) Oedipus Lex. Berkeley

Hafner K, Lyon M (2008) ARPA KADABRA Die Geschichte des Internet. 3. Auflage, Heidelberg/ Berlin

Hasebrook J (1995) Multimedia-Psychologie. Heidelberg

Hoffmann-Riem W (2006) Governance im Gewährleistungsstaat – vom Nutzen der Governance-Perspektive für die Rechtswissenschaft. In: Schuppert GF (Hrsg.) Governance Forschung. 2. Aufl. Baden-Baden, S.195 ff

Katsh EM (1995) Law in a digital world. New York/Oxford

Katzer C (2016) Cyberpsychologie. München

Kurer P (2016) Der Jurist im digitalen Holozän. In: NZZ v. 19.4.2016, www.nzz.ch/bildstrecken/bildstrecken-feuilleton/rechtsanwaelte-und-neue-technologien-der-jurist-im-digitalen-holozaen-ld.14594 (28.2.2018)

Luhmann N (1993) Das Recht der Gesellschaft. Frankfurt/Main

Mitchell WJT (1994) When is seeing believing? In: Scientific American. Februar. 1994, S. 44 ff

Mitchell WJT (1996) City of bits: Space, place, and the infobahn. Cambridge University Press, Cambridge

Myers DG (2008) Psychologie. 2. Aufl. Heidelberg

Popitz H (1992) Phänomene der Macht. Tübingen

Rehbinder M (2014) Rechtssoziologie. 8. Aufl. München

Röhl KF, Böhm A, Böhnke M, Langer T, Machura S, Marfels G, Ulbrich S, Weiß M (2005) Das Projekt “Recht anschaulich”. In: Hilgendorf E (Hrsg.): Beiträge zur Rechtsvisualisierung. Berlin, S. 51 ff

Schuppert GF (2004) Bilderverbot? Einige verfassungsrechtliche Überlegungen zu den Grenzen visueller Kommunikation. In: Frankenberg G, Niesen P (Hrsg) Bilderverbot. Recht, Ethik und Ästhetik der öffentlichen Darstellung. Münster, S. 70 ff

Simmel G (1968) Soziologie. Untersuchungen über die Formen der Vergesellschaftung. 5. Auflage, Berlin

Spiecker gen. Döhmann I (2016), Zur Zukunft systemischer Digitalisierung – Erste Gedanken zur Haftungs- und Verantwortungszuschreibung bei informationstechnischen Systemen. CR, S. 698 ff

Stein U (1995) Lex Mercatoria – Realität und Theorie. Frankfurt/Main

Susskind R, Susskind D (2015) The future of the professions. Oxford University Press, Oxford

von Rohr A (2001) Evolutionsbiologische Grundlagen des Rechts: Zum Einfluss neurogenetischer Information auf das Recht. Berlin

Weiser M (1991) The Computer for the 21st Century. Scientific American, September, S. 66 ff

Wellner P (1993) Interacting with paper on the digital desk. Cambridge University Press, Cambridge

Wendehorst C (2016) Die Digitalisierung des BGB. NJW, S. 2609 ff

Part II
Production

Industry 4.0: Agile Development and Production with Internet of Production

Günther Schuh, Michael Riesener, Andreas Gützlaff, Christian Dölle, Seth Schmitz, Johanna Ays, Shari Wlecke, Jonas Tittel, and Yuan Liu

Contents

1 Industry 4.0 As Enabler for Higher Agility in Manufacturing Companies

With increasing volatility of the global competitive environment, manufacturing companies need to be able to quickly adapt to changes, e.g., shortened product life cycles, increasing price pressure, higher customization, higher product complexity, new competitive situations or new technologies (Stark 2016). This ability for quick

G. Schuh · M. Riesener · A. Gützlaff · C. Dölle · S. Schmitz · J. Ays · S. Wlecke · J. Tittel · Y. Liu (✉)
Laboratory for Machine Tools and Production Engineering WZL of RWTH Aachen University, Aachen, Germany
e-mail: y.liu@wzl.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_19

reaction culminates in the company's "agility" which empowers fast, flexible and proactive activities (Lindner and Leyh 2018).

First agile approaches for product development were already formulated in the year 1986 when Takeuchi and Nonaka presented a scrum-method that targeted iterations in the product development to achieve faster and better results (Takeuchi and Nonaka 1986). With this approach, the authors deliberately turned away from the conventional, sequential development method. In the year 1992, the vision of the "agile manufacturing enterprise" was introduced in the USA (Nagel 1992). The authors argue that the predictable end of the mass production era will cause the beginning of a new era which will be characterized by high volatilities and in which agility will become a necessity. However, the ultimate breakthrough of agile principles was marked no earlier than the year 2001 by the "manifesto for agile software development" (Beck et al. 2001). The manifesto was formulated as a countermovement to the conventional, linear, and inflexible product development process. Within the software industry, the manifesto was quickly and widely adopted. Due to the increasing integration of software into physical products, agile principles also found their ways into the manufacturing industry. Manufacturing companies were faced with the challenge to integrate the traditional sequential development process of physical products with the non-linear agile principles of the software components. The general feasibility was shown by Karlström and Runeson through use case analyses at the companies ABB and Ericsson (Karlström and Runeson 2006). Cooper and Sommer developed the "agile-stage-gate" approach to combine the two different development approaches. It serves as a powerful tool for agile product development in manufacturing companies (Cooper and Sommer 2016). Through consequent implementation of the agile principles, companies can achieve a faster response as well as higher customer orientation, thus generating stronger competitive advantages.

The achievement of agility requires two substantial premises. On the one hand, the need for action must be recognized with minimal latency and ideally in advance. On the other hand, the consequences of the various potential countermeasures must be transparent to enable a fast and sound decision process. Both premises can only be achieved with a detailed and real-time knowledge of all relevant processes within the company. This can be achieved by interconnecting all objects and systems to form a cyber-physical system (Klötzer and Pflaum 2019).

This form of interconnectivity is generally known as the Internet of Things (IoT). The IoT has become the epitome of an Internet-based global network which has already dramatically changed our everyday life. The core characteristic of the IoT is the mass generation and use of data (Tsai et al. 2014). It enables precious insights which can be turned into new product and process innovations. The generation of Big Data within the IoT has two major advantages. First, the applications do not significantly differ from user to user. They are well comparable and highly scalable. Second, the data can be described by a relatively small number of parameters. However, in the manufacturing industry, the development and production conditions vastly differ from company to company and are therefore hardly comparable. The relevant processes can also only be described by a large number of parameters. This results in a high degree of complexity. Consequently, a direct application of the IoT principles in manufacturing industries, which is commonly known as "Industry 4.0," is nontrivial (Schuh et al. 2017e).

In manufacturing companies, the source of applicable data is often reduced to the company itself, in some cases, even to single factories only. As a result, companies require a long time to gather a relevant amount of data from which desirable data-based insights can be generated. Although many companies own a considerable amount of data without Industry 4.0, the data often lacks the required data integrity and can therefore not be leveraged. Against this background, it becomes essential for manufacturing companies to gather data to a fuller extent, to use the existing data intelligently and to combine data across different domains. Consequently, an appropriate data infrastructure is required.

2 The Internet of Production As Infrastructure for Industry 4.0

On the RWTH Aachen Campus, scientists from production technologies, computer science, material science and economics have developed the vision of the Internet of Production (short: IoP, see Fig. 1) (Schuh et al. 2017e). The goal is to achieve a new level of cross-domain collaboration with semantically correct and context-aware data. The collaboration is set to occur not only singularly, but continuously and iteratively with real-time data and with the adequate granularity. By doing so, the IoP specifically considers the circumstance that available data in manufacturing companies are stored in independent, proprietary systems. These systems can usually only be used by experts with specific knowledge within one domain. The IoP provides the appropriate data infrastructure to combine all relevant systems and thus enables a cross-domain data analysis (Schuh et al. 2017b).

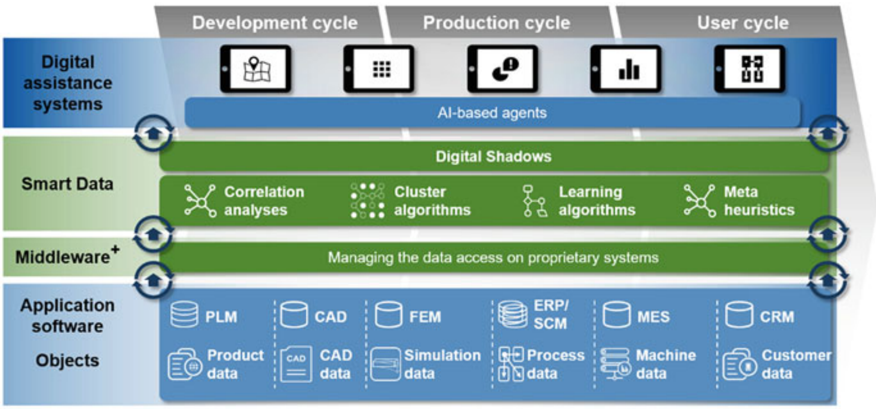


Fig. 1 The IoP as data infrastructure for Industry 4.0 (source: WZL of the RWTH Aachen)

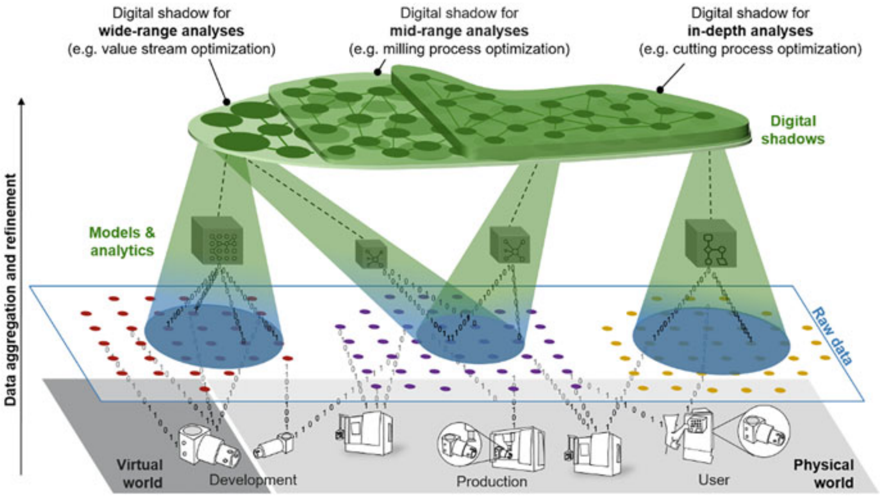


Fig. 2 Concept of digital shadows for the provision of information on the required aggregation level (source: WZL of the RWTH Aachen)

The IoP is horizontally divided into three areas along the product life cycle: development cycle, production cycle and user cycle. Vertically it is divided into the layers application software with the respective objects, Middleware+, Smart Data and digital assistance systems.

The vertical structure of the IoP distinguishes between different data granularities. The lowest layer describes a variety of *application software* which is often used in companies. Here, proprietary systems with respective objects are predominant, e.g., Product Lifecycle Management (PLM), Computer Aided Design (CAD), Finite-Element-Method (FEM), Enterprise Resource Planning (ERP), Manufacturing Execution System (MES), Customer Relationship Management (CRM), etc. The data sets of these systems may have different levels of completeness, correctness, semantic and granularity. To conduct cross-domain data analyses, an intermediate layer is required which can purposefully filter and connect the proprietary systems. Otherwise this can only be done with great manual effort (Atzori et al. 2010). Within the IoP, the *"Middleware+"* forms this intermediate layer. Through its automation, data can be provided with minimum latency and according to specific purposes and relevance (Smart Data).

The goal of the layer *Smart Data* is to create digital shadows which are sufficiently accurate real-time projections of production related models and correlations. In contrast to digital twins as exact images of the reality (Hehenberger and Bradley 2016), not all process details are projected, but only relevant information from heterogeneous data silos along the product life cycle (see Fig. 2) (Schuh et al. 2017b). This reduced data granularity leads to a significantly simplified and accelerated data processing. Digital shadows allow both historical and real-time analyses to identify optimization potentials and to verify experience-based hypotheses. By doing so, the above mentioned two premises for agility can be fulfilled. On the one hand, companies can identify the need for action fast and even proactively due to the holistic understanding of the production system. On the other hand, the consequences of different countermeasures can also be made transparent.

Digital assistance systems consume the high complexity degree of Smart Data and supports users in their decision processes. They reduce the complexity to what is essentially required without eliminating it. The design of intuitive user interfaces is key to high acceptance of the assistance systems. For example, apps in production can be designed in a similar way to consumer apps which are already widely common. The use of artificial intelligence (AI) in this layer is also intended for the purpose of smart decision support. Here, AI describes a toolbox with intelligent, computer-assisted agents. These agents are capable of solving problems more effectively and efficiently than humans (Russell and Norvig 2016). For example, they can learn from historic human decisions to improve future decision suggestions or even to automate the decision process for recurring problems.

The IoP offers the required data infrastructure to successfully implement Industry 4.0. The following chapters show how a higher agility can be achieved in product development and production. Finally, the concept of subscription-based business models is introduced which strongly effects the product application phase.

3 Agility in Product Development

As already introduced, the business environment in many industries is nowadays characterized as unstable which makes it difficult to predict market and technological changes. Simultaneously, the progressive networking of system components results in an increased complexity to be managed. This change is often described with the term VUCA (Volatility, Uncertainty, Complexity, Ambiguity) (Bennett and Lemoine 2014).

The generally formulated conditions of VUCA show visible effects also on the product development of manufacturing companies. Customers change their product requirements at short notices and demand individual solutions. Innovative products are crucial for the establishment of competitive advantages. After comparatively short periods of time, customers perceive formerly admired features only as performance features or basic features. This results in shorter product life cycles, which in turn require an acceleration of the development processes. At the same time, the dynamic business environment increases uncertainty in product development as well as variety of product functions (Schuh 2012).

Against this background, agile development processes are also becoming increasingly common in the manufacturing industry following the example of software development. Concrete weaknesses of plan-driven development processes in the manufacturing industry will be presented in the following sections. Based on this, the potentials and challenges of agile development processes are identified. Finally, we will explain how the IoP enables companies to shape agile development processes.

3.1 Weaknesses of Plan-Driven Development Processes

Plan-driven development processes, such as the now withdrawn but nevertheless widely applied industry standard VDI 2221 (VDI Verein Deutscher Ingenieure 1993) or the classic stage-gate process according to Cooper (Cooper 1990), describe common procedures for the development of physical products. Within plan-driven product development, the development task is divided into different work steps. Common sequential work steps are planning, conception, design and elaboration of the product to be developed (VDI Verein Deutscher Ingenieure 1993). The mentioned processes are to be executed in a given sequential order, and phases are separated by milestones according to the stage-gate approach. To proceed into the consecutive phase, the required degree of maturity must be proven based on predefined criteria. This procedure is proven for development projects with low to medium uncertainty. However, an agile response to the dynamic business and development environment is hardly possible (Cooper 2016).

The starting point of plan-driven product development processes is the creation of a complete specification sheet, which collectively documents the fixed requirements of the customer (Feldhusen et al. 2013a). A large part of the requirements is often transferred from already completed projects. Therefore, there is a risk of not sufficiently considering changed market conditions or new customer requirements due to a lack of customer integration into the development process. In industrial practice, corrections to the actual result are often necessary at a late stage of the plan-driven development process involving a high financial and time expenditure.

Against the background of the identified weaknesses, current plan-driven development processes are unsuitable for meeting the challenges described above.

3.2 Potentials and Challenges of Agile Product Development

To survive in competition, companies need to address the weaknesses described above. The revised version of VDI 2221 emphasizes a more iterative approach within development processes, but product development is still divided into individual phases by predefined milestones (VDI Verein Deutscher Ingenieure 2019). Targets set at the beginning of the development may therefore be pursued for too long without adjustment. The flexibility and the degree of customer integration of the product development process must be enhanced. At the same time, the development speed must be increased. A promising idea is the creation of a situation-specific ability to react to the uncertainties of product development with the help of agility. In this context, agility describes the ability of a company to adapt flexibly and iteratively to a given situation in a volatile and dynamic environment (Diels 2017).

In the course of formulating the Agile Manifesto in the early 21st century, experienced software developers prioritized individuals and interactions over processes and tools, working software over documentation, working with the customer over contract negotiation and responding to change over following a plan (Beck et al. 2001). The agile approach increases effectiveness and efficiency in software development (Serrador and Pinto 2015).

However, a direct transfer of agile methods from the software industry to the development of physical products in the manufacturing industry is not possible due to various restrictions. The manufacturing industry considers the marketable delivery of functional but not final technical products to the customer as not applicable (Schuh et al. 2018b). Instead, the use of prototypes for an intensive and short-cycle involvement of customers and other stakeholders is reconsidered (Böhmer et al. 2018). The existing plan-driven approach to test planning, prototype construction, test execution and test evaluation cannot easily be combined with the agile idea. In addition, agility is often hindered by fixed hierarchical levels and extensive approval loops within the existing organizational structure.

Another decisive restriction are human behavior patterns that have become established during the plan-driven processes and that make the introduction of an agile development more difficult. People generally tend to understand progress as the evolution of known solutions. Product innovations are therefore often derived from evolution of existing products (Murray 2011). Consequently, the solution space is limited, and existing innovation potentials are often neither recognized nor used. Furthermore, psychological research describes a human urge to complete activities that were started once. In product development, this results in the urge to specify the scope of products to be developed in detail and as early as possible, without gaining relevant customer knowledge in advance (Riesener 2017). Furthermore, the human brain filters and categorizes environmental impressions (Chabris and Simons 2010). The perception of reality is influenced by the collected subjective experiences. This results in a so-called silo thinking of the individual departments which hinders the interdisciplinary cooperation.

Due to the described restrictions, a simple transfer of the agile approach of the software industry to the product development of the manufacturing industry is not purposeful. It is rather recommended to transfer the basic agile values and principles and to develop a specific methodical approach based on those aspects. The implementation is supported by the IoP and is presented in the following section.

3.3 Agile Product Development in the Manufacturing Industry

Despite the restrictions described above, there are successful examples of agile product development processes in the manufacturing industry based on the technological framework of the IoP. Before explaining the principles of the IoP in the context of agile development processes, Fig. 3 presents an overview of the approach adopted by leading users of agile development processes. A distinction is made between the macro and micro level.

From the detailed documentation based on User Story Cards, numerous questions with relevance for product development arise. These questions result from incomplete requirements and describe the uncertainties of the properties and behavior of the product to be developed. According to Design Thinking, the objectives of the questions can be divided into three types: Desirability ("What do customers and markets need?"), Feasibility ("Can it be done from a technical or organizational perspective?") and Viability ("Can we make money doing it?") (Neck et al. 2018). Considering the strength of the uncertainty as well as the criticality of a question resulting from this uncertainty, these questions can be prioritized among themselves. By deriving product scopes directly from the prioritized questions, development activities can be synchronized in an agile development process. Since questions in technical development projects can usually only be answered across disciplines, cross-departmental collaboration is required.

In the context of the IoP, this interdisciplinary collaboration achieves a previously unimaginable productivity by resolving semantic conflicts. This requires comprehensive data availability as well as a fundamental systemic understanding of individual experts. Consequently, development data can be provided and explained to experts as knowledge which is reduced to the essential content for the respective application context. This enables an interdisciplinary approach to complex and multilayered questions.

At the macro level of agile product development, answering questions forms a so-called development cycle with a typical duration of a few months. A development cycle is concluded by the completion of a so-called Minimum Viable Product (MVP). Common definitions from software development describe an MVP as a preliminary version of the product which is implemented sufficiently to illustrate the value of the product for the user (Rancic Moogk 2012). Generally, this includes selling the MVP to early users. In the agile development of the manufacturing industry, an MVP is created for a specific part of the product. Therefore, the MVP cannot be sold but serves to answer the questions of the development cycle. In addition, the answer to a question is not necessarily provided by early users, but also with the involvement of internal stakeholders who belong to, for instance, production or sales. Insights and decisions for further development are derived from the analysis of the MVP. The conscious confrontation with undesirable developments as well as the acceptance of resulting changes in the development process require a paradigm shift in development organizations. This paradigm shift leads to the ability

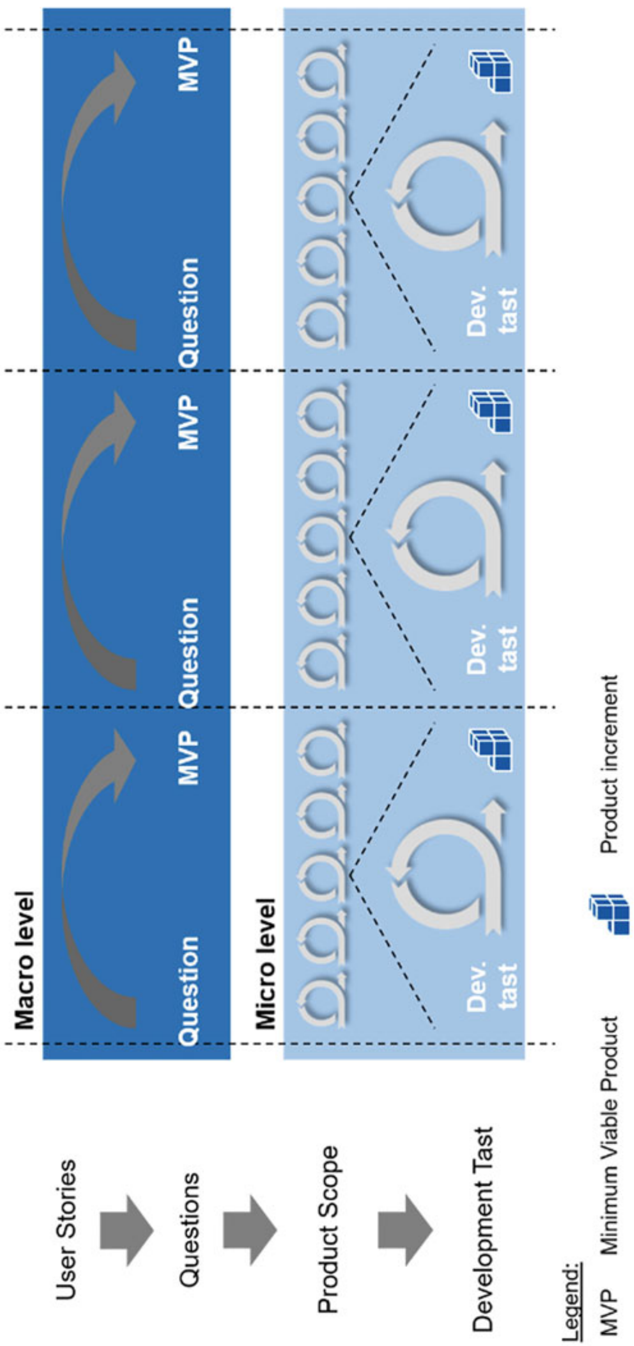


Fig. 3 Agile approach to product development in manufacturing companies (source: WZL of the RWTH Aachen)

of a company to quickly implement technical changes. The cross-discipline and cross-system data consistency realized in the IoP enables companies to implement rapid engineering change requests. The integration and implementation of short-cycle feedback from customers and other stakeholders for a consistent product implementation that meets the requirements is a logical consequence of the data and system consistency realized in the IoP. If the MVP shows that a concept is not feasible and a possible change process is not purposeful, an early termination of the project prevents financial waste. The risks of agile development processes are thereby reduced (Feldhusen et al. 2013b).

To answer questions at the macro level, concrete development tasks must be derived from the questions at the micro level. The tasks are integrated into an iterative and incremental process framework, which is based on the widely used Scrum approach according to Schwaber (Schwaber 2007). Development tasks are processed within sprints and result in the completion of associated product increments. The sprint period is usually limited to about 30 working days. In the sprint planning phase, the Product Owner responsible for the product determines the tasks to be processed together with the actual development team. The development team is composed interdisciplinary and works self-organized. Increased data and information availability through cross-system data consistency between different disciplines in the IoP increases the efficiency of the operative development work in the sprint by significantly reduced latency, search and waiting times. In addition, the IoP provides possibilities for ubiquitous communication even in globally distributed development networks, which makes a central contribution to the dissolution of the silo thinking established in many development organizations. In addition, the actual development work in the IoP is completely digitalized so that alternative solutions can be simulated, evaluated, and finally selected based on data. The product increment of a sprint is presented in the final Sprint Review as the content-related result of the development task. Product increments are realized in the IoP either physically by the availability of additive manufacturing technologies or virtually by means of visual methods in form of virtual reality.

The benefit of agile product development is therefore an increase in customer value through a stronger fulfillment of customer needs in combination with a reduction of time-to-market by reducing the development times. This is done by an increased flexibility in the development process and the associated possibility of an early reaction to undesirable developments due to changing customer needs. Fig. 4 illustrates the differences between plan-driven and agile product development in terms of the potential benefits.

4 Agility in Production

To benefit from agile product development and achieve a significantly reduced time-to-market, the production also needs to follow agile principles. First, the latency caused by prototype validation cycles during the iterative product development

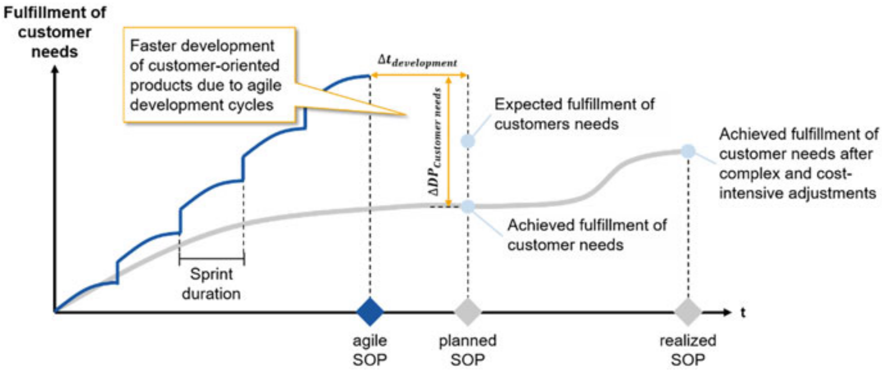


Fig. 4 Customer needs fulfillment in comparison of plan-driven and agile product development (Schuh et al. 2017a)

needs to be reduced. Second, the planning of the production and assembly processes need to be accelerated to achieve series production maturity within minimum time. Third, production needs to be generally enabled to react quickly and proactively to environmental changes and to achieve high productivity on a continuous basis (Schuh et al. 2017c). In this sequence, the following chapters describe how agility in production can be achieved.

4.1 Agility in Prototyping in Support of the Highly Iterative Product Development

One element of agile product development is the building of physical prototypes for various purposes. The transition to the production department requires the conversion of the engineering bill of material (eBOM) from CAD data to a manufacturing bill of material (mBOM) and the creation of production work sheets. These tasks are often based on subjective experience and therefore conducted manually (Curie 2018). They can take up to several days and cause long latencies for the entire development process.

Automated transition processes from the product development to the production department would significantly reduce the latencies. To do so, various data systems from development and production need to be interconnected. Although the automated conversion of eBOM to mBOM is generally possible and first commercial solutions are already available, they often struggle with the complexity of all relevant information (Schuh et al. 2017d). Therefore, the length (required parts) and the width (required attributes) of both the eBOM and mBOM still need to be manually secured. The automated creation of production work sheets through Industry 4.0 is also generally possible by matching potential product features with potential production tasks. However, it is still a major challenge in practice due to the variety of product features and the complexity in production (Kaipa et al. 2018).

The cross-domain data integration of the relevant systems within the IoP offers an appropriate infrastructure to meet these challenges. One solution is the application of machine learning (ML), a subarea of artificial intelligence, which is explicitly supported by the IoP. Based on historical data, ML agents can generate task specific insights (Russell and Norvig 2016). Its application on historic CAD data, eBOMs, mBOMs and production worksheets may help to automate the transition processes from product development to production. Furthermore, the IoP supports the recirculation of prototyping insights back to product development. Design flaws that are only discovered on prototypes can be visually documented with mobile apps and directly placed into development systems. By doing so, an agile data-based communication between prototyping and product development is enabled.

4.2 Agility in the Integrated Product and Production Development

An agile prototype development enables an accelerated integration of customer requirements, improved product quality and reduced time-to-market. It is followed by production and assembly planning. Due to the time intensity of the planning processes and to shorten the time-to-market, in industrial practice, these processes usually do not start strictly after the product development, but already in parallel. The goal is to place machine orders and to start recruiting processes as early as possible. In doing so, the highly iterative product development causes high planning uncertainties for the production development process. The short-term development sprints cause continuous product changes which need to be accommodated with direct long-term implications for the production processes. This conflict between late product changes and long-term, often irrevocable investment decisions poses great challenges for manufacturing companies. It often results in high costs, suboptimal production processes, and intense disputes. Thus, a holistic approach integrating both agile product and agile production development is necessary to create a sound long-term investment decision basis (Diels et al. 2015; Schuh et al. 2018a).

A solution to fulfil this need is offered by a customer-focussed concept which is characterized by a highly iterative and integrated product and production development (HIPPD), combined with an iterative validation approach (see Fig. 5). In the development phase, the maturity degree of the product is continuously increased while the degree of freedom of the production process is decreased due to different product and production hypotheses. Product and production hypotheses are developed together and represent an integrated conception of the product and the production system throughout the development cycle. The hypotheses consist, for example, of product information such as bill of materials and dimensions as well as production information such as machines, operating times, and logistical factors. Within the

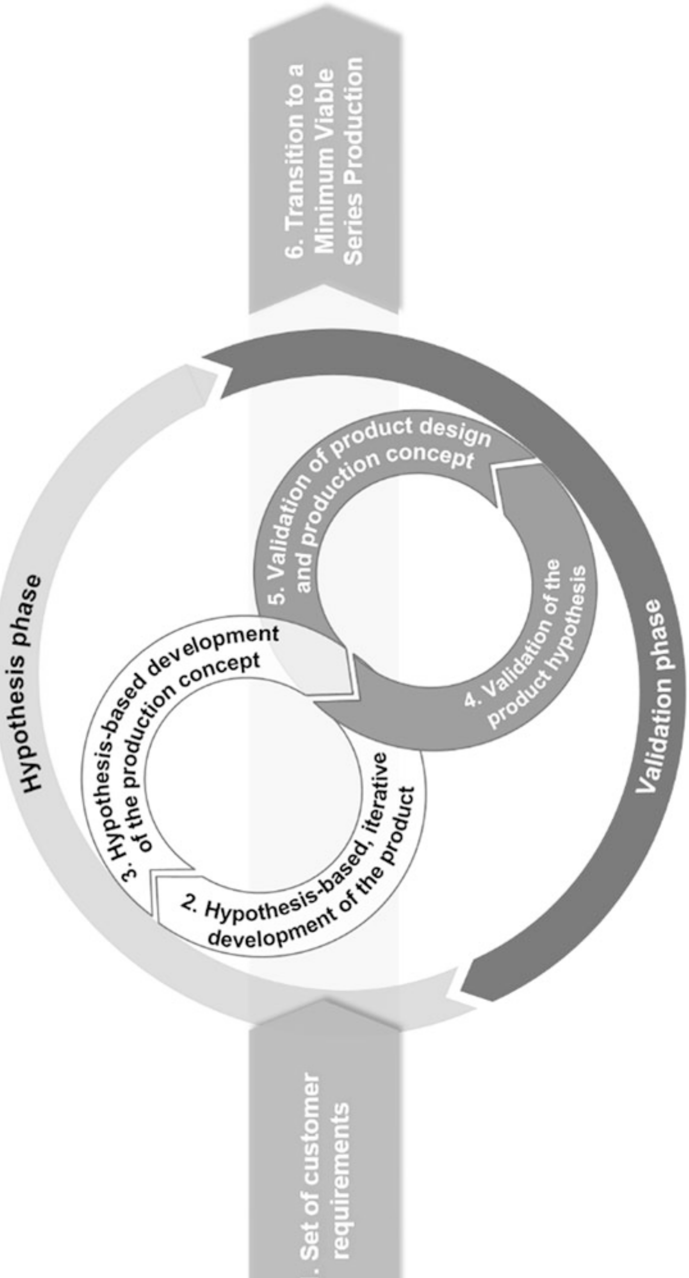


Fig. 5 Approach for the highly iterative and integrated product and production development (HIPPD) (source: WZL of the RWTH Aachen)

integrated hypotheses, all interdependencies between the product and production process development domains are considered to always ensure a transparent overview of each hypothesis. These interdependencies are modelled to track the consequences of a change in a product or production hypothesis and to enable a proactive improvement of both domains simultaneously. Upon reaching a maturity degree that is ready for validation, the integrated product and production hypotheses are transferred into the validation cycle, where prototyping or simulations are used. The development and validation cycles are carried out until the product has reached its maximum degree of maturity and the production system the minimum degree of freedom, i.e., the pre-series prototype is validated, and the series production system is detailed. This state is the core enabler for the transition to a Minimum Viable Series Production (MVSP). The idea of a MVSP allows the early production of pre-series prototypes under series conditions and creates higher flexibility for adaptations. The MVSP describes a production state, optimized in terms of capacity utilization with a minimum number of units and minimum use of resources. On this basis, the first Minimum Viable Products (MVPs) can be produced and brought to market, thus generating sales, and allowing further changes to both the product and the production process. In the MVSP, the planned machines, plants, and employees represent the potential solution area for the internal depth of added value. Within this solution field, it is necessary to determine the minimum number of production factors that allow individual MVPs to be manufactured close to series production. On the machines and plants of the MVSP, strategic degrees of freedom are considered to convert agile occurring product and production-conditioned change requirements. The MVSP is continuously optimized until the output quantity can no longer be increased or reaches its maximum. Based on this maximum output quantity of the MVSP, the temporarily completed series production planning results are checked and adjusted if necessary. In this way, the MVSP enables the efficient implementation of product changes and further series production development even in a high degree of production process maturity, e.g., during pre-series production, and thus enables an optimal investment decision basis (Wlecke et al. 2019).

Through the highly iterative and integrated product and production development (HIPPD), all relevant product changes can be quickly transferred to the production development in intended systematic iterations at an early stage. This avoids significantly more complex and expensive iterations at later stages which could lead to suboptimal compromises and production solutions. By doing so, the HIPPD enables not only a reduced time-to-market but also minimal production costs. Therefore, the continuous cost tracking and forecasting is a central element in the HIPPD. This cost analysis is strongly supported by the infrastructure of the IoP. In the user cycle, customer requirements are continuously identified and translated into user stories. These are used to derive target costs for the product structure and serve as input information for the development cycle. Total costs are identified in the iterative loops between the development and production cycle by defining process, procurement, and equipment costs in an end-to-end cost view. This enables a holistic, yet agile approach for the cost calculation. Therefore, the application of HIPPD in the context of Industry 4.0 and IoP enables the continuous consideration of volatile customer requirements and product specifications, while providing the necessary cost transparency and manage the complex interdependencies in the integrated development (Schuh et al. 2020a).

4.3 Agility in Series Production Through Real-Time Decision Support

As previously described, the IoT principles cannot be directly transferred to the manufacturing industry due to the overall high complexity in production systems. This complexity results from the interdependencies of a vast number of parameters and is even increased by external influences. The medium-term production planning (production demand planning) is directly impacted by customer behavior, general economic and seasonal effects (Schuh and Stich 2012; Nyhuis 2018). While these impacts can be partly anticipated, numerous unplanned disturbances may also impact the production, resulting in the need to quickly adapt the short-term production planning (in-house production planning). Typical unplanned disturbances include machine breakdowns, tool breakdowns and staff shortages (Schuh and Stich 2012). Regardless of the production type (line production, group production or job shop production), 28%–43% of all companies face more than three disturbances each day with an interruption of the production for more than 10 min (Niehues 2016). The impact of these disturbances on the production are dramatically increased with growing number and complexity of the production processes and production networks (Geisberger and Broy 2012). Therefore, the ability to quickly adapt to disturbances is an effective competitive advantage for companies.

The first premise for fast adaptation and higher agility is to recognize the need for action. A high degree of real-time transparency supports companies to quickly identify or even forecast disturbances (Schmitz and Krenge 2014). Today's sensors and systems are already capable of monitoring numerous machine parameters to determine the probability of the next breakdown and to maximize machine availability through predictive maintenance (Auf der Mauer et al. 2019). However, other complex correlations in production may remain hidden from the human intellect. They may be discovered by various means of data analytics to identify and observe further critical parameters and to sustainably increase the process stability. While doing so, the infrastructure of the IoP can be used to support the data collection and integration into digital shadows.

Yet some disturbances are difficult to anticipate in advance despite all available technologies, especially short-term disturbances as, e.g., staff absence, delays in logistics, short-term technical and logistical changes of customer demands, as well as emergency production orders (Steinlein et al. 2020). The potential countermeasures are numerous. They include additional shifts, adaptation of order schedules, flexible work times, use of additional resources, batch size split, etc. While selecting from these potential countermeasures, production engineers need to consider their

Fig. 6 Simulation-based "EkuPro" tool for the objective selection of countermeasures upon disturbances in production (source: WZL of the RWTH Aachen)

impacts on the overall production. These may include higher costs, delivery delays for other orders, the resulting conventional penalties, increased work in progress (WIP), increased storage, decreased quality, etc. Therefore, upon unplanned disturbances, it becomes even more important to gain transparency on the impacts of the disturbances as well as the consequences of potential countermeasures as fast as possible.

This ability forms the second premise for higher agility in manufacturing companies. Today, the countermeasures are usually selected based on the experience and subjective evaluation of single employees. Although the human experience should not be underestimated, it may lead to non-optimal decisions and therefore to avoidable negative impacts. Despite modern planning software, e.g., Advanced Planning and Scheduling (APS), an automated selection of optimal countermeasures is barely supported. With Industry 4.0 and the IoP, appropriate decision support systems become possible. The holistic analysis of the impacts of potential countermeasures is enabled due to the interconnection of various data systems to form purposeful digital shadows. On the level of digital assistance systems, only relevant information is presented to the user to identify the optimal countermeasures.

The WZL of the RWTH Aachen University has developed the tool "EkuPro" which supports the objective selection of countermeasures while also leveraging the human experience (see Fig. 6). Upon unplanned disturbances, a defined interface extracts all relevant information from the company's data systems, e.g., ERP and MES, to depict the current status of the production system. Within the tool, the user can configure different, potentially promising countermeasures and combine them in different scenarios. Upon activating the simulation, the scenarios are automatically transferred and simulated in the tool Siemens Plant Simulation. The simulation results are generated within a short period and visualized in the "EkuPro" tool. Based on the results, the user can analyze the impacts of each scenario and make an objective decision (Schuh et al. 2019). In the future, the use of AI can potentially automate and improve the configuration of scenarios, as well as the decision-making process based on the simulation results.

5 Agility Through New Business Models: Subscription in the Machinery and Equipment Industry

To sustain in the increasingly fierce global competition, manufacturers, especially from high wage countries, cannot rely on traditional technological advantages alone. This is especially true for the machinery and equipment industry in which both the global revenue share and the technological advantage of companies from high wage countries are diminishing (VDMA 2019; Dutta et al. 2019). Instead of product and process innovations, studies show that business model innovations are more successful in the long-term (Lindgardt et al. 2009). In recent years, subscription business models have been especially impactful in creating new champions and changing the competitive landscape, e.g., in the software and entertainment industries. The essence of their success is a consequent customer focus in all business activities which creates a stronger long-term customer relationship (Tzuo and Weisert 2018). With Industry 4.0 and IoP, subscription business models may also be applied in the machinery and equipment industry. Through stronger customer relationships and continuous data-based interaction, manufacturers can achieve higher agility in fulfilling even volatile customer requirements. In the long-term, they can support manufacturers from high-wage countries to extend their competitive differentiation.

Traditionally, subscription business models describe recurring payments for recurring values. They were first applied in the book and newspaper industry in the seventeenth century (Gassmann et al. 2014). Today's successful subscription models emphasize the importance of consequent customer focus and therefore suggest a wider definition to capture the nature of their success while also considering industry-specific characteristics. For the machinery and equipment industry, the definition is shown in Fig. 7 (Liu et al. 2021).

Subscription business models in the machinery and equipment industry describe a long-term relationship between suppliers and customers who receive recurring values for recurring payments. Based on consequent customer centricity, all generated values are designed towards the customers' problems and needs, especially to a continuous increase of their productivity. The values are transported not only by physical products, but also by all digital products and services that are required and desired by the customers.

Fig. 7 Definition of subscription business models in the machienry and equipment industry

5.1 Potentials of Subscription in the Machinery and Equipment Industry

Subscription models in the machinery and equipment industry can potentially create manifold benefits for both customers and manufacturers. The essential benefits for customers can be described by five core value propositions (Liu et al. 2021). The consequent customer focus of subscription results in the continuous innovation of subscription offers. With the manufacturers' and customers' know-how combined, a *continuous productivity increase* of the machines and equipment can be achieved by the collaboration of both parties together. Since all customers are different, manufacturers need to create *customized offers*. This can be done by either customizing the physical and digital products as we as services, or by advising customized combinations of standardized offers and thus maintain an economical cost structure. With subscription, customers can also achieve higher business *flexibility*. Manufacturers need to allow their customers to change between products and services which consequently need to be designed modularly. Customers also enjoy a significant financial flexibility since capital expenditures (CAPEX) for purchasing machines are shifted to operational expenditures (OPEX). Depending on the degree of customer flexibility that is offered, and depending on the revenue model, the customers' business risks are partly transferred to the manufacturers. This *risk sharing* needs to be reflected in a smart pricing strategy. Also, it creates an intrinsic motivation for manufacturers to improve their offers and the customers' productivity continuously (Schuh et al. 2020b). However, manufacturers need to be aware to carefully select customers who are eligible for subscription and rule out those with worrisome economic prospects. Eventually, the subscription offers need to be provided to the customers in a *convenient* way so customers can focus more on their core value-add processes. This not only involves a user-friendly design of physical and digital products. Manufacturers can also carry out the customers' non-value-add processes completely. For example, with predictive maintenance, customers can be completely disburdened from the planning and conduction of maintenance tasks.

The benefits for manufacturers are also manifold. It is scientifically proven that customers show an higher willingness to pay for integrated solutions to solve their problems (Tukker 2004). By doing so, manufacturers can extend their competitive differentiation, increase the customer satisfaction, and establish a sustainable customer relationship. In the long-term, this may lead to stable revenue streams and, over time, to higher aggregated overall profitability compared to one-time transaction-based revenues. With Industry 4.0 and the user cycle of the IoP, manufacturers can gather data on how customers interact with the machines and gain a better understanding of the customers' implicit needs. These insights may reveal adaptation needs of the subscription offers to continuously achieve high customer satisfaction. This way, manufacturers can agilely fulfil even volatile customer requirements. The data-based learning can also be used to increase the revenues by boosting both upselling (increase the utilization of an offer category) and cross-selling (sale of additional offer categories) (Lah and Wood 2016). A profitable example for upselling is the offer of consumable materials of the machines, e.g., papers for industrial printing machines or screws and anchors for construction equipment. Often the choice of the consumables may affect the machine performance. Due to the detailed knowledge about the machines, manufacturers may determine the optimal consumables and contribute to higher machine productivity.

The data infrastructure of the IoP can strongly support the implementation of subscription models in the machinery and equipment industry. To fulfill the value propositions economically, a data-interconnection of machines and machine components is required (Herterich et al. 2015). Not only can manufacturers analyze the customers' interaction with the machines, but the customers can also gain higher transparency of their processes. To do so, the user cycle of the IoP offers the required infrastructure. For high complexity analyses, algorithms based on AI can be used. The horizontal connection with the development cycle enables a systematic recirculation of generated insights. These insights can be used to develop of the next generation of machines and to increase the product innovation advantage, which eventually may also support the one-time sale of machines.

5.2 Challenges in Implementing Subscription Business Models

Manufacturers may face major challenges in implementing subscription business models. First, a company-specific subscription business model needs to be developed (Liu et al. 2021). The core value propositions serve as a guideline and may need to be prioritized and extended if necessary. Companies also need to develop specific offers which can contribute to these value propositions. As a major part of any business model, the pricing strategy may form the most difficult task in developing subscription models (Tzuo and Weisert 2018). Traditionally, manufacturers tend to calculate sales prices based on costs and intended margins. This approach may be difficult in subscription due to the spread of revenue streams over time and due to the partly intangible value propositions, e.g., risk sharing. Instead, manufacturers need to develop an approach for a value-based pricing strategy.

Another significant challenge in implementing subscription models results from the recurring revenue streams. Compared to one-time sales, the profit is not generated instantly but only after a long period. This causes a significant drop of the company's overall revenues in the first few years while also facing increased costs, e.g., for adapting the IT infrastructure in the sales, finance, or service departments. Therefore, it is utterly important to secure the company's financial liquidity, e.g., by forming an intermediate subsidiary with a strong financial partner.

Manufacturers also face major organizational challenges which eventually result in the need to adapt the organizational structure. Service and after-sales activities, which are often seen as complementary business in the machinery and equipment industry, will become key business activities in subscription models (Gölzer and

Gepp 2016; Isaksson et al. 2009). At the same time, data analytics functions also need to gain significant priority to derive product and service innovations from customer data. Furthermore, the need for consequent customer focus requires a process-oriented organizational structure although most companies of the machinery and equipment industry today are functionally structured.

Further challenges may occur in customer interaction. Industrial companies are traditionally very reluctant to share data that can potentially reveal their performance and know-how. Many customers need to be first convinced from the benefit of new services before granting access to their data. This also restricts manufacturers in equipping their machines with technologies for collecting data in the first place (Kampker and Stich 2019). Therefore, a dedicated data sharing contract needs to be placed with the customers. Only then can manufacturers learn from customer data and improve their offers over time.

Despite all challenges, the potential benefits of subscription models may outweigh the potential downsides. With Industry 4.0 and the IoP, subscription providers have the chance to fulfill customer needs better and to sustain in the increasingly fierce global competition. The systematic increase of the machines' productivity may reduce the overall required number of machines on the global market and lead to significant market consolidation. Therefore, manufacturers who can quickly introduce subscription models may greatly benefit from first mover advantages.

6 Summary

Industry 4.0 is the key enabler for higher agility in manufacturing companies. Originating from the software industry, agile principles have been adopted by the manufacturing industry to react to the increasingly volatile global competition and due to the increasing integration of software components in physical products. With the Internet of Production (IoP), an appropriate data infrastructure for the successful implementation of Industry 4.0 can be used. The IoP is the result of an interdisciplinary development of the RWTH Aachen Campus and continuously refined within the German Excellence Cluster "Internet of Production." Its implementation supports manufacturing companies to act agilely and successfully even in highly volatile competitive environments.

During the product development, traditional plan-driven development processes have only limited effectiveness and efficiency due to volatile customer requirements and short product life cycles. As an alternative approach, agile product development principles can be applied which aim at increasing customer benefits while simultaneously reducing development times. The interdisciplinary and cross-system data consistency in the IoP offers the technological prerequisites to realize agile product development in the manufacturing industry.

To harvest the benefits of agile product development and realize a significantly shorter time-to-market, the production also needs to apply agile principles. Based on the cross-domain data interconnection between the IoP's development and production cycle, latencies in the prototyping processes can be reduced to support a highly iterative product development. The application of agile principles for the downstream iterative production planning process solves potential conflicts between late product changes and the resulting high costs from late production planning changes. The integration of product and production development into the presented HIPPD approach leads to overall reduced planning time and costs. It forms a data-based approach and utilizes the infrastructure of the IoP for modelling the interdependencies of the two development domains. Within the series production, the analysis of all relevant processes enables an agile identification of needs for action upon unplanned disturbances, as well as the objective selection of optimal countermeasures. To do so, digital shadows for the processes need to be created within the IoP. This way, production employees are supported objectively in their decision processes.

The introduction of subscription business models based on Industry 4.0 and IoP is an innovative way to create additional competitive advantages for both machine manufacturers and customers. Vertical and horizontal data interconnections form the basis to continuously improve the customers' productivity. With subscription models, manufacturers especially from high-wage countries have the chance to establish sustainable customer relationships and thus to sustain in the increasingly fierce global competition.

The application of Industry 4.0 and IoP for product development, production and business models enables manufacturing companies to achieve an overall higher agility throughout the entire product lifecycle. By doing so, companies are capable to react and to sustain in the increasingly fierce and volatile global competition.

References

Atzori L, Iera A, Morabito G (2010) The internet of things. A survey. Comput Netw 54(15): 2787–2805

Auf der Mauer M et al (2019) Applying sound-based analysis at Porsche production: towards predictive maintenance of production machines using deep learning and internet-of-things technology. In: Urbach N, Röglinger M (eds) Digitalization cases. How organizations rethink their business for the digital age. Springer, Cham, pp 79–97

Beck K et al (2001) Manifesto for Agile Software Development. http://agilemanifesto.org/, zuletzt geprüft am 07.01.2019

Bennett N, Lemoine GJ (2014) What a difference a word makes. Understanding threats to performance in a VUCA world. Bus Horiz 57(3):311–317

Böhmer AI et al (2018) Towards agile product development - the role of prototyping. In: Maier A et al (eds) Design methods and tools. Curran Associates Inc., Red Hook, pp 1–10

Chabris CF, Simons D (2010) The invisible gorilla. And other ways our intuitions deceive us, 1. paperback ed. Broadway Paperbacks, New York

Cooper RG (1990) Stage-gate systems. A new tool for managing new products. Bus Horiz 33(3): 44–54

Cooper RG (2016) Agile–stage-gate hybrids. Res-Technol Manage 59(1):21–29

Cooper RG, Sommer AF (2016) Agile-Stage-Gate: new idea-to-launch method for manufactured new products is faster, more responsive. Ind Market Manage 59(2016):167–180

Curie C (2018) Why automation is necessary for your EBOM and MBOM. Manually keeping track of manufacturing process planning changes is no way to create a digital thread and digital twin. Ind Syst Eng Work 50(8):40–42

Diels F (2017) Indikatoren für die Ermittlung agil zu entwickelnder Produktumfänge. Dissertation. WZL der RWTH Aachen; Fraunhofer-Institut für Produktionstechnologie IPT

Diels F, Rudolf S, Schuh G (2015) Highly iterative product development process for engineering projects. AMM 794:532–539

Dutta S, Lanvin B, Wunsch-Vincent S (2019) Global Innovation Index 2019. Stronger Innovation Linkages for Global Growth. Cornell University; INSEAD; World Intellectual Property Organization, Geneva

Feldhusen J et al (2013a) Vorgehen bei einzelnen Schritten des Produktentstehungsprozesses. In: Feldhusen J, Grote K-H (eds) Pahl/Beitz Konstruktionslehre. Methoden und Anwendung erfolgreicher Produktentwicklung. Springer, Berlin, pp 291–409

Feldhusen J, Grote K-H, Thon J (2013b) Grundsätzliche Überlegungen zur Rationalisierung. In: Feldhusen J, Grote K-H (Hrsg) Pahl/Beitz Konstruktionslehre. Methoden und Anwendung erfolgreicher Produktentwicklung. Springer, Berlin, pp 773–816

Gassmann O, Frankenberger K, Csik M (2014) The business model navigator. 55 models that will revolutionise your business. Pearson, Harlow

Geisberger E, Broy M (3/2012) agendCPS. Integrierte Forschungsagenda Cyber-Physical Systems. acatech, München

Gölzer P, Gepp M (2016) Anlagenbauer reagieren zurückhaltend auf aktuelle Trends im Engineering - eine empirische Studie. In: Schulze S-O, Muggeo C, Abulawi J (Hrsg) Tag des Systems Engineering. Ulm, 11.-13. November 2015. Hanser. München, pp 205–214

Hehenberger P, Bradley D (2016) Mechatronic futures. Springer International Publishing, Cham

Herterich M, Uebernickel F, Brenner W (2015) Nutzenpotentiale cyber-physischer Systeme für industrielle Dienstleistungen 4.0. HMD 52(5):665–680

Isaksson O, Larsson TC, Rönnbäck AÖ (2009) Development of product-service systems. Challenges and opportunities for the manufacturing firm. J Eng Design 20(4):329–348

Kaipa KN et al (2018) Toward automated generation of multimodal assembly instructions for human operators. In: Madni AM et al (eds) Disciplinary convergence in systems engineering research. Springer International Publishing, Cham, pp 885–897

Kampker A, Stich V (2019) Service-Innovation. 6 Prinzipien für erfolgreiche, datenbasierte Service-Innovation in Industrieunternehmen. Whitepaper. Unter Mitarbeit von Marco Husmann, Tobias Harland und Philipp Jussen. FIR e.V. an der RWTH Aachen, Aachen

Karlström D, Runeson P (2006) Integrating agile software development into stage-gate managed product development. Empir Softw Eng 11(2):203–225

Klötzer C, Pflaum A (2019) Cyber-Physical Systems (CPS) als technologische Basis einer digitalen Supply Chain der Zukunft. In: Becker W et al (eds) Geschäftsmodelle in der digitalen Welt. Springer Fachmedien Wiesbaden, Wiesbaden, pp 381–396

Lah T, Wood JB (2016) Technology-as-a-Service Playbook. How to Grow a Profitable Subscription Business. Point B Inc, s.l

Lindgardt Z et al (2009) Business model innovation. When the games gets tough, change the game. The Boston Consulting Group

Lindner D, Leyh C (2018) Organizations in transformation: agility as consequence or prerequisite of digitization? In: Abramowicz W, Paschke A (eds) Business information systems: 21st international conference, BIS 2018, Berlin, Germany, July 18-20, 2018. Proceedings. Springer, Cham, pp 86–101

Liu Y et al (2021) Design of tailored subscription business models – A holistic guide for machinery and equipment manufacturers. In: Behrens B-A et al (Hrsg) Production at the leading edge of technology. Proceedings of the 10th Congress of the German Academic Association for Production Technology (WGP), Dresden, 23–24 September 2020. Springer, Berlin, in press

Murray PN (2011) How Steve Jobs Knew What You Wanted. Can market research predict consumer response to new technology? https://www.psychologytoday.com/intl/blog/

inside-the-consumer-mind/201110/how-steve-jobs-knew-what-you-wanted, zuletzt geprüft am 14.01.2019

Nagel RN (1992) 21 st century manufacturing enterprise strategy report. Iacocca Institute, Lehigh University, Arlington, Virginia

Neck HM, Neck CP, Murray EL (2018) Entrepreneurship. The practice and mindset. SAGE, Los Angeles

Niehues MR (2016) Adaptive Produktionssteuerung für Werkstattfertigungssysteme durch fertigungsbegleitende Reihenfolgebildung. Dissertation. Technical University Herbert Utz Verlag GmbH, München

Nyhuis P (Hrsg) (2018) PPS-Report 2017/18. Studienergebnisse, 1. Erstausgabe. TEWISS, Garbsen

Rancic Moogk D (2012) Minimum viable product and the importance of experimentation in technology startups. Technol Innovat Manage Rev 2(3):23–26

Riesener M (2017) Agile Produktentwicklung. In: Schuh G et al (Hrsg) Engineering valley - Internet of production auf dem RWTH Aachen Campus. Festschrift für Univ.-Prof. em. Dr.-Ing. Dipl.-Wirt. Ing. Dr. h.c. mult. Walter Eversheim. 1. Auflage. Apprimus Verlag. Aachen, pp 13–28

Russell SJ, Norvig P (2016) Artificial intelligence. A modern approach, 3rd edn. Global edn. Pearson, Boston

Schmitz S, Krenge J (2014) Big Data im Störmanagement. In: Plödereder E et al (Hrsg) Informatik 2014. Big Data - Komplexität meistern: 22.-26. September 2014, Stuttgart: proceedings. Gesellschaft für Informatik. Bonn, pp 1229–1240

Schuh G (2012) Innovationsmanagement. Handbuch Produktion und Management 3, 2., vollst. neu bearb. und erw. Aufl. Springer, Berlin

Schuh G, Stich V (Hrsg) (2012) Produktionsplanung und -steuerung, 4., überarbeitete Auflage. Springer Vieweg, Berlin

Schuh G et al (2017a) Enabling agility in product development through an adaptive engineering change management. Procedia CIRP 63:342–347

Schuh G et al (2017b) Generative Stücklistenerstellung in der manuellen Montage. ZWF 112(6): 392–395

Schuh G et al (2017c) Agile Produktentwicklung. In: Brecher C et al (Hrsg) Internet of Production für agile Unternehmen. AWK Aachener Werkzeugmaschinen-Kolloquium 2017. 1. Auflage. Apprimus Verlag. Aachen, pp 29–51

Schuh G et al (2017d) Change Request im Produktionsbetrieb. In: Brecher C et al (Hrsg) Internet of Production für agile Unternehmen. AWK Aachener Werkzeugmaschinen-Kolloquium 2017. 1. Auflage. Apprimus Verlag. Aachen, p 131

Schuh G, Prote J-P, Dany S (2017e) Internet of production. In: Schuh G et al (Hrsg) Engineering valley - Internet of production auf dem RWTH Aachen Campus. Festschrift für Univ.-Prof. em. Dr.-Ing. Dipl.-Wirt. Ing. Dr. h.c. mult. Walter Eversheim. 1. Auflage. Apprimus Verlag. Aachen

Schuh G et al (2018a) Process model for generative assembly planning in the highly iterative product development. Procedia CIRP 72:363–368

Schuh G, Dölle C, Schloesser S (2018b) Agile Prototyping for technical systems – Towards an adaption of the Minimum Viable Product principle. In: Ekströmer P, Schütte S, Ölvander J (Hrsg): DS 91: Proceedings of NordDesign 2018, Linköping, Sweden, 14th - 17th August 2018. Linköping University (NordDESIGN)

Schuh G et al (2019) Verbesserung der Liefertermintreue durch Simulation. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 114(12):819–822

Schuh G et al (2020a) Concept for the cost prognosis in the industrialization of highly iteratively developed physical products. In: Nyhuis P, Herberger D, Hübner M (Hrsg) Proceedings of the 1st Conference on Production Systems and Logistics (CPSL 2020). Publish-Ing. Hannover, pp 99–109

Schuh G et al (2020b) Subskriptionsmodelle im Maschinen- und Anlagenbau. In: Internet of Production - Turning Data into Value, pp 282–301

Schwaber K (2007) Agiles Projektmanagement mit Scrum, dt. Ausg. Microsoft Press, Unterschleißheim

Serrador P, Pinto JK (2015) Does Agile work? — A quantitative analysis of agile project success. Int J Project Manage 33(5):1040–1051

Stark J (2016) Product lifecycle management, Vol 2. The Devil is in the details, 3rd edn. Springer, Cham

Steinlein F, Liu Y, Stich V (2020) Development of a decision support app for short term production control to improve the adherence to delivery dates. In: Nyhuis P, Herberger D, Hübner M (Hrsg) Proceedings of the 1st Conference on Production Systems and Logistics (CPSL 2020). Publish-Ing. Hannover, pp 438–447

Takeuchi H, Nonaka I (1986) The new product development game. Harv Bus Rev 1986:137–146

Tsai C-W et al (2014) Data mining for internet of things: a survey. IEEE Commun Surv Tutor 16(1): 77–97

Tukker A (2004) Eight types of product–service system. Eight ways to sustainability? Experiences from SusProNet. Bus Strat Env 13(4):246–260

Tzuo T, Weisert G (2018) Subscribed. Why the subscription model will be your company's future - and what to do about it. Portfolio/Penguin, New York

VDI Verein Deutscher Ingenieure (Mai 1993) Methodik zum Entwickeln und Konstruieren technischer Systeme und Produkte. VDI-Richtlinien 2221, Mai 1993

VDI Verein Deutscher Ingenieure (November 2019) Entwicklung technischer Produkte und Systeme. VDI-Richtlinien 2221 Blatt 1, November 2019

VDMA (2019) Maschinenbau in Zahl und Bild 2019. VDMA Verband Deutscher Maschinen- und Anlagenbau e.V., Frankfurt am Main

Wlecke S et al (2019) Concept for the industrialization of physical products in the highly iterative product development. In: Wulfsberg JP, Hintze W, Behrens B-A (Hrsg) Production at the leading edge of technology. Springer, Berlin, pp 583–592

Data-Based Quality Management in the Internet of Production

Chair of Production Metrology and Quality Management

Robert Heinrich Schmitt, Max Ellerich, Peter Schlegel, Quoc Hao Ngo, Dominik Emonts, Benjamin Montavon, Daniel Buschmann, Rebecca Lauther, Nils Klasen, and Raphael Kiesel

Contents

1 Introduction

Today the implementation of Industry 4.0 is shaping the competition of manufacturing companies on global markets. Those who want to meet their competitors on an equal footing are permanently required not only to record all available information securely and in real-time but also to process it to realize a precise and continuous analysis. Having the right information available at any time is an enormous

R. H. Schmitt (✉)
Laboratory for Machine Tools and Production Engineering WZL of RWTH Aachen University, Aachen, Germany

Fraunhofer Insitute for Production Technology, Aachen, Germany
e-mail: R.Schmitt@wzl.rwth-aachen.de

M. Ellerich · P. Schlegel · Q. H. Ngo · D. Emonts · B. Montavon · D. Buschmann · R. Lauther · N. Klasen · R. Kiesel
Laboratory for Machine Tools and Production Engineering WZL of RWTH Aachen University, Aachen, Germany

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_20

challenge. To this end, it is essential to recognize patterns in the data stream, to learn from them, and to be able to derive and implement the right forecasts for the company, processes, and production.

The **Internet of Production** (IoP) describes real-time, secure information available at any time and any place. Precise and continuous data analysis, pattern recognition, and reliable decision-making support the production systematically and sustainably. Agile and highly iterative product development becomes just as possible as the fast, error-free response to individual customer requests and product changes within series production (Brecher et al. 2017).

The IoP infrastructure shown in Fig. 1 consists of four underlying layers: The **Raw Data** layer as well as the raw data access via the respective Application Software, the **Middleware+** for managing the data access to different proprietary systems, the **Smart Data** layer for generating knowledge based on the Digital Shadow, and the **Smart Expert** layer on which the domain-specific use of the aggregated knowledge takes place.

With data-based quality management now located in the IoP, quality-relevant data gets collected in real-time from various data sources at the raw data level. Quality-relevant data in this context includes not only quality data but also process or sensor data from the respective production unit as well as product and order data (Schmitt and Pfeifer 2015). As in currently used working environments, the Application Software with the respective proprietary raw data remains on this lowest infrastructure layer. On the one hand, the Application Software's functional scope cannot simply be replaced and, on the other hand, should not be changed due to the high migration effort and the lack of release capability. The Application Software will also find use in the future infrastructure for low-interface processes. Within the framework of the IoP, however, information from different domains is now to be analyzed together. For this purpose, the Application Software level gets extended by three superordinate layers.

Due to the system-side diversity of Application Software in manufacturing companies, access via a Middleware+ and the subsequent preparation of data require a high initial effort in terms of cleansing, aggregation, filtering, contextualization, and synchronization. Within this data preprocessing scope, a **Digital Shadow** is created by selecting or extracting informative, differentiable, and independent **quality features** based on which subsequent modeling of the product and process quality in higher levels of the IoP is made possible. In this context, the term Digital Shadow describes the combination of preprocessed data and merely copied raw data and thus the relevant image of reality in real-time (Bauernhansl et al. 2016). The inherent knowledge can now be extracted to use this contextual information, among other things, through interpretation utilizing **data analytics** at the Smart Data level. Within the meaning of quality management, this knowledge can be used at the Smart Expert level to diagnose the quality of the product and the process in real-time (**diagnostic quality**), to forecast it (**predictive quality**), and ultimately to control it proactively (**prescriptive quality**). This enables knowledge-based decision support for humans in various quality management areas through the end-to-end use of data.

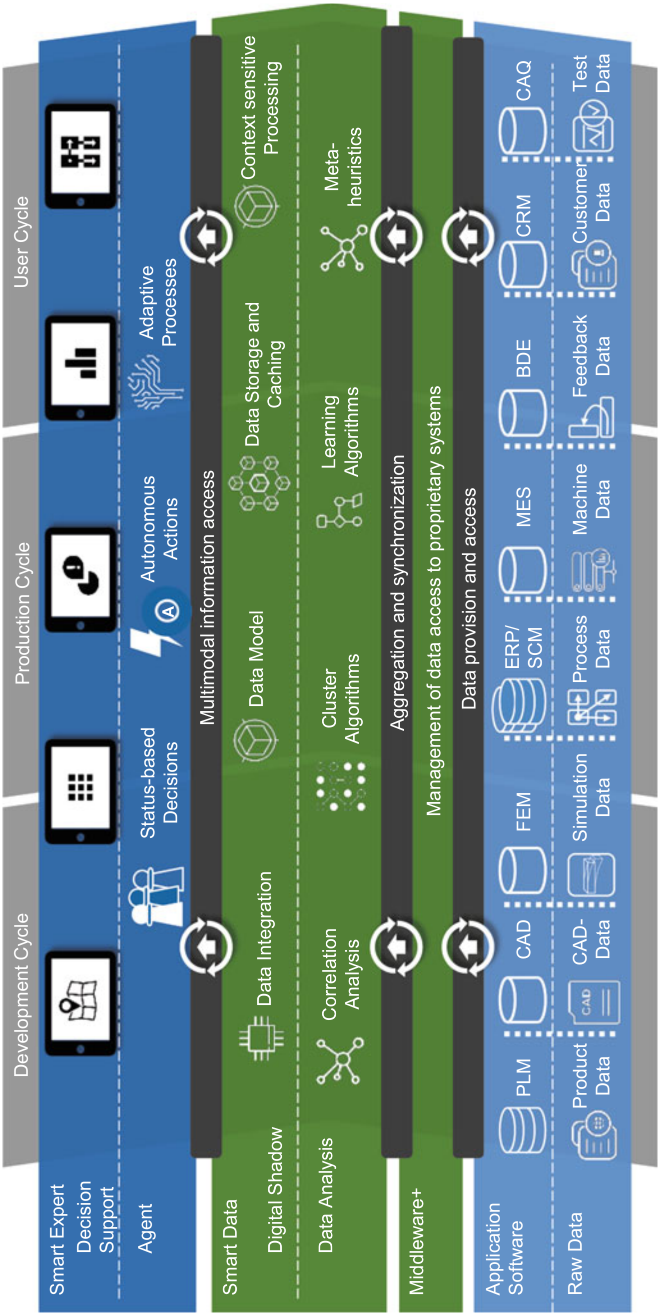


Fig. 1 The Internet of Production (Brecher et al. 2017)

In the following article, the content, objectives, and challenges at different levels of IoP are discussed in detail, embedded in the context of data-based **quality management.** Furthermore, it is explained how **data-based quality management** can be implemented across levels. Chapter 2 describes the identification of quality-relevant data and their associated data sources at the raw data level and provides an outlook on refinement into Smart Data. Afterward, Chapters 3 and 4 discuss the acquisition of sensor data in production and the connection to cloud-based **IoT operating systems** using **edge devices.** Chapter 5 describes the extension of existing quality methods to include data-based quality predictions and the Smart Expert level. Subsequently, Chapter 6 highlights how humans are empowered to make knowledge-based decisions in quality management, applying the described contextual data usage. Finally, a cross-level conclusion is drawn, and an outlook on future work and potentials of quality management in the Internet of Production is given.

2 Quality-Relevant Data: Synergy Between Quality Management and the Internet of Production

The following chapter describes how the Internet of Production infrastructure can enable data-based quality management. In this context, the common point of connection is represented by the quality-relevant data and its data sources, which are identified within the framework of quality management on the one hand and made available by the IT infrastructure of the IoP on the other. There is an immanent synergy between the Internet of Production and data-based quality management in a certain sense. This synergy is explained below by briefly defining the term **quality** and explaining how quality management can be implemented.

The DIN EN ISO 9001 standard defines quality as the degree to which a set of inherent characteristics meets the requirements (DIN EN ISO 9000:2015). From a metaphorical perspective, the standard considers quality as the degree of overlap

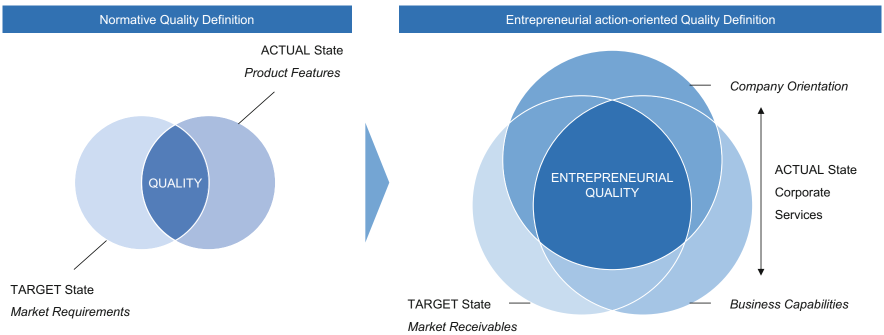


Fig. 2 From the normative to the entrepreneurial action-oriented quality definition

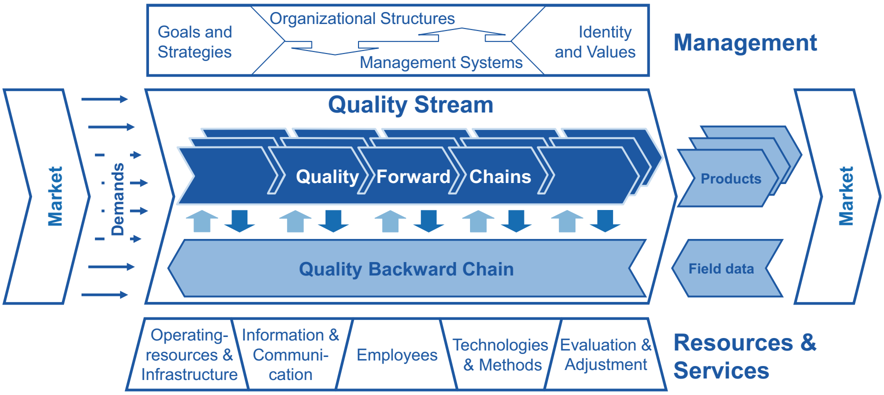


Fig. 3 The Aachen quality management model

between the product features offered and the customer's requirements (Fig. 2). The implementation of quality in the sense of the standard is not trivial in business practice. No concrete recommendations for action to increase the degree of coverage described above can be derived from the norm. In addition, the normative definition must be further developed against the background of **Total Quality Management**, according to which quality should relate not only to products but also to processes and systems, since neither process nor system quality can be derived from the standard (Pfeifer and Schmitt 2014). Schmitt takes this idea and develops the **Aachen quality management model**, which acts as a guide for action in implementing quality management in the company (Schmitt and Pfeifer 2015).

The Aachen quality management model (Fig. 3) extends the classic quality definition from the standard by taking the company's goals and boundary conditions into account. In this sense, the product should fulfill the market requirements and be in line with the company's orientation and feasible with the company's capabilities. In the Aachen quality management model, the aim is to achieve continuous improvement in corporate quality, which is determined by the degree of overlap between market requirements and corporate performance. Corporate performance is the degree of overlap between corporate orientation and corporate capabilities (Schmitt and Pfeifer 2015).

In the Aachen Quality Management Model, quality-related activities are assigned to three core elements: Management, Quality Stream, Resources and Services. Management is responsible for defining and pursuing corporate goals and strategies, defining quality policy, and designing organizational structures (Pfeifer and Schmitt 2014). Linked to the market is the **Quality Stream**, which aims to transform market requirements into innovative products. The Quality Stream consists of several forward-looking Quality Forward Chains (preventive QM) and a backward-looking Quality Backward Chain (reactive QM). Resources and services are assigned all essential activities required to execute the other two core elements (Schmitt and Pfeifer 2015).

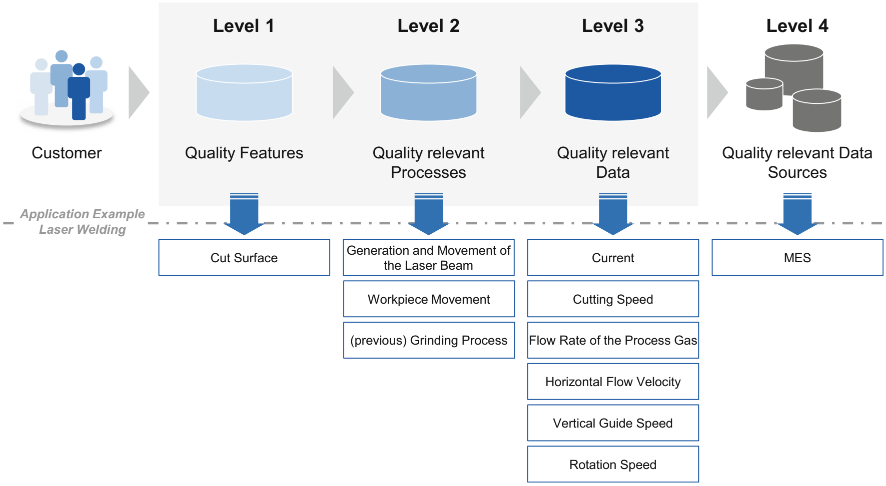


Fig. 4 Four-stage process model for identifying quality-relevant data and associated data sources

Based on the Aachen quality management model, it is clear that a large part of the activities within quality management has a data-oriented character. Decision-making processes in quality management are based on extensive data collection and analysis. The **Six Sigma** methodology is cited as an example, which effectiveness heavily dependents on a valid database. The sequence of Six Sigma projects can be described in principle with the **DMAIC cycle**, which consists of five phases. Two of the phases (Measure and Analyze) deal with systematic data collection and analysis. In many cases, these take up about 55% of the total project duration (Franchetti and Yanik 2011). Since quality management has a data-based character, the Internet of Production favors its implementation by providing the required data more efficiently and with adequate granularity. In the way the Internet of Production works, a Middleware+ provides access to raw data, which is then refined into Smart Data using various data analysis methods. The Digital Shadow is formed based on the pre-processed data, a data construct that has a granularity appropriate to the intended use. The Digital Shadow then enables the so-called Smart Expert to make knowledge-based decisions with the information it contains. Against the backdrop of quality-oriented use, there must be clarity about the definition of quality-relevant data when data is extracted via the Middleware+ and when the Digital Shadow is formed so that it can be refined into Smart Data. To address the overriding goal of improving corporate quality based on the Internet of Production, we, therefore, need a methodology for identifying quality-relevant data and data sources (Fig. 4).

At this point, the four-stage procedure for identifying quality-relevant data and associated data sources shown in the figure is presented and illustrated using laser cutting as an application example. On the one hand, some suitable methods already exist at the individual stages. On the other hand, the modification of existing

Table 1 Derivation of quality-relevant data using laser welding as an application example

Quality relevant processes	Quality relevant data
Generation and Movement of the Laser Beam	Amperage (A)
	Cutting Speed (m/s)
	Flow Rate of the Process Gas (cm^3/s)
Workpiece Movement	Horizontal Guide Speed (m/s)
	Vertical Guide Speed (m/s)
(Previous) Grinding Process	Speed of Rotation (Hz)

methods or the complete development of new methods is necessary (Refflinghaus et al. 2016).

Stage 1: Identification of Quality Characteristics
Starting with the customer, the quality characteristics of the product under consideration are first identified. Since **customer requirements** change permanently, specific quality characteristics do not remain forever but must be temporarily re-identified. In quality management and market research, various methods already exist for identifying quality characteristics, such as the focus group method, the lead user method, or empathic design. In the application example of laser cutting, the resulting cut surface is a quality characteristic.

Stage 2: Identification of Quality-Relevant Processes
The realization of quality characteristics on a product is mostly dependent on the immanent interactions of several processes, which must be analyzed and monitored to ensure high product quality. Here, the focus is mainly on the quality-relevant processes that can be determined using quality management methods. For example, these can be derived from the quality characteristics by applying Quality **Function Deployment**. Concerning the application example of laser cutting, the following quality-relevant processes can be derived from the quality characteristic "cut surface": Generation and movement of the laser beam, movement of the workpiece, and preceding grinding process.

Stage 3: Identification of Quality-Relevant Data
The quality of production processes can be assessed with the help of quality-relevant data. One possible method for deriving quality-relevant data is the statistical **Design of Experiments**. An analysis based on the design of experiments can be used to determine the critical process variables whose reference variables could lead to the quality-relevant data. For example, the manufacturing process CO2 laser cutting has the quality-relevant data cutting speed, current, or flow rate of the process gas. Table 1 shows further examples.

Stage 4: Identification of Quality-Relevant Data Sources
In this stage, the data sources from which the quality-relevant data get extracted are identified. Possible data sources are, for example, the **Manufacturing Execution System** (MES) or other operating systems that provide the required data in the form of tables, documents, other media (images, audio, video), or online data streams. If

the required data is not available, it must be recorded using appropriate sensor technology. The quality-relevant data of the laser cutting application example specified in Table 1 are usually stored automatically in the MES.

With the four-stage process model's help, the quality-relevant data and the preferred data sources from which it can be extracted are identified. This information is essential so that Middleware+ can be used to extract the relevant data and refine it into application-specific Smart Data. Against this background, quality management can benefit from the Internet of Production infrastructure and develop a synergetic interaction with it. However, the necessary preconditions must be created for this, both by implementing a functioning Middleware+ and forming structures within the corporate organization that enable continuous identification, recording, and validation of quality-relevant data.

3 Sensor Data Acquisition in Production Systems

In the last stage of the presented four-stage approach for the identification of quality-relevant data and associated data sources, the results are divided into the following different classes of data sources (Fig. 5).

A large part of quality-relevant data can be taken directly from the machine control systems or the higher-level MES systems. A general overview of the data streams typical for machine tools and data sources is given in Fig. 6 (Institute of Electrical and Electronics Engineers et al. 2015). They belong to the class of in-process data and are directly linked to the process by a closed-loop or open-

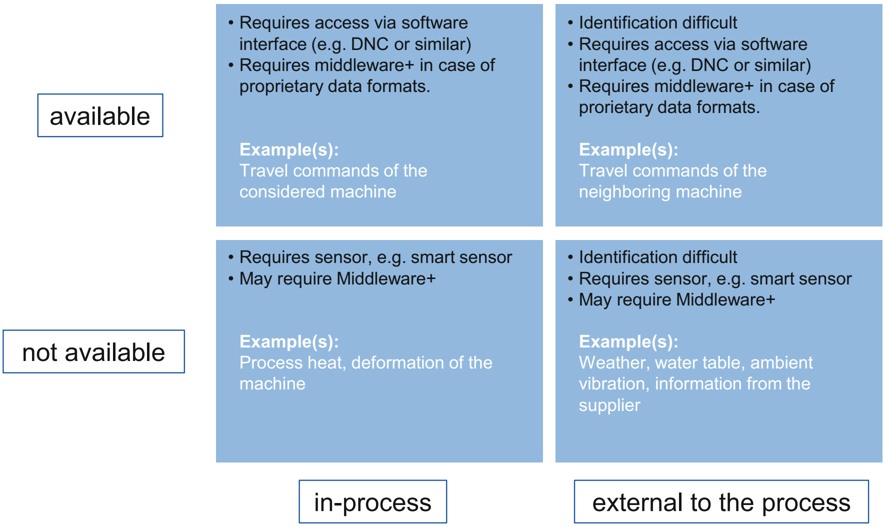


Fig. 5 Four different classes of data sources

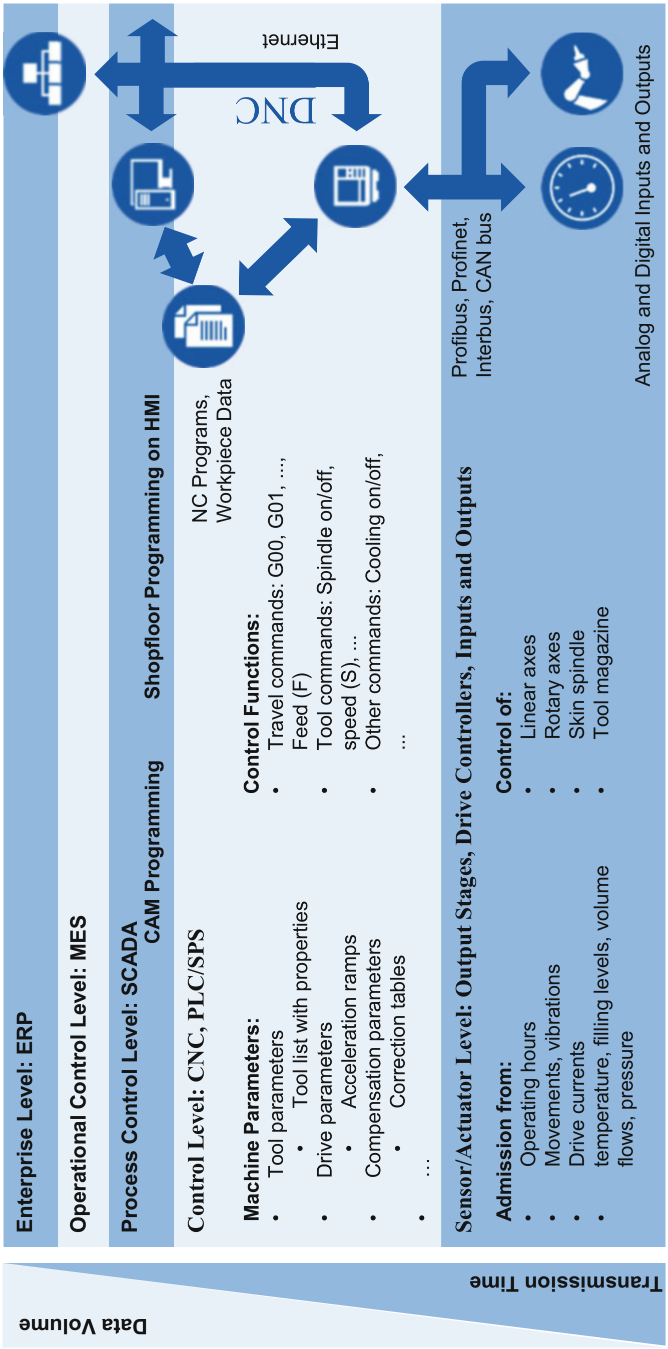


Fig. 6 Information flows in the Machine Tool System (adapted from: Institute of Electrical and Electronics Engineers et al. 2015)

loop control system. In modern systems, they can typically be used for higher-level quality management with little effort. However, the prerequisite for this is that the machine or control manufacturer enables access to this data. This is typically done via the option of a proprietary software interface, which in many cases can be found under the established term **Distributed Numerical Control** (DNC). The challenge is to convert the data structure, which is often also proprietary, into a readable format. A Middleware+ used for this purpose is described in more detail in the following chapter.

Furthermore, data external to the process can also be relevant to quality. An example of this is vibration generated by machines adjacent to the examined production machine, which influences the production process. In this case, the data sources might be external to the process but already available in the DNC network. Quality-relevant data can also be generated in upstream process steps. In the case of information about material properties, the data source may already be located at the supplier. This clearly shows the advantage of a horizontally and vertically networked production landscape.

Particularly in processes where the highest demands are placed on process stability, identifying and recording disturbances and thus quality-relevant data can become arbitrarily complicated. If data sources are not available, suitable sensors for recording the data must first be found and integrated. Along with the significantly increased importance of process data, the requirements profile for sensors has changed, which led to the development of so-called **Smart Sensors.** In classic sensors, which generate an electrical, usually analog signal from a measured physical variable, sensor calibration, signal processing, and transmission typically have to be performed separately. Smart Sensors, in contrast, perform a variety of tasks and integrate several functions in a single component. As highly integrated and intelligent systems, they can partially enable or provide self-monitoring, self-tuning, signal conditioning, digitization, digital signal processing, and digital system interfaces and contain several layers at once in the IoP diagram (Fig. 1). With Smart Sensors, the boundaries to edge devices and cloud-based IoT systems become fluid. The main advantages of Smart Sensors are that they not only provide pure data but that these can also be exchanged with the necessary consistency from the **shopfloor** to various company levels using integrated communication technology. This simplifies production monitoring and optimization in relative real-time and provides direct decision support for the machine operator or even experts elsewhere. A key challenge is integrating or retrofitting such technologies into existing production systems, also known as **retrofit**.

A good example of this is the temperature measurement system developed at the Chair of Production Metrology and Quality Management, which can be easily retrofitted into existing industrial plants and used in many areas (Fig. 7a).

Temperature represents a quality-relevant influencing variable in many processes, which is often shaped by the environment and is not part of the process control or regulation. During production and quality testing in the form of dimensional

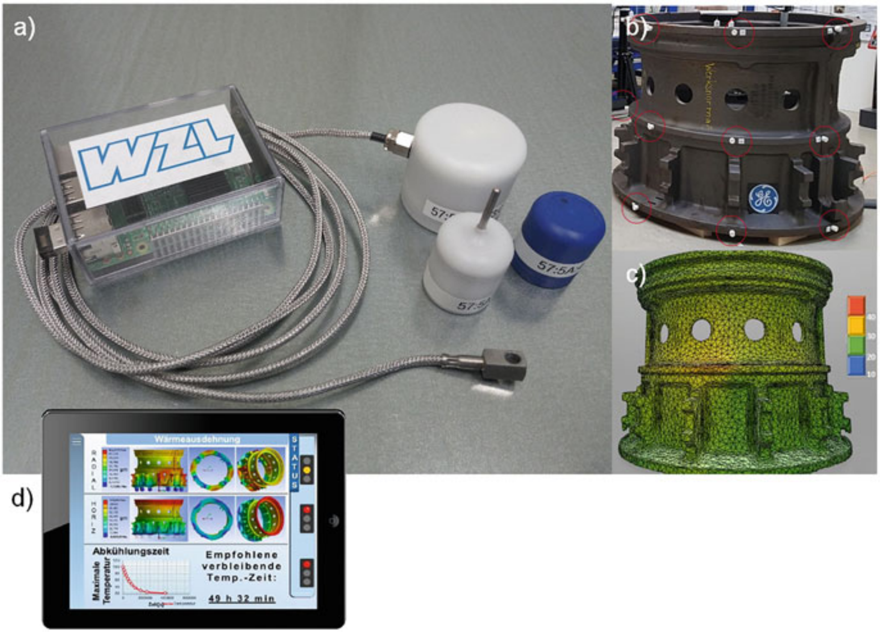


Fig. 7 WZL Temperature Measurement System: (**a**) Smart Sensors with gateway, (**b**) Large component, (**c**) Coupled FEM model, (**d**) Decision support via tablet with visualization software

Table 2 Technical specifications of the sensors

Feature	Specification	
Temperature sensor type	3 different sensor types	Contact sensor wireless (PT1000)
		Ambient sensor wireless (PT1000)
		Contact sensor wired (PT1000)
Dimensions (D × H)	D = < 50 mm, H = < 50 mm	
Housing material	Thermoplastic material (POM), suitable for industrial use	
Mounting type	Magnetic	
Energy supply	Battery operated (life >1.5 years)	
Communication interface: Sensor—Gateway	Bluetooth Low Energy 4.2, Class II, nominal range 80 m	
Gateway	Raspberry Pi™, various communication protocols (MQTT, HTTP)	
Data format	*JSON* among others	
Other properties	Dust and waterproof, shockproof	

measurements, temperature fluctuations can also cause large variations in process quality due to the thermo-elastic behavior of physical systems. Although many systems have temperature sensors, the spatial coverage is usually not high enough to gain knowledge and subsequent process optimization, which results in the need for an integrable smart temperature sensor.

Due to the often-harsh environmental conditions in the production environment, the sensors have been provided with design and features according to the requirements. The technical specifications of the sensors are listed in Table 2. Features such as data transmission via radio, battery-powered, dust- and waterproof, availability of various communication protocols, and open data formats are exemplary for high integrability.

One use case of the developed temperature measurement system is explained below using the dimensional measurement of large components as part of quality inspection. In the production of large industrial components (Fig. 7b), such as turbine housings, tolerances' requirements are disproportionately high compared to the component dimensions. For this reason, the influence of temperature and the resulting thermal expansions are significant. Due to heterogeneous and transient temperature distributions, large components' thermal condition is often not sufficiently known. Depending on the size of the component, the tempering process can take several days to weeks. Regardless of this, in practice, the times until complete temperature control are rarely observed due to the necessary, long machine downtimes, storage capacities, or delivery deadlines. The measurements are subject to increased uncertainty, which cannot be tolerated from the quality assurance side.

To address this shortcoming, a method for "in-process" monitoring of transient thermoelastic states was developed (Fig. 7-a-d). With the presented Smart Sensors' help, the component surface temperature distribution is determined at equally distributed points and transmitted to a gateway via Bluetooth (Fig. 7-a). By coupling to a FEM software, the geometric displacement is calculated based on the real-time temperature distributions and an evaluation and prediction of the component condition is enabled (Fig. 7-c). The appropriate positions of the temperature sensors are determined using the CAD model of the component. The temporal temperature distribution on the component surface and inside the component is calculated using an integrated model, which is iteratively fed and optimized using ongoing temperature measurements on the component surface. Based on the calculated temperature distribution, the component's 3D deformation can be continuously redetermined concerning the stored CAD model for a given homogeneous reference temperature (usually 20 °C). For each current and future point in time, the software provides a CAD model of the deformed component as well as a correction vector field in a standardized, industry-standard data format. A mobile device provides **Decision Support**, which provides a visualization of the thermally induced displacements for the Smart Expert to assist in decision-making (Fig. 7-d). The temperature measurement system offers considerable savings potential for producing large components, both in machine time and in eliminating the need for air conditioning in the production and test rooms. Furthermore, thermally induced measurement uncertainties are significantly reduced, significantly reducing the risk of producing defective parts.

The temperature measurement system's application can easily be extended to other scenarios, for example, to the monitoring of critical thermally induced deformations of production systems. In addition to temperature, many other influencing variables can be relevant to quality. Typical examples are Deformations, geometric or kinematic machine faults, vibrations, and electromagnetic radiation (Montavon et al. 2018). Regardless of which variable is involved, which measurement principle is used to collect the measured variable, and whether the required data is internal or external to the process, the preparation and provision of the data takes on the central role in terms of IoP, and this is discussed in more detail below.

4 Edge Devices and Cloud-Based IoT Operating Systems

The establishment of cloud-based platforms in technical ecosystems with a large, heterogeneous group of users, contributing software developers, and individual devices (e.g., smartphones) shows potentials that can also offer added value for **cyber-physical production systems** (Schmitt and Voigtmann 2017). Therefore, the development of corresponding platforms for industrial ecosystems is of great interest and is being pursued by several companies and consortia around the keyword IoT operating system. This term is initially understood to mean a bundling of various software applications that are directly compatible with each other for general use in the area of the industrial **Internet of Things.** The typical functional components of such systems can also be identified in part in the Aachen model of the Internet of Production:

- A portfolio of communication interfaces makes it possible to receive data both from distributed sensors or machines and from other software units (raw data). The interfaces can be realized via independent standards as well as via proprietary gateway applications (Middleware+).
- A uniform database system (e.g., based on SQL) is used to manage the data after a system-specific transformation of the data. A redundant backup of the latter is usually part of the intended database infrastructure.
- Integrated, dynamic visualization software makes it possible to provide application- and target group-oriented representations of the information obtained (decision-support).
- With the help of programming interfaces (SDKs), users and possibly third parties can develop their own applications for data evaluation and processing, which are executed directly within the IoT operating system. These can also include automated notifications and actions based on parameters identified as quality-relevant (**Agents**).
- An operating system's basic functions are also implemented according to the technology's current state, for example, a general network capability, user administration, and security infrastructure.

The provision of such an operating system including corresponding hardware in the **PaaS** (**Platform-as-a-Service) business model is** referred to as a cloud-based Industrial Internet-of-Things (IIoT) operating system. It is offered as a product by

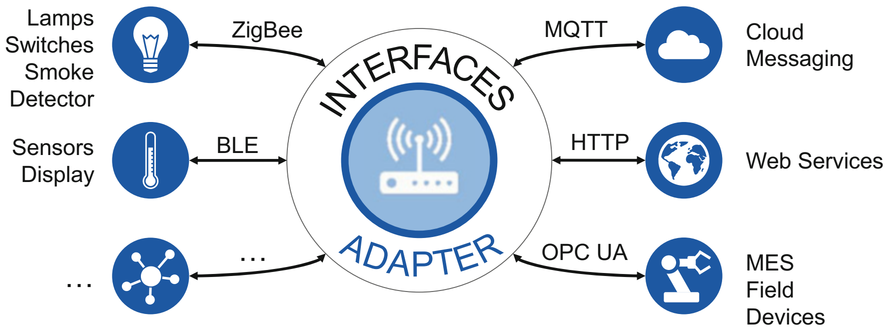


Fig. 8 Principal sketch of a gateway for cross-protocol IoT communication

Siemens under the name MindSphere®, by GE under the name Predix, and by a larger consortium as ADAMOS, among others (Peters and Schäfer 2018; Boniface et al. 2010). These also compete with general offerings for IoT platforms, such as Microsoft Azure, Google Cloud Platform, and Amazon Web Services (Peters and Schäfer 2018). In principle, a cloud-based IoT operating system can also be individually designed by companies themselves from individual software and hardware components.

The heterogeneity in the area of platforms is also reflected in the number of available standardized interfaces and protocols for sensors and field devices: Already with OPC UA, MQTT, HTTP/CoAP REST as well as Bluetooth Low Energy, the selection is so large that an application-specific implementation of all options protocols is hardly possible, especially for small and medium-sized enterprises (SMEs). Simultaneously, in the course of Industry 4.0, the highest possible compatibility is necessary to increase competitiveness. Model-based software gateways offer an approach to solving this problem: The information and functions provided by the respective sensors, machines, and software solutions are mapped in a protocol-independent modeling language (Pfrommer et al. 2016). The transformation into the address spaces and data formats of the specific protocols is executed as a bijective adapter, including the additional required configuration settings (e.g., certificates) (Fig. 8). Thus, any number of interfaces can be provided for a single device with low implementation overhead. In addition, the central modeling can be used to establish a harmonization of the data also on the physical level (including format, unit, and uncertainty understanding) with the help of profiles.

The operation of the gateway logic presented here on dedicated hardware with corresponding network capability is a typical example of edge devices, which thus serve as enablers for the use of cloud-based IoT operating systems. An example of this is the temperature measurement system presented in the previous chapter. A Raspberry Pi™ makes the measurement data transmitted via Bluetooth Low Energy via various protocols for further analysis applications in the network. If edge devices are defined more generally as technical entry points into cloud-based systems, their role is not limited to the Middleware+ level of the Internet of Production model. Any

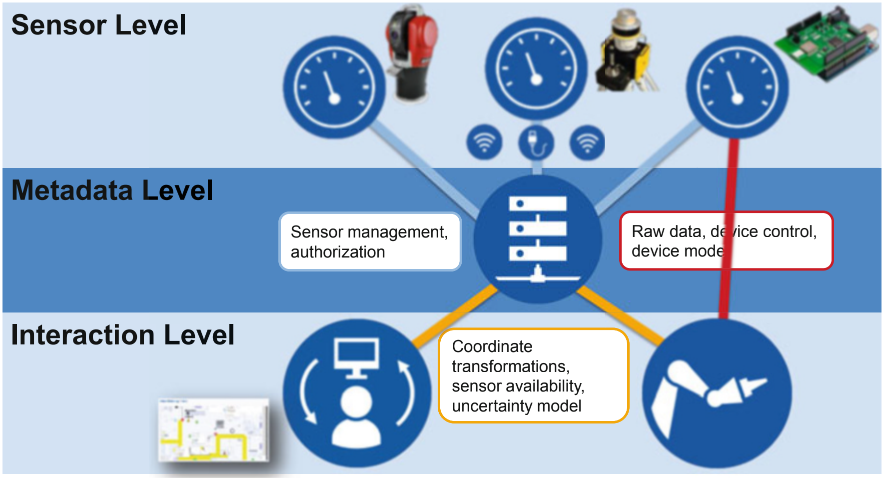


Fig. 9 Three-level architecture of the virtual metrology frame (Montavon et al. 2017)

entity, i.e., also measuring devices, machines, actuators, and HMIs, can be interpreted as an edge device.

The interaction of edge devices and a cloud-based IoT operating system will be illustrated by the use case of the **Virtual Metrology Frame** developed by Schmitt et al. (Fig. 9) (Montavon et al. 2017). This provides a metrological reference frame of the shopfloor based on distributed, spatially measuring systems. Their position information can be retrieved as a service via a uniform interface in a standard coordinate system. The communication architecture required for this is executed on an individually configured IoT operating system. In corresponding applications of the metrological reference frame, for example, metrology-based assembly tasks or inline intermediate feature inspections, the cloud-based architecture significantly reduces the integration time and achieves more efficient use of the measuring devices through networking (Montavon et al. 2017; Maropoulos et al. 2014; Schmitt et al. 2016). The last occurs as edge devices, which provide their position information after pre-processing and abstracting proprietary software components in a model-based gateway. Two database-based services are used as additional Middleware: On the one hand, the position information is stored as historical data series after a reduction, among other things, to keep them available as potentially quality-relevant data. On the other hand, the metadata of the individual measuring devices is managed, i.e., temporal and spatial availability, transformation parameters to the global coordinate system, and corresponding uncertainty models. Visualization and planning applications and actuators access the measurement data in a decentralized manner and use the centrally available metadata for their interpretation. This topology enables the time-critical and data-intensive transmission of position information directly between sensor and actuator and avoids infrastructural peak loads.

The discussion and choice of topology in the area of cloud-based IoT operating systems follows the same criteria as in the general design of information technology networks: resilience, high availability, as well as sufficient performance are set against increasing complexity and costs due to decentralized and redundant system.

At the same time, falling hardware costs and rising capacities in the field of edge devices, and the establishment of scalable PaaS offerings in the subscription model, are pushing operating and investment costs into the background as decision criteria. In addition, modern data transmission standards such as 5G offer new possibilities in terms of flexible architecture as well as achievable bandwidths and latencies (Andrews et al. 2014). The continuous technical development additionally motivates the content-driven and platform-independent design of IoT systems, in which IoT operating systems are used as basic vehicles similar to conventional PC operating systems.

In addition to the design of a cloud-based infrastructure within a company or unit, the current heterogeneity of IoT operating systems will lead to the question of how to design communication across different cloud systems (also referred to as **cloud of clouds**) in the future (Verissimo et al. 2012). In practice, this problem arises, for example, when quality-relevant data accumulates along the process chain at different suppliers, but an added value can only be gained from combining the data (Robinson and Malhotra 2005). Regarding the level model of the Internet of Production, whether this overarching networking can be mapped at the Smart Data level and which new types of trust mechanisms and business models will develop in this context.

5 Data-Based Quality Forecasts

The comprehensive digitization of industrial production opens the gates to more and more data, leading to an exponential increase in the available data volumes. On the one hand, production processes are being equipped with more and more sensor technology. On the other hand, there is ever-increasing networking of suppliers, producers, and customers (Huber 2018). Above all, implemented inline measurement technology enables the collection of product and process data, such as temperature, acceleration, or the force-displacement curve, directly during production. The collected data can be precisely assigned using auto-ID technologies. This makes it possible to link all existing information, such as order data, process data, product data, or quality data, with the currently produced product (Bauernhansl et al. 2014). With the hardware and software available today, the data can not only be collected and stored in real-time, but also pre-processed and analyzed in many cases (Erner 2018).

The quality of a product depends largely on the interaction of the individual production steps and the condition of the components used in each case. Due to increasingly complex production processes, the number of immanent interactions of individual processes is rising. In addition, the increasing individualization of products is leading to a significant increase in process variance. A data-based approach makes it possible to take a closer look at the factors relevant to quality. Instead of a

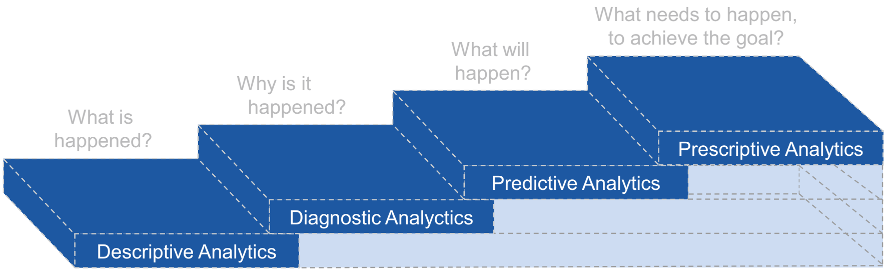


Fig. 10 Stages of data analysis

pure assessment based on process data, existing information about intermediate products and the individual assembly can also be considered. The use of existing data leads both to an increase in efficiency and to production in line with requirements and makes it possible to improve process and product quality in the long term (Refflinghaus et al. 2016; Schuh und Riesener 2017).

Data **analytics** is of fundamental importance for data-based quality prediction as a tool for examining large volumes of data of different types. A distinction is made between four different data analysis levels, some of which build on each other: **descriptive analytic**s**, diagnostic analytic**s**, predictive analytics,** and **prescriptive analytics** (Lin 2015). The individual stages are shown in Fig. 10 (Stimmel 2015). The individual stages' overriding goal is to extract the information contained in the data and make it usable for forecasts and decision-making aids (Schmitt 2016).

Descriptive analytics describes and summarizes the data obtained. The information generated in the process maps the past production process. This type of analysis has been part of classic quality management for several decades. For example, derived characteristic values are used for statistical process control. In Diagnostic Analytics, the information is used to determine the causes of specific events. In this way, possible reasons for any process or quality deviations can be uncovered. One example of this is the determination of the cause of an increased reject rate. Modern production systems already independently provide the required measured values via integrated sensors. Predictive analytics is then used to make forecasts based on existing data and the information obtained for the first time. Here, for example, process sequences are predicted considering a certain model error. While the first two data analytics stages refer to a point in time that can no longer be influenced and thus exclusively to historical data, the focus is now on events that will occur in the future with a certain probability. Prescriptive analytics then makes it possible, based on the forecast models' findings, to compare decisions and measures with their respective effects. This leads to a targeted influencing of processes. In this way, for example, quality losses can be proactively avoided, and ultimately the quality of the currently manufactured product can be controlled in a targeted manner. Targeted **quality control** paves the way for demand-oriented production, in which individual customer wishes and expectations can be taken into account in the form of flexible requirements. Due to the huge amounts of data, methods of **artificial intelligence**,

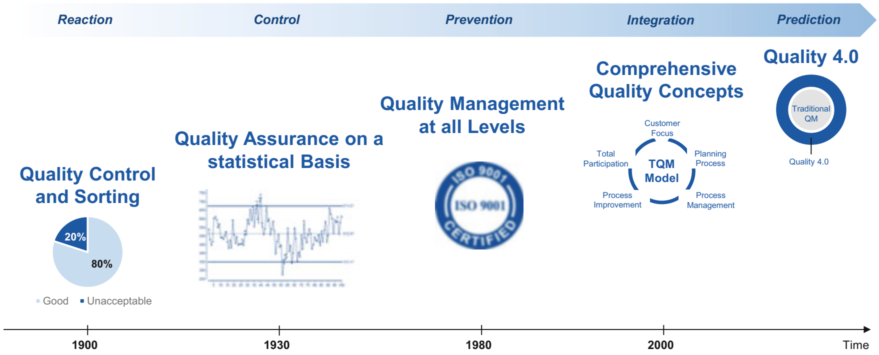


Fig. 11 Evolution of quality management

particularly **machine learning**, are increasingly being used at the third and fourth analytics levels, in addition to conventional methods such as regression analyses or simulations. These enable effective and efficient knowledge identification and extraction, especially with large amounts of data (Elser et al. 2018; Schmitt 2016; Lin 2015; Stimmel 2015).

With its data-based approaches, data analytics offers quality management possibilities that cannot be realized with classic Six Sigma methods. Most of the current Six Sigma approaches make statements about the primary population based on one or more samples. Since, in the context of the Internet of Production, information and measured values on all produced parts are already available in many use cases, it makes sense to use and analyze this data effectively instead of making sample-based statements. In this way, sample-based uncertainties of the respective observation can be avoided. Furthermore, data analyses using Six Sigma methods often only allow a limited number of parameters to be tested. The necessary intensive test planning and execution result in high time and cost expenditure. This problem is mostly circumvented using data analytics since the analyses and investigations are based on the existing database, permanently expanding during operation. However, a high level of data pre-processing must be considered for the results-oriented application of data analytics. Another important aspect is the prediction of the produced quality, the so-called predictive quality. Using classical approaches, which are primarily limited to the assessment and control of the current product and process quality, it is only possible to react proactively to future deviations and disturbances to a minimal extent. The potential to close this gap lies in predictive and especially prescriptive analytics. In the future, the new approaches of data-based quality prediction should make it possible to close quality control loops in a superordinate manner through targeted visualization and suitable measures so that quality-related failures can be avoided. For example, future quality deviations could be pointed out to the responsible employee via **smart devices** so that an adjustment and correction of the necessary parameters can take place before quality losses occur (Alp Kucukelbir 2018; van der Aalst 2016; van Dijk 2018).

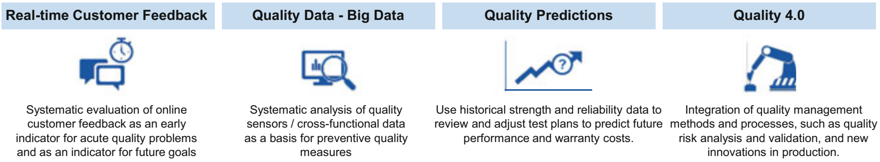


Fig. 12 Quality management through innovative methods

With the change in focus from reactive to proactive and predictive shown in Fig. 11, a new development era of quality management is beginning. When looking at the individual evolutionary stages of quality management, especially in manufacturing companies, the introduction and use of data-based approaches mark another paradigm shift. During the first two evolutionary stages, there was a shift from product-based approaches and methods of quality control to process-based approaches, with a focus on quality planning and defect prevention. With the effective use of the data available today and specifically the benefit of all information and measurements about every component manufactured and every product produced, the product focus moves back into the spotlight. Data-based decisions on a continuous basis enable the focus to be directed more strongly on the individual product again.

The actual approaches of artificial intelligence and machine learning algorithms are often decades old but are only finding their way into many other areas with current technology (Jones 2014). These technologies must be adapted for the production environment within the framework of the Internet of Production to be able to ultimately use the existing potential. Particularly in the area of quality management, only a few solutions have been actively used to date. To change this, it will be necessary for the next few years to expand the solutions already implemented to include new approaches and thus to provide people (Smart Experts) with knowledge-based decision support in various areas of quality management through the **consistent use of data.**

Figure 12 shows some examples of applications for data-based quality management methods. For example, the systematic use of online customer feedback data makes it possible to improve a company's own understanding of its products and identify acute quality problems more quickly than possible using conventional complaints and feedback loops. In addition, the information collected can be used as an indicator of customer opinion to formulate future goals (Schmitt et al. 2014). The use of quality data in the form of a systematic analysis of measured values from quality sensors and especially cross-functional data serves as a basis for preventive quality measures and estimation of optimal parameters (Refflinghaus et al. 2016). Using historical strength and reliability data, test plans can be controlled and adjusted as well as future performance and warranty costs can be predicted. Employing the so-called Quality 4.0, new innovative quality management methods and processes are integrated into production and enable, for example, quality risk analyses and validations. In addition, this offers the possibility of taking individual customer-specific quality profiles into account in the future. Over the long term, the new data-driven methods lead to fact-based decisions on a continuously expanding knowledge base (Dirlea et al. 2019; Kleinemeier 2016; Schmitt 2016).

In addition to the advantages and potentials mentioned, there are still some limitations and problems in implementing predictive quality. Especially in applying approaches from the field of artificial intelligence, there is still a considerable need for research. The data itself is particularly critical. Approximately every 2 years, the amount of data available doubles (Mayer-Schönberger 2015). However, most of it is available in unstructured form, which is why direct processing is usually only possible to a limited extent (Schmitt 2016). As already mentioned initially, there is often a lack of compatibility due to many different, often proprietary systems, which is why the data required for modeling may be available but not accessible. Another aspect is the computing power needed to compute models. Even training smaller neural networks, a popular branch of machine learning, quickly leads to enormous computational requirements. For large complex models, even modern computing centers still reach their limits. Another challenge is the increasing complexity of processes in both production engineering and data analytics. Employees who want to implement data-based approaches within production must have expert knowledge in production engineering and data analysis (Elser et al. 2018). However, the most significant criticism of methods from the field of artificial intelligence remains the lack of theoretical background. For example, artificial neural networks lack transparency due to their black-box character model. It is often impossible to exactly understand or predict how and especially why the model reacts in a certain situation and under varying circumstances. The lack of security and reliability is still an exclusion criterion for the use of artificial intelligence solutions for many industries, especially for critical processes (Wolfangel 2016).

6 Decision Support for Humans

Growing global competition is creating increasing cost and deadline pressure for companies. In addition to the continually growing number of variants, the demand for customized products and services is also rising. Particularly in the context of **Industry 4.0**, there is also a growing demand for increased performance and flexibility of processes, which also enable efficient use of resources. Overall, this leads to significantly increased complexity in manufacturing companies (Steinhilper et al. 2012). Demographic change and its consequences also pose further challenges for Germany as an industrial location (Plorin et al. 2013). To survive and remain competitive in the long term in such a dynamic environment, manufacturing companies are forced to increase productivity and flexibility and to integrate new technologies into production successfully (Steinhilper et al. 2012). These changes in the production environment in the context of Industry 4.0 are strongly characterized by the integration of new technologies, which increasingly pose major challenges to manufacturing companies (Lerch and Gotsch 2014).

The importance of the "human" factor in these work and interaction processes is made clear in a 2016 study by the management consultancy Accenture. Eighty-three percent of German consumers say they still want to be served by people and see interaction with people as a significant factor in customer satisfaction (Accenture 2016). This suggests a continued very high or even growing need for trained employees about direct interaction with customers. In addition to social skills in direct customer contact, new job profiles require an increased ability to act situationally, particularly for employees in the production environment. When interacting with people and systems, a situation analysis is needed, i.e., the existing necessities and needs must be recognized and recorded quickly and correctly. Furthermore, cognition and decision processes must be supported, and a continuous reassessment of the situation must be performed (Haas et al. 2012). In the process, factors that are often unforeseeable, such as specific requirements, need to be recorded ad hoc and integrated into decision-making processes in a meaningful way. This is the only way to achieve adequate decision quality for different types of decisions. The human ability to deal with unpredictable situations and incomplete information cannot yet be reliably imitated by technology. Thus, the human factor continues to be one of the most important variables for corporate success (von Hahn 2018). In relation to the Internet of Production, human decision support starts at the Smart Expert level. The employee should be supported in the situational decision-making process by processing the various stages of data analysis and ultimately becomes the internal customer of data-based quality forecasts.

To meet these changing demands on the **competence profile** of employees, the question arises as to how employees who previously had little or no contact with these tasks can be enabled to cope with new work tasks. The 2015 survey shows that companies that focus on hybrid value creation are increasingly investing in employee skills development (Bahrke and Kempermann 2015). SMEs particularly face challenges in this context, where the availability of resources is limited. Therefore, there is a need to empower employees effectively and efficiently for new work tasks in direct customer interaction to ensure high decision-making quality and ultimately employability. The integration of **digital support tools**, such as tablets or smart glasses, is one way to address this need, according to the principle of "**training on the job**." In principle, some approaches involve the use of new technologies to increase the quality of decision-making. However, these often focus on the management level, such as Big Data applications in conjunction with mobile devices for strategic planning Optimizations for decision quality combined with smart glasses are also already taking place on the store floor itself (Krauß et al. 2016; Lindner et al. 2017).

Since, in addition to the technical development potential created by the use of Industry 4.0 technologies, humans are a decisive factor in the design and efficient use of the entire production system. They must be optimally integrated through the data-based support of situational decision-making processes. Human flexibility and creativity are difficult to replace by autonomous systems, which is why humans must be equipped with intelligent decision-making systems in Industry 4.0 to process the complexity that has arisen in the production process and analyze it and make the right decisions. This requires a collaborative form of **work organization**, in which people must control, coordinate and optimize spatially distributed and interconnected production resources. The decisions needed for this take place in considering the situation- and context-dependent objectives (Botthof and Hartmann 2015). Due to the increasing complexity and the multitude of information and decision options, people in networked production increasingly need **decision support systems** such as tablets or smart glasses (smart devices), which enable people to make an informed decision based on data (Smart Expert level).

Accordingly, a decision support system working in the field does not have to make the decision itself. Still, it must support the development of people's decision-making ability by increasing their understanding of the environment. Therefore, their role is knowledge mediating. This is because although the decision process itself cannot necessarily be represented in a fully structured way, decision support systems can provide models that help humans gain insight into the relationships between decisions and the goals to be pursued (Gorry and Scott Morton 1971). Correspondingly, problems can either be prepared in a fully structured way, i.e., the knowledge required to solve them is represented in a fully encoded way by a mathematical model, or it is prepared as a semi-structured or structured subproblem that can be processed by a computer and used as decision support (Prenzel 2018).

Such decision support systems are abbreviated as **DSS (Decision Support Systems)** and are highly organized information systems that serve the purpose of decision support. Figure 13 shows different components that work together in the context of Industry 4.0 and are interlinked by DSS. They represent data-based solutions used to support and solve complex tasks (Shim et al. 2002). DSS operates in a bounded decision environment and initially includes a goal system that explains much of the decisions' preferences. Furthermore, the decision context itself must be specific and unambiguously delimited and have definable decision procedures. Humans, which can have different roles, and a suitable working environment for the preparation, analysis, and documentation of decision proposals are also required for such a DSS. For a DSS to function as such and enable the human to make the optimal decision based on data, it must have a broad knowledge base or acquire this during its activity. For this purpose, external data sources (e.g., specialized databases), mathematical models, and methods with a modeling environment that is as open as possible, as well as procedures and search engines are used, as are internal utilities and reporting systems (Haettenschwiler 2001).

A knowledge-based DSS has the human expert's decision logic stored and simulates the human way of thinking. This makes the solution path transparent and comprehensible for the user. By integrating the familiar language, the decision-making process and the result are presented in a way that is easy to understand. In addition, the transparency and familiar presentation strengthen the confidence in the correctness of the procedure. Thus, the human user in the production environment is supported in his decision-making ability and enabled to make an **optimal decision**. As possibilities of solution strategies are available for selection, these are presented logically and comprehensibly, and possible consequences are

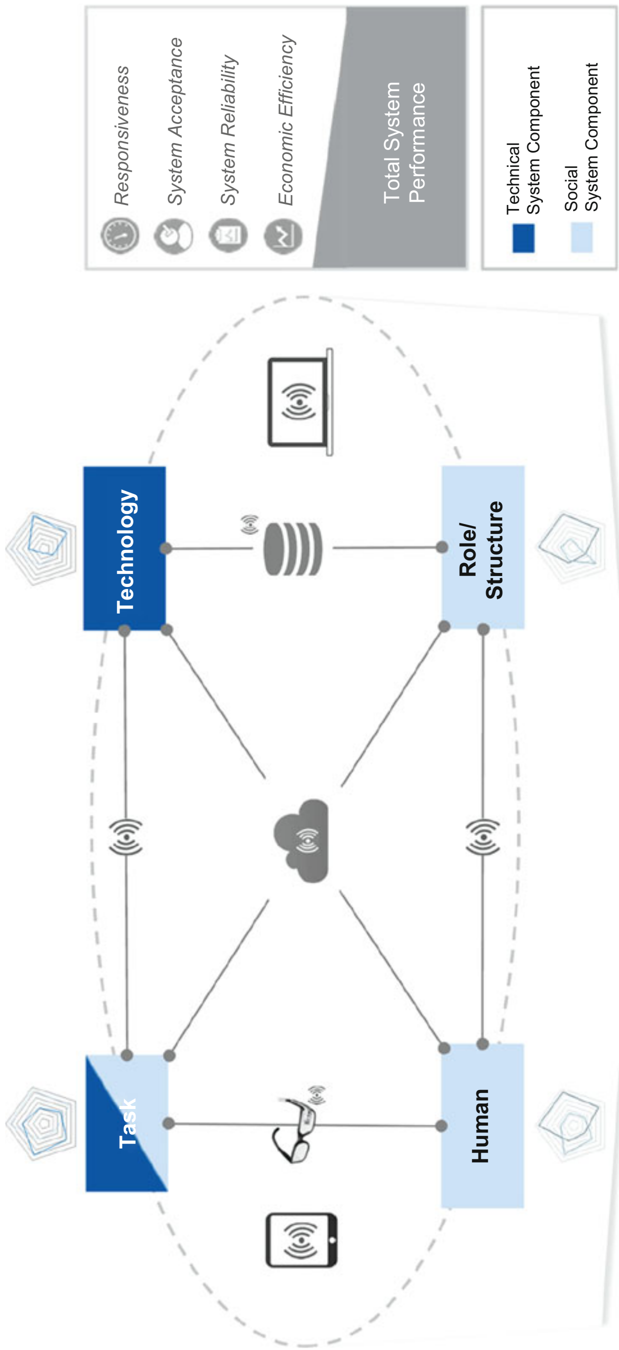


Fig. 13 Linking various Industry 4.0 components through DSS

simulated and mapped, and the decision-making itself is fully prepared (Klein and Methlie 2009).

A prerequisite for the function of such a DSS in the Internet of Production is the implementation of the two upstream levels (raw data and Smart Data). Only when the raw data has been prepared (cleansing, aggregation, filtering, contextualization, and synchronization) and processed into Smart Data for quality forecasts, a DSS can use the data to enable people to make optimal decisions. The DSS can support the employee in all four types of quality prediction (Descriptive, Diagnostic, Predictive, and Prescriptive). If the first two levels of the Internet of Production are sufficiently integrated, a DSS can be used not only in operational processes but also in the planning and design of production systems and open up optimization potential (HAB Conference Proceedings 2014).

The primary use of DSS today is seen in the area of management. However, in addition to managers, more and more skilled workers and production workers are also being integrated into the decision-making process in the context of Industry 4.0. With the help of **mobile assistance systems,** for example, employees in production receive context-intensive information about current production performance data to plan and implement the next steps (Kagermann et al. 2013). Technical assistance systems also support optimizations by using specific reflection and decision-making capabilities (Spath 2013). Consequently, manufacturing companies need to integrate decision support systems to remain competitive in the future. Depending on the existing competence profile, the use of such systems increases the decision-making quality of employees and thus contributes to more efficient use of existing production resources.

7 Conclusion and Outlook

The digitization and implementation of Industry 4.0 in manufacturing companies result in entirely new quality management requirements due to the rapidly increasing availability of data in real-time. The use of classic QM methods, which are primarily limited to the assessment and control of current and past product and process quality, does not allow the full potential of the available data volumes to be exploited. New approaches to a data-based quality prediction that apply predictive and prescriptive analytics are intended to close the quality control loops in the future using targeted visualization and suitable measures, even on a higher level, to avoid quality-related failures. With this paradigm shift from reactive to proactive and predictive, a new development era of quality management is beginning. During implementation in the next few years, the focus will be on adding new approaches to solutions that have already been implemented and thus enabling knowledge-based decision support for people (Smart Experts) in various areas of quality management through the consistent use of data.

For the successful implementation of data-driven quality management, the Internet of Production provides the necessary IT infrastructure, which enables real-time, secure information available at any time and any place. To this end, the existing raw data layer or Application Software is extended by three superordinate layers, which enable joint analysis of information from different domains. The central element of predictive and prescriptive models in data-driven quality management is the Digital Shadow. This data construct has the granularity appropriate to its intended use as a combination of pre-processed data and merely copied raw data.

Due to the diversity of proprietary systems in manufacturing companies, accessing and processing the data requires a high initial effort in cleansing, aggregation, filtering, contextualization, and synchronization. Further challenges of the intended cross-domain analysis of large amounts of data often lie in missing context information, low information density, limited accessibility, and the domain knowledge required to understand the data at hand. In addition to the purely data-related hurdles, further challenges arise from the high computing power often required for machine learning-based prediction models, their black-box nature, and the increasing complexity of production processes in general. Thus, there is still a considerable need for research, especially concerning the use of artificial intelligence methods in quality management, to increase the safety and reliability in the application, which is relevant for many industries, and to consolidate trust in new solutions.

References

Accenture (2016) 11th annual Accenture Global Consumer Pulse Survey. Dublin

Alp Kucukelbir (2018) Beyond Six Sigma: Process Improvement with Machine Learning. Ferolabs. Available online at https://www.ferolabs.com/post/beyond-six-sigma-process-improvement-with-machine-learning, last updated on 2018-09-04, last reviewed on 2019-01-06.

Andrews JG, Buzzi S, Choi W, Hanly SV, Lozano A, Soong ACK, Zhang JC (2014) What will 5G be? IEEE J Select Areas Commun 32(6):1065–1082. https://doi.org/10.1109/JSAC.2014.2328098

Bahrke M, Kempermann H (2015) Hybride Geschäftsmodelle - als Lösungsanbieter zum Erfolg. eine vbw Studie. Munich

Bauernhansl T, Hompel M, Vogel-Heuser B (eds) (2014) Industrie 4.0 in Produktion, Automatisierung und Logistik. Application, technologies, migration. Springer Vieweg (SpringerLink). Available online at, Wiesbaden. https://doi.org/10.1007/978-3-658-04682-8

Bauernhansl T, Reinhard G, Schuh G, Krüger J (2016) WGP - Standpunkt Industrie 4.0

Boniface M, Nasser B, Papay J, Phillips SC, Servin A, Yang X et al (2010) Platform-as-a-service architecture for real-time quality of service management in clouds. In: Guadalupe Ortiz Bellot (Hg.): Fifth International Conference on Internet and Web Applications and Services (ICIW), 2010. 9–15 May 2010, Barcelona, Spain; Proceedings 2010 Fifth International Conference on Internet and Web Applications and Services. Barcelona, Spain. International Academy, Research, and Industry Association; International Conference on Internet and Web Applications and Services; ICIW. IEEE, Piscataway, pp 155–160

Botthof A, Hartmann EA (eds) (2015) Zukunft der Arbeit in Industrie 4.0. Springer Vieweg, Berlin

Brecher C, Klocke F, Schmitt R, Schuh G (eds) (2017) Internet of Production für agile Unternehmen. AWK Aachen Machine Tool Colloquium 2017, May 18-19. IIF - Institut für Industriekommunikation und Fachmedien GmbH; Aachener Werkzeugmaschinen-Kolloquium; AWK. 1st edn., Apprimus Verlag, Aachen

DIN EN ISO 9000:2015, 2015: Quality management systems - Basic principles and terminology (ISO 9000:2015)

Dirlea V, Kidambi R, Krubasik S, Sachseneder C (2019) Quality 4.0 Preventive, Holistic, Future-Proof. Hg. v. ATKearney. Available online at unter https://www.atkearney.com/industrial-goods-services/article?/a/quality-4-0-preventive-holistic-future-proof, last updated on 06.01.2019

Elser H, Heutmann T, Lindemann M, Schmitt R (2018) Clairvoyance in Four Steps. Data analytics for digital quality management. QZ Quality and Reliability, 02/2018. Available online at https://www.qz-online.de/qz-zeitschrift/archiv/artikel/data-analytics-fuer-das-digitale-qualitaetsmanagement-4912074.html

Erner M (2018) Management 4. 0 - Unternehmensführung Im Digitalen Zeitalter. Gabler, Berlin

Franchetti M, Yanik M (2011) Continuous improvement and value stream analysis through the lean DMAIC Six Sigma approach: a manufacturing case study from Ohio, USA. IJSSCA 6(4):S278. https://doi.org/10.1504/IJSSCA.2011.040676

Gorry GA, Scott Morton MS (1971) A framework for management information systems

Haas M, Schiersmann C, Weber PC (2012) Kompetenzprofil für Beratende. With the collaboration of members of the working group Quality Characteristics and Competencies. Hg. v. nfb/ Forschungsgruppe Beratungsqualität. Bielefeld

HAB Conference Proceedings (ed) (2014) Human-Centered Decision Support in Intelligently Networked Production Systems

Haettenschwiler P (2001) New user-friendly concept of decision support. vdf Hochschulverlag AG, Zurich

Huber W (2018) Industrie 4.0 kompakt - wie Technologien unsere Wirtschaft und unsere Unternehmen verändern. Transforming and changing the entire enterprise. Springer Vieweg (SpringerLink Books), Wiesbaden. https://doi.org/10.1007/978-3-658-20799-1

Institute of Electrical and Electronics Engineers; IEEE Conference on Emerging Technologies & Factory Automation; IEEE International Conference on Emerging Technologies and Factory Automation; ETFA (2015) 2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA). 8 - 11 Sept. 2015, City of Luxembourg, Luxembourg. 2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA). Luxembourg, Luxembourg, 8/9/2015 - 11/9/2015. IEEE, Piscataway, NJ

Jones N (2014) Computer science. The learning machines. Nature 505(7482):146–148. https://doi.org/10.1038/505146a

Kagermann H, Helbig J, Hellinger A, Wahlster W (eds) (2013) Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Deutschlands Zukunft als Produktionsstandort sichern. Final Report of the Working Group Industry 4.0. Working Group Industry 4.0

Klein M, Methlie LB (2009) Knowledge-based decision support systems with applications in business: a decision support approach

Kleinemeier M (2016) Von der Automatisierungspyramide zu Unternehmenssteuerungs-Netzwerken. Handbook Industry 4.0: Production, Automation and Logistics, pp 1–8

Krauß J, Permin E, Schmitt R, Große Böckmann M, Plutz M (2016) Smartglasses im Einsatz auf dem Shopfloor – Erfahrungen, Potentiale und Limitatinen. Werkstatttechnik Online 11/12:794–798

Lerch C, Gotsch M (2014) The role of digitalization in the transformation from producer to manufacturing service provider. Die Unternehmung 68(4):250–267

Lin N (2015) Applied business analytics. Integrating business process, big data, and advanced analytics. Pearson (Safari Tech Books Online), Upper Saddle River. Available online at http://proquest.safaribooksonline.com/9780133481549

Lindner F, Kostyszyn K, Grunert D, Schmitt R (2017) Smart devices in manufacturing. ZWF J Econ Factory Operat 112(10):662–665

Maropoulos PG, Muelaner JE, Summers MD, Martin OC (2014) A new paradigm in large-scale assembly-research priorities in measurement assisted assembly. Int J Adv Manuf Technol 70(1–4):621–633. https://doi.org/10.1007/s00170-013-5283-4

Mayer-Schönberger V (2015) What is big data? On the acceleration of the human cognitive process. Aus Politik und Zeitgeschichte 65(11–12):14–19

Montavon B, Peterek M, Schmitt RH (2017) Communication architecture for multiple distributed large volume metrology systems. In: 2017 IEEE International Symposium on Systems Engineering. ISSE 2017: Vienna, Austria, October 11–13, 2017: 2017 symposium proceedings. 2017 IEEE International Systems Engineering Symposium (ISSE). Vienna, Austria. IEEE International Symposium on Systems Engineering; Institute of Electrical and Electronics Engineers; Systems Council; Annual IEEE International Symposium on Systems Engineering; IEEE International Systems Engineering Symposium; ISSE. IEEE, Piscataway, pp 1–5

Montavon B, Dahlem P, Peterek M, Ohlenforst M, Schmitt RH (2018) Modelling machine tools using structure integrated sensors for fast calibration. JMMP 2(1):14. https://doi.org/10.3390/jmmp2010014

Peters D, Schäfer PM (2018) Overview of software development topics for the digitalization of industry. German Aerospace Center. Available online at https://core.ac.uk/download/pdf/154791828.pdf

Pfeifer T, Schmitt R (2014) Masing Handbuch Qualitätsmanagement. 6th, revised edn. Available online at https://doi.org/10.3139/9783446439924.

Pfrommer J, Grüner S, Goldschmidt T, Schulz D (2016) A common core for information modeling in the industrial internet of things. at-Automatisierungstechnik 64(9):1. https://doi.org/10.1515/auto-2016-0071

Plorin D, Jentsch D, Riedel R, Müller E (2013) Ambient Assisted Production; Konzepte für die Produktion 4.0 unter Berücksichtigung demo-graphischer Entwicklungen. Werkstatttechnik online 103:135–138

Prenzel A (2018) Designing human-computer interaction for decision support in planning systems

Refflinghaus R, Kern C, Klute-Wenig S (eds) (2016) Qualitätsmanagement 4.0 - Status quo! Quo vadis? Report on the GQW annual conference 2016 in Kassel. Society for Quality Science; Society for Quality Science Annual Conference; GQW Annual Conference. Kassel University Press, Kassel (Kasseler Schriftenreihe Qualitätsmanagement, Vol. 6)

Robinson CJ, Malhotra MK (2005) Defining the concept of supply chain quality management and its relevance to academic and industrial practice. Int J Prod Econ 96(3):315–337. https://doi.org/10.1016/j.ijpe.2004.06.055

Schmitt R (ed) (2016) Smart Quality - QM im Zeitalter von Industrie 4.0. 20th Business Forum Qualität; 12. and 13. September 2016, Aachen. Business Forum Qualität, 1st edn. Apprimus Verlag, Aachen

Schmitt R, Pfeifer T (2015) Qualitätsmanagement. Strategies - methods - techniques. 5th, updated edn. Hanser, Munich. Available online at http://www.hanser-elibrary.com/isbn/9783446440821.

Schmitt R, Heinrichs V, Laass MC (2014) Opinion always with you. How posts from social networks can benefit companies. QZ Quality and Reliability, 06/2014. Available online at https://www.qz-online.de/qz-zeitschrift/archiv/artikel/wie-beitraege-aus-sozialen-netzwerken-unternehmen-nutzen-koennen-737318.html, last checked on 07 Jan. 2019

Schmitt RH, Voigtmann C (2017) Sensor information as a service - building block of networked production. tm –Technisches Messen 84(5). https://doi.org/10.1515/teme-2017-0006

Schmitt RH, Peterek M, Morse E, Knapp W, Galetto M, Härtig F et al (2016) Advances in large-scale metrology - review and future trends. CIRP Annals 65(2):643–665. https://doi.org/10.1016/j.cirp.2016.05.002

Schuh G, Riesener M (2017) Produktkomplexität managen. Strategies - Methods - Tools. With the collaboration of Stefan Breunig, Christian Dölle, Manuel Ebi, Michael Gerrit Schiffer, Sebastian Schloesser and Elisabeth Schrey. 3rd, completely revised edition. Hanser, Munich. Available online at https://doi.org/10.3139/9783446453340.

Shim JP, Warkentin M, Courtney JF, Power DJ, Sharda R, Carlsson C (2002) Past, present, and future of decision support technology. Decision Support Syst 33(2):111–126

Spath D (ed) (2013) Produktionsarbeit der Zukunft-Industrie 4.0. Fraunhofer Verlag, Stuttgart

Steinhilper R, Westermann HH, Butzer S, Haumann M, Seifert S (2012) Making complexity measurable. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 107(5):360–365

Stimmel CL (2015) Big data analytics strategies for the smart grid. CRC Press, Boca Raton. Available online at http://proquest.tech.safaribooksonline.de/9781482218282

van der Aalst W (2016) Process mining. Data science in action. Second edition. Springer, Heidelberg. Available online at http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1203872

van Dijk E (2018) Process Control & Automation A new approach to data-driven process improvements. How self-service industrial analytics combined with a DMAIC cycle approach can fuel continuous improvement projects. Hg. v. Processing. Available online at https://www.processingmagazine.com/new-approach-data-driven-process-improvements/, last updated at 25.10.2018, last checked at 06.01.2019

Verissimo P, Bessani A, Pasin M (2012) The TClouds architecture. Open and resilient cloud-of-clouds computing. In: 2012 IEEE/IFIP 42nd International Conference on Dependable Systems and Networks workshops (DSN-W). 25 - 28 June 2012, Boston, MA, USA. 2012 IEEE/IFIP 42nd International Conference on Dependable Systems and Networks Workshops (DSN-W). Boston, MA, USA. IEEE Computer Society; International Federation for Information Processing; IEEE/IFIP International Conference on Dependable Systems and Networks; DSN-W; Workshop on Open Resilient Human-aware Cyber-physical Systems (WORCS); Workshop on Fault Tolerance for HPC at Extreme Scale (FTXS); Workshop on Recent Advances in Intrusion Tolerance and Resilience (WRAITS); International Workshop on Dependability of Clouds, Data Centers and Virtual Machine Technology (DCDV). IEEE, Piscataway, pp 1–6

von Hahn N (2018) Die digitale Zukunft wird nicht von allein gut-Der essenziell menschliche Faktor der Digitalisie-rung und Führung in digitaler Transformation. Springer Gabler (Digitale Führungskräfteentwicklung), Berlin

Wolfangel E (2016) Die Grenzen der künstlichen Intelligenz. Spectrum 2016-05-03 (18). Available online at https://www.spektrum.de/news/die-grenzen-der-kuenstlichen-intelligenz/1409149, last verified 06 Jan 2019

Manufacturing Technology 4.0

Achieving a Production Economy with Secure Audit Trails and Distributed Production Chains

Thomas Bergs, Fritz Klocke, Daniel Trauth, and Jan Rey

Contents

"Digital is the main reason just over half of the companies on the Fortune 500 have disappeared since the year 2000." — Pierre Nanterme, CEO of Accenture

---

T. Bergs (✉) · F. Klocke · D. Trauth · J. Rey
WZL, RWTH Aachen University, Aachen, Germany
e-mail: t.bergs@wzl.rwth-aachen.de; F.Klocke@wzl.rwth-aachen.de; D.Trauth@wzl.rwth-aachen.de; J.Rey@wzl.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_21

1 Introduction

Industrial Revolutions are defined as fast and far-reaching socioeconomic transformations caused by disruptive technologies (Gehrke et al. 2015, p. 6). The First Industrial Revolution shifted society from an agricultural to an industrial model, thanks to advances in transportation and early mechanization through steam power. This was the beginning of machine work (see Fig. 1).

The Second Industrial Revolution expanded the first. Electricity enabled the mass production of automobiles, trains, and telegraphs, which led to the age of mobility and telecommunication. In the Third Industrial Revolution, digitalization was introduced through the invention of the computer and the Internet. These inventions helped to link physical objects and value chains together. The emergence of a new degree of efficiency and automation transformed entire sectors, displaced established corporations, and destroyed traditional business models.

Today, we are on the brink of the Fourth Industrial Revolution. This next phase melds the physical realm (the First and Second Industrial Revolutions) with the

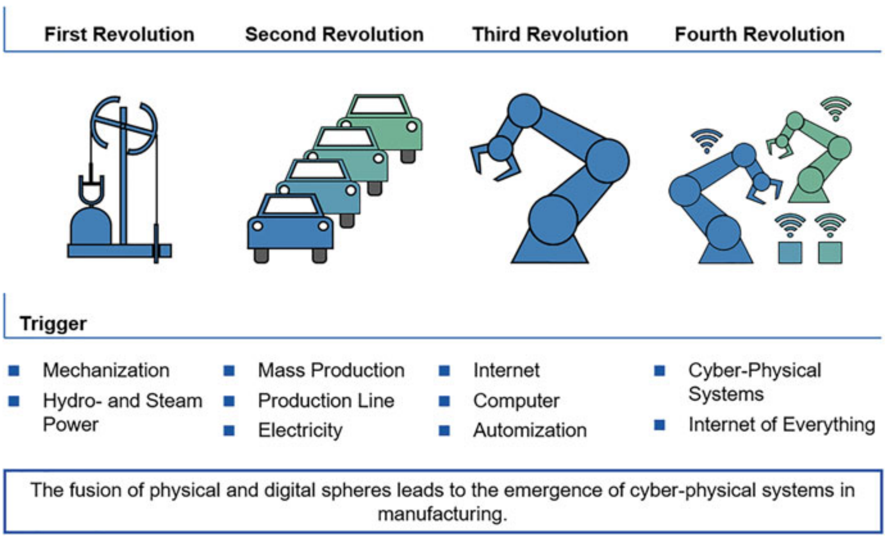


Fig. 1 Industrial development [1]

digital realm (the Third Industrial Revolution), which is made possible especially through the use of real-time data.

This results in the emergence of what are known as cyber-physical systems. In the context of production, these refer to physical objects (e.g., machine tools) that are equipped with microcomputers, as well as sensors and actuators (Bischoff 2015, p. 31). This means that, just as with earlier Industrial Revolutions, machines (in the form of these cyber-physical systems) will play a dominant role. The chief difference lies in the fact that, thanks to the increasing convergence of various new technologies, these systems will become increasingly intelligent without actually reaching the point of self-consciousness.

1.1 The Industrial Internet of Things

The Internet of Things (IoT) is the network consisting of all physical devices and systems that are connected to the internet. It enables interaction and data exchange between people and devices. The IoT market has experienced strong growth in the past, and it will continue to grow exponentially in the future as well. Current studies predict that there will be between ten and twelve connected devices per person by 2025. Smart cities and the Industrial IoT sector (Industrial Internet of Things—IIoT) are the main drivers of the IoT market, with market capitalization of approximately USD 267 billion in the year 2020 (see Fig. 2) (Cisco 2017; Statista 2019).

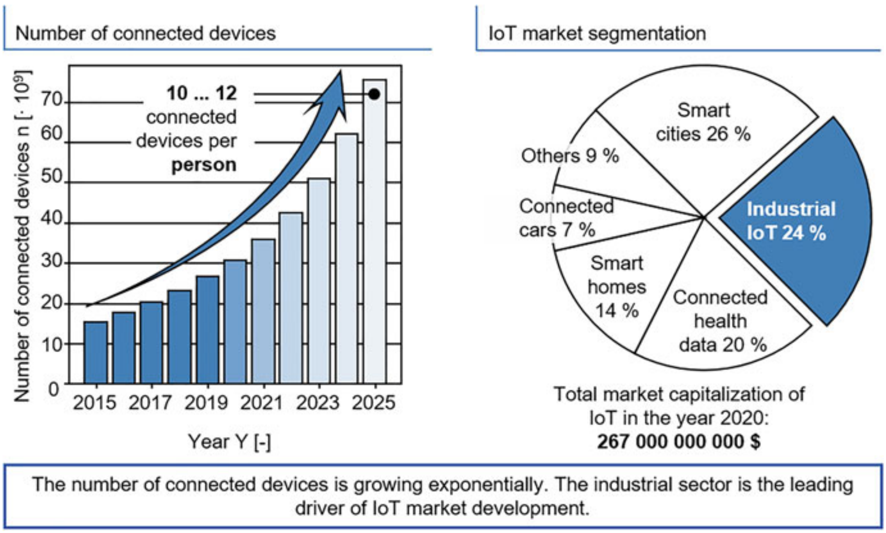


Fig. 2 Growth and market capitalization of the IoT market

1.2 Internet of Production

A direct application of the IoT approach in manufacturing engineering is not sufficiently feasible at present, since this area has significantly more parameters and considerably less available data than other big-data application domains (e.g., the financial sector). Although modern production is, in fact, characterized by large amounts of data, this data is neither easily accessible and interpretable nor sufficiently interconnected that knowledge may be generated from it. With the Internet of Production (IoP) initiative, the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University and RWTH Aachen University aim to pave the way for a new dimension of cross-domain cooperation where semantically adequate and context-sensitive data from production, development and usage are provided in real time at an appropriate level of granularity (Schuh 2017, p. 7) The core scientific approach is the introduction of digital shadows (also known as *digital twins*) as purpose-oriented, aggregated, multiperspectival, and persistent data sets. In the future, the *Internet of Production* cluster of excellence will design and implement a conceptual reference infrastructure for the Internet of Production that will make the generation and application of digital twins possible. To make the IoP a reality, Aachen's renowned research scientists from the areas of production engineering, IT, materials engineering, and other critical disciplines are working together to solve interdisciplinary challenges. This includes, for example, the integration of reduced production-engineering models into data-driven machine learning for cross-sector knowledge generation and context-adaptive initiatives (Schuh 2017, p. 8). The IoP is used by production engineers to support a new kind of holistic work in and with systems through the design and further development of engineering tools, methods, and processes. An integrated development approach is therefore a necessity for the entire field of production engineering.

1.3 The Production Economy

In terms of content, the Internet of Production particularly encompasses the areas of data collection, data modeling, and data exploitation. These topics are spanned by the overarching question of how the value of data can be quantified and how this data can be traded between machines. This leads to the introduction of the term *production economy*, where data is understood as an economic resource rather than strictly as information. In a production economy, self-monitoring and autonomous machines, devices, and systems are capable of ordering services, organizing their own production activities, and making decisions with the trust of their owners (Rajasingham 2017, p. 7). The offer of these services will initially involve humans as well, but will increasingly be handled by other machines (see Fig. 3). A production economy therefore offers the potential to capitalize the data collected and

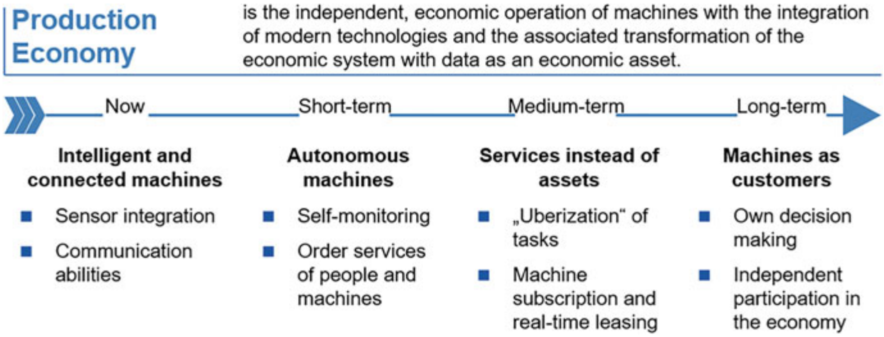


Fig. 3 Definition of production economy and their long-term effects

processed in the IoP and to trade it in an economic system. In this respect, the data can be understood as a refined and traded resource ("data as the new oil").

In the future, industrial companies will increasingly avoid purchasing expensive devices and machines. Instead, the expectation is of an "Uberization" of self-managing facilities that share their services in a decentralized ecosystem. In this context, so-called *machine-subscription models* will play an important role in the future. The basic idea behind this kind of model is that a machine user will no longer pay for the machine as a capital good, but rather for a value proposition for a specific benefit, such as for a defined productivity level or a share in the finished parts produced and accepted. Also to be expected is that machines will increasingly become independent market participants and financial actors with their own bank accounts and payment systems. Smart contracts are the prerequisite for this. A *smart contract* refers to a secure and non-terminable computer program that constitutes an agreement.

This computer program is run automatically, and the agreement is automatically enforced. Thus, breach of contract—even between parties that do not know or trust one another—is made impossible through the use of a smart contract (Bashir 2018, p. 53).

Six pillars define a production economy. Machines and facilities must be digitalized with the help of *sensors* (1) that make machine conditions visible and enable *machine-to-machine communication* (2) in both directions, transmission and reception. With the help of *artificial intelligence* (3), these machines can work on their own in a *decentrally* (4) structured *sharing economy* (5) in which the operators no longer define value through ownership (Rüther 2017). The backbone of this kind of production economy is any form of *distributed ledger technology* (6) that enables the use of smart contracts and a trustworthy exchange of data between devices.

1.4 Interim Conclusion

As a result of the Fourth Industrial Revolution, smart, interconnected and autonomous cyber-physical systems will emerge. In light of this, autonomous machine-to-machine (M2M) transactions will make a new production economy possible that is characterized by the independent, economic operation of machines. The most important resource in a production economy is data. However, the unalterable recording and storage of this data poses a challenge. Distributed Ledger Technologies (DLT) represent a promising approach toward solving this problem. These technologies will be described in more detail in the following chapter.

2 Distributed Ledger Technologies

Ledgers, the foundation of accounting, are as old as writing and currency. In the past, their storage devices were clay, wood, stone, papyrus and paper. After the widespread distribution of computers in the 1980s and 1990s, paper files became digitalized, often through manual data entry. These digital books imitate the cataloging and bookkeeping of the paper-based world. Paper-based institutions remain the backbone of our society. Examples of this are cash, seals, signatures, invoices, certificates and the use of double-entry accounting. Increases in computing power and breakthroughs in cryptography, together with the discovery and use of new algorithms, made the development of distributed ledgers possible. In its simplest form, a distributed ledger is a decentralized database that is administered and updated independently by every participant (or node) in a large network. Its distribution is truly unique. Data sets are not communicated to various nodes from a central location, but rather are created and maintained independently by each node. This means that every individual node in the network processes every transaction, reaches its own conclusions and votes on these conclusions to ensure that the majority agrees with what is known as the consensus. As soon as this consensus is reached, the distributed ledger is updated and all nodes retain their own identical copy of the ledger. This architecture enables a new kind of data storage, processing and use that goes beyond the usefulness of a simple database (Bauerle 2018a).

2.1 From Data Silos to Distributed Ledgers

Distributed ledgers are a dynamic form of data storage devices and possess characteristics and capabilities that make it possible to formalize and to secure new kinds of relationships in the digital world (Yaffe 2017). The core of this new kind of relationship is that the costs of trust (previously borne by notaries, lawyers, banks,

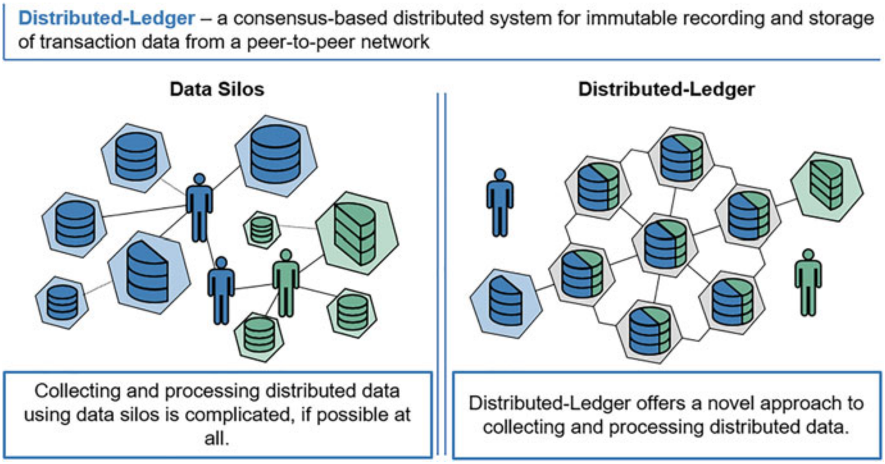


Fig. 4 From data silo to distributed ledger

supervisory authorities, governments, etc.) are avoided through the architecture of the distributed ledger (see Fig. 4).

The limited automated interoperability between the data silos is offset by the human interface between the silos. The invention of distributed ledgers represents a revolution in the way information is gathered and communicated. This applies to static data (e.g., registers) as well as dynamic data (e.g., transactions). Distributed ledgers make it possible for users to go beyond the simple use of a database for storage, modification and extraction. DLTs are not made possible by the employment of a new technology.

They are based on a unique combination of three existing technologies: peer-to-peer networks, cryptography and data structure.

2.2 *Characteristics of Distributed Ledgers*

A distributed ledger is a chain of time-stamped, cryptographically secured, unalterable blocks of consensus-validated digital data. These blocks exist in numerous synchronized and geographically distributed copies (Bauerle 2018b, c). The technology behind distributed ledgers can be applied in the following areas: (a) *to secure the immutability of the data,* (b) *to create a digital identity,* (c) *as a recording system, and* (d) *to be made available as a platform* (see Fig. 5).

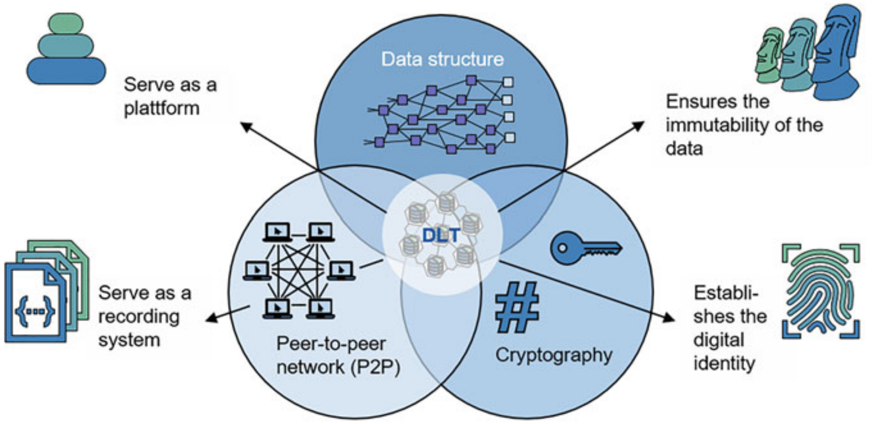


Fig. 5 Characteristics of distributed ledger

2.2.1 Securing the Immutability of Data

The chief characteristic of DLT is that its own change history is saved. For this reason, it is often described as unalterable or immutable. While changing an entry in the database would theoretically be possible, it would take enormous effort, since all subsequent data would have to be changed at every single node. The strengths of Distributed Ledger Technology can therefore be exploited especially when using it as a recording system rather than as a database.

2.2.2 Creating a Digital Identity

A digital identity in a Distributed Ledger Technology is achieved using cryptographic keys. The combination of a public and a private key creates a powerful digital identity reference based on how the ownership of the private key is structured. Like a mailbox, a public key is publicly known. A private key, in contrast, involves consent to an interaction, like a mailbox key. This leads to the creation of a digital fingerprint for the data. Cryptography is therefore the key technology for making DLT secure.

2.2.3 Use As a Recording System

DLTs represent a major innovation in the areas of information gathering and distribution. They are suitable for collecting static data (register) as well as dynamic data (transactions) and thus represent the next evolution of recording systems. In the case of registers, the data can be stored in one of three ways (Bauerle 2018c): *Unencrypted* data can be read by every participant in the DLT and is fully transparent. *Encrypted* data can only be read by participants with a decryption key. The key provides access to the data and can prove who added the data at what point in time. *Hash data* is created through the application of a so-called hash function (also known as a hashing algorithm) to a data set. During this process, the data set is transformed into a character string with a predefined length. Hash data can be used to demonstrate that the data has not been manipulated. For this reason, it is often saved in the ledger as a digital fingerprint, whereas the majority of the original data is stored in a database outside of the ledger.

2.2.4 Use As a Platform

The first DLT-based platform was a cryptocurrency. Recently, however, smart contracts have gained increasing importance. These smart contracts expand the areas of application for platforms based on DLT (see chapter 1.3).

2.3 DLT Data Structures

The most well-known Distributed Ledger Technology is the blockchain. The analogy of a book can help to clarify the principle behind a blockchain. Every block represents a page in a book, and the blockchain represents the binding. A transaction within a block is like an entry in a line on this page. The difference is that another block can only be added to the blockchain if a distributed, decentralized consensus (validation) has been reached (see Fig. 6). When the book goes to print, the bookbinder puts the pages together and secures validation centrally. The validation ensures that the pages, or blocks, are in the correct order.

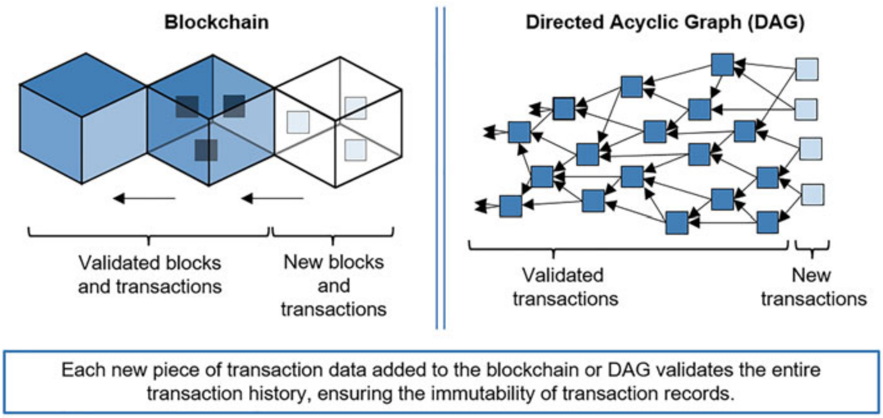


Fig. 6 Types of DLTs: Blockchain and DAG

Just as with book binding, the validation of the blocks is very costly in terms of time and energy. The pages of the book must be pressed together under high pressure for considerable periods until all the pages hold together and the blocks become a chain. In the world of the blockchain, this process is referred to as *Proof-of-Work* (PoW). This refers to a consensus mechanism that ensures that a specific unit of computing power is expended by solving a cryptographic puzzle. Thanks to the hash value calculated because of this process, every participant can verify that the corresponding unit of computing power has been spent (Bashir 2018, p. 54). As compensation for the resulting calculation costs, the computer that first reaches a consensus receives financial compensation, typically in the form of the system's own cryptocurrency.

Alongside blockchain, there are other Distributed Ledger Technologies that are based on a Directed Acyclic Graph (DAG). From a technical perspective, a blockchain is a one-dimensional DAG. However, since "blockchain" has already become an established term, this article will use the terms "blockchain" and "DAG" for the sake of simplicity. DAG, then, is usually understood as a graph with at least two dimensions. Unlike blockchain, the DAG does not group transactions together in blocks, but instead connects them directly to a structure. The result is no longer a one-dimensional line, but rather a graph that expands horizontally and vertically. With blockchain, a comparatively difficult cryptographic puzzle must be solved to attach a specific number of transactions to the blockchain in the form of a block. The cryptographic puzzles employed by the DAG, in contrast, are significantly easier to solve, so that even individual transactions can be attached to the structure. This reduces the resources necessary to achieve PoW. Since it is not worth paying compensation for the resources that had to be used during PoW, transactions are frequently free of charge. These so-called feeless transactions lead to significantly better scalability of transaction throughput compared to the blockchain (Tam 2018). Furthermore, the transaction rate increases along with the number of participants in the system, since this means larger numbers of participants are involved in the validation processes, which likewise improves the scalability of the system.

In addition to their data structure, DLTs are also defined in terms of their accessibility. The most popular implementation of the blockchain is Bitcoin. Bitcoin refers to a DLT network that anyone can participate in (permissionless). A freely accessible system offers a number of advantages, such as a lower risk of being hacked by malicious participants, while simultaneously making access to its contents possible for everyone. This explains the existence of blockchain implementations for regulating access, for example to protect the sensitive data of certain institutions from abuse.

The possibility of network participation for any participant (permissionless) also exists in the DAG. The most well-known permissionless DAG implementation is the Tangle network run by the IOTA Foundation in Berlin, Germany. The IOTA Foundation's distributed ledger is called Tangle. In theory, all DLTs are fundamentally suitable for a production economy; however, the aims of the IOTA Foundation and the perfected implementation of the Tangle are particularly well suited to the

Industrial Internet of Things. In contrast to blockchain-based systems, the Tangle involves no transaction costs and achieves network scalability.

2.4 Interim Conclusion

Distributed Ledger Technology (DLT) is a novel approach to gathering and storing transaction data. DLT is based on the use of special data structures (blockchain or DAG), cryptography, and peer-to-peer network architectures that are controlled by an algorithm. DLT transactions are transparent, reliable, and incorruptible. DLTs possess special functions. They can be used to secure immutability, to create a digital identity, and to serve as a platform or recording system. IOTA refers to a permissionless distributed ledger developed specifically for IoT and the production economy.

3 From the Internet of Production to a Manufacturing Economy

Now that the fundamentals of DLTs and IOTA as drivers of a successful production economy have been discussed, the following section will present the prerequisites for the integration of DLTs into business models as well as some practical case studies based on DLTs. While IOTA does represent a suitable option for monetizing data transactions, the question of whether DLTs are advantageous for a given business model or use case must be examined for every individual case. Since the focus of application in this article is on the use of data from value-adding manufacturing processes and non-value-adding production engineering stages (e.g., transport stages) are not taken into consideration, the following will no longer speak of the production economy, but of the manufacturing economy.

3.1 Prerequisites for DLT Business Models

Based on a study by PricewaterhouseCoopers (PwC), various conditions must be met before DLTs can usefully be integrated into a business model (PwC and Long Finance 2016) (see Fig. 7). DLTs exhibit great potential for a use case when the following conditions apply:

– Multiple participants must be able to view common information. This is why they share a common dataset.
– Multiple participants perform actions that must be recorded and update the data. This is why a decentralized update policy is necessary.

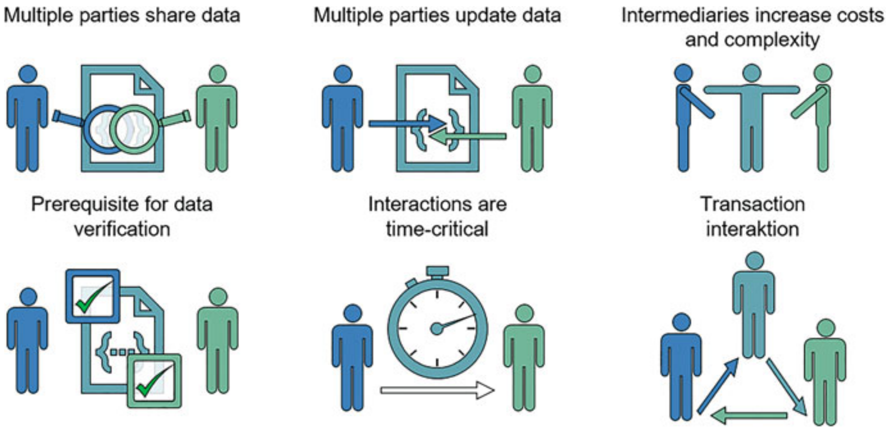


Fig. 7 Characteristics of a feasible DLT use case

- Participants must be able to trust that the recorded actions are valid. This is why they need a platform for data verification.
- The removal of central intermediaries promises the potential to reduce costs (e.g., fees) and complexity (e.g., the number of necessary voting processes in a system).
- Participants must be able to act in a timely manner; for time-sensitive tasks, the reduction of delays is therefore economically advantageous (e.g., lower settlement risk, better liquidity).
- Transactions created by different participants are independent of one another.

3.2 Use Cases

By using a combination of the PwC characteristics (business level) and the previously described technological features (technology level) possessed by IOTA (immutability, digital identity, platform, recording system), it is possible to identify potential use cases (see Fig. 8).

As a rule, use cases are interesting opportunities for DLTs when there is need of a common data set that must be updated and verified by multiple participants. In such instances, DLTs can provide a powerful acceleration of business operations. DLT use cases that can be derived from this therefore include asset sharing, M2M communication, data marketplaces, distributed manufacturing, supply chain tracking, digital product memory, the inspection of spare parts, and quality documentation. Select use cases are described in more detail below.

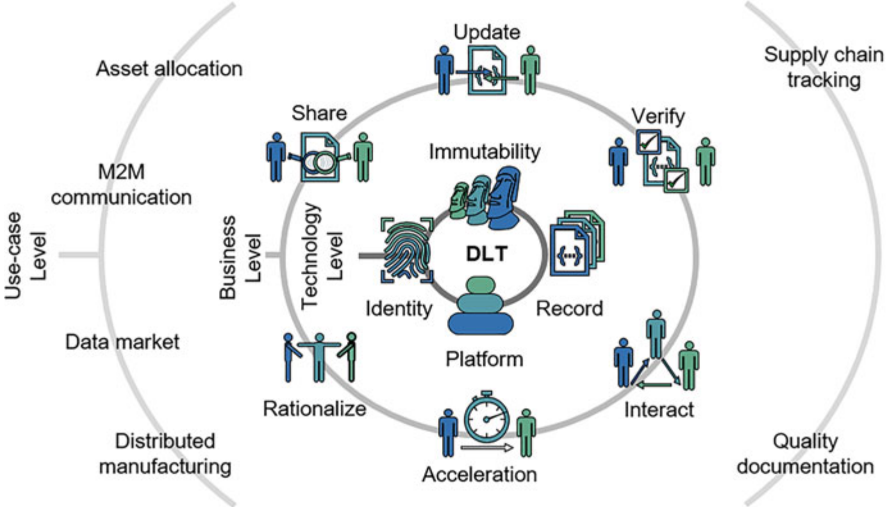


Fig. 8 Derivation of DLT use cases

3.3 Example: Marketplace for Unalterable and Verifiable Data

The concept of a data marketplace can be illustrated using an example from quality assurance. Across all companies engaging in production, various measuring tools are used to check functionally relevant component properties for their required specifications. Some of the measurements that occur in a company have already been conducted by other entities in the supply chain. Because no company-wide access to such quality control data has been available to date, measurements must be repeated at significant cost and measuring equipment must also be acquired, partly at high investment costs (e.g., air-conditioned measuring rooms). SMEs in particular stand to profit therefore from extensive economic benefits thanks to the exchange of data along the supply chain (Sønstebø 2018).

Furthermore, artificial intelligence in general and advances in machine learning harbor the potential of significantly improving company processes through a data lake spanning the entire value chain. The possibility of discovering new process patterns and visualizing implicit knowledge arises as a result. This leads to improved products and processes as well as to new business models. The prerequisite for this is a secure data marketplace that makes unalterable and verifiable data available in real time.

Here it is of secondary importance whether the data itself is stored in the Tangle or only its signature. In addition, the question of whether the data is stored locally, in the cloud, or on a portable drive is irrelevant. Thanks to the data signature that is referenced in the Tangle, the integrity of the data can be ensured at all times (Pritam

2018). This means that the reliability of the data is always guaranteed in the case of distributed storage or computing power.

3.4 Example: DLT-Based Supply Chains

Currently, within supply chains themselves, there is a frequent lack of trust and transparency along the upstream and downstream value creation stages. Documented products of origin for items can ensure that these items cannot be altered, counterfeited, reproduced, or stolen. They therefore harbor the potential to significantly boost trust and transparency in a supply chain. An example for a documented product of origin is a digital twin used jointly by participants in the DLT system. A digital twin in this context refers to the digital model of an actual process, product, machine, or service that possesses a unique, unalterable identity. In this regard, the DLT makes data integrity possible—and thereby secure audit trails (Waschbusch 2018).

3.5 Validated Manufacturing Sequences

One example of a DLT-validated manufacturing sequence was established by Fujitsu. In this case, a so-called component audit trail was implemented, whereby only authentic components that had passed through all manufacturing stages in the required order were accepted and passed on by robots. This guaranteed proof of authenticity and origin, as well as data access from any location. Additionally, DLT enabled the unalterable storage of robot sensor data in a secure communication channel. This data taken from the entire life cycle of a robot could then be used and monetized for various applications, allowing for the establishment of an auditable robot life cycle. In this way, usage data from the robots, for example, was used to adapt maintenance cycles more efficiently to the individual robots (Waschbusch 2018).

3.6 Interim Conclusion

DLT or IOTA applications reveal their potential particularly when interactions between multiple independent parties are based on the same set of data. DLT and IOTA allows for the creation of a secure data marketplace for machines and humans where the integrity and verifiability of data is guaranteed. The characteristics of DLT and IOTA enable the use of digital twins that create secure audit trails within decentralized systems. DLT and IOTA are new technologies and their applicability in the manufacturing context remains largely unexplored.

4 Example Application: Secure Audit Trails for Fine Blanking

The object of this example application is a fine blanking press made by the company Feintool, type XFT 2500 speed. Fine blanking is a separative manufacturing process for sheet metal products. It is primarily used in serial production in the automotive and aerospace industries and in medical technology, as well as in special-purpose machine construction. Typical parts include, for example, the brake caliper carriers or the metal tongue of a safety belt. As a result, fine blanked parts must commonly fulfill safety-critical functions, which means special requirements for manufacturing. Fine blanking is usually performed from the coil. The process consists of a coil and reel system, a guide plate, a lubrication unit, and a fine blanking press with a triple-action tool. Guiding this process requires an average of 1000 signals.

Due to variations in the physical material and process uncertainties, it is statistically impossible to manufacture parts with identical metallurgic and geometric characteristics. Although all parts must fulfill the same stringent requirements, not all parts can fulfill these requirements due to their individual characteristics.

However, if it were possible to create a digital twin of every individual part and to share this distinct information across the value chain, subcontractors and suppliers could better adapt their downstream steps to the component characteristics and improve the end customer's trust in safety-relevant parts, assemblies, or products.

The aim of this example application is to extract the production data of fine blanked parts from machine control in real time, to store it in encrypted form in IOTA's Tangle technology, and to access it via a web-based front end. The example application is limited to select data, such as punching force, press stroke, and material designation. To achieve this, three essential steps are required: data acquisition, data processing and data visualization.

4.1 Data Acquisition

In addition to the over 1,000 control signals from the stored program control of the fine blanking press, the fine blanking tool was instrumentalized by including additional measuring technology. Force measurement sensors were employed at a sampling rate of 10 kHz: the resulting ring tooth force on the ring tooth plate was measured using the Kistler 9021A, the cutting force with the Kistler 9041A, and the counterforce with the Kistler 9031A. The laboratory device LabAmp 5167A was used as a charge amplifier and A/D converter (see Fig. 9).

The charge amplifier has an integrated network interface that enables the collection of sensor data via network stream. This data can be used for various applications parallel to normal test operation. Additionally, all sensors with analog voltage output and Integrated Electronics Piezo-Electric (IEPE) mode can be connected. LabVIEW2018 by National Instruments was used as the data collection software.

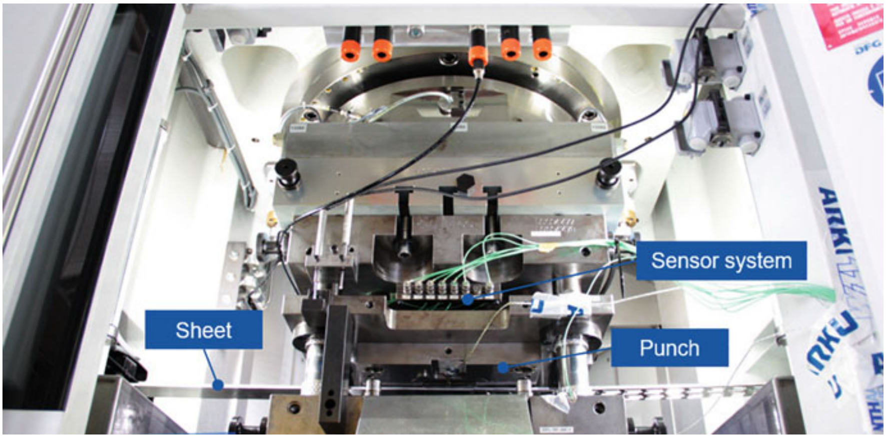


Fig. 9 Instrumentization of the fine blanking tool

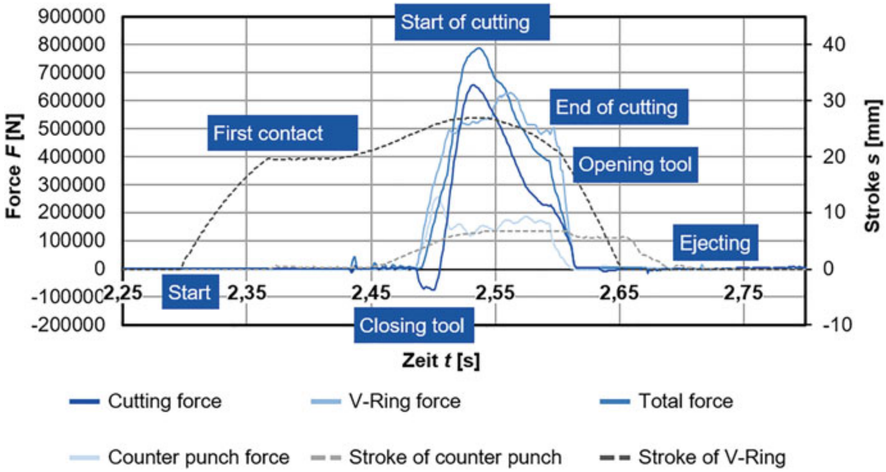


Fig. 10 Forces of fine blanking

The data collection software is connected to LabAmp and reads distinct data that is only assigned once for each device, such as the serial numbers for various components.

Together, these characteristics form a digital fingerprint. Figure 10 illustrates the high complexity involved in the process. Within a very short time (0.5 s), several forces act on the tool, all of which take on a defined task.

4.2 Data Processing

The data package in this example application includes measured process data as well as metadata. The metadata contains a distinct ID (type: integer) that is detected using digital image processing on the surface of the part, along with a material name corresponding to the international standard for the material used (type: string). Further metadata, such as the product name, the machine operator, the manufacturer, the customer, etc., can also be listed. Real machine data was also measured. This includes the maximum stamping force (type: float, unit: kN) of the ram resulting from the cutting contour and the material, as well as the punch stroke (type: float, unit: mm). Since the material characteristics and plate thickness differ from part to part due to material variations, the stamping force and the punch stroke are different for every stroke. At present, die roll (type: float, unit: mm) is estimated on the basis of existing analytical models. The time stamp (type: unsigned integer) is calculated using the UNIX time stamp, and denotes the time of production, not the time of upload. In data preparation, data signing is particularly relevant. The data should neither be stored on the Tangle in plain language nor be freely accessible to everyone. For this reason, the SHA-2-256 hashing algorithm was chosen due to its suitability for hashing large data volumes. Additionally, RSA or PKCS 1.5 was chosen as the public key encryption technology.

Data storage occurred with the help of Amazon Web Services (AWS). All data sets were stored in a DynamoDB via an Amazon API Gateway and the Amazon Lambda Service (see Fig. 11).

Calculation of Proof-of-Work (PoW) also occurred in AWS. Amazon's SQS service was also used for this purpose. Through the AWS implementation of PoW, approximately ten transactions per second were achieved. In the future, however,

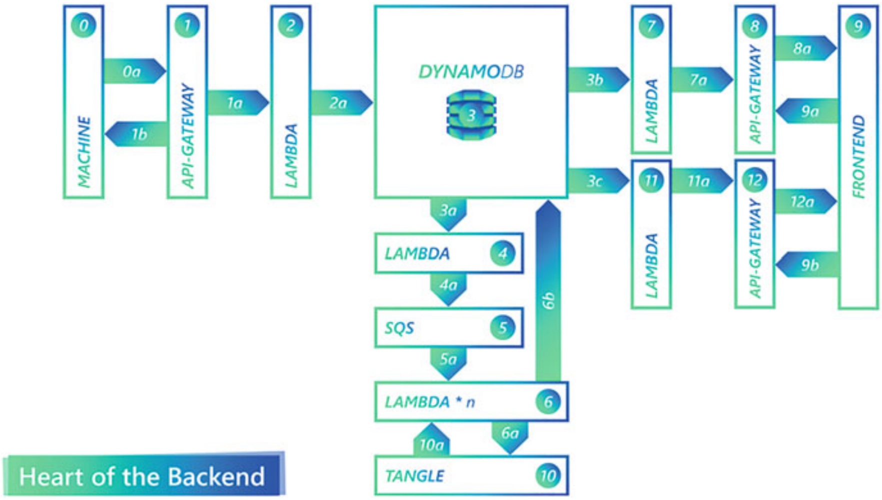


Fig. 11 Structure of the backend

Field Programmable Gate Arrays (FPGA) and EDGE servers will be tested as alternatives, which promise even higher transaction rates.

Publicly available IOTA libraries were used to attach the data to the Tangle. Encryption was achieved using Masked Authenticated Messaging (MAM).

4.3 Data Visualization

There are several ways to make transactions visible. The simplest way is to use a public explorer. All transactions listed in the Tangle network can be viewed using the Tangle explorer thetangle.org (see Fig. 12).

To make them easier to find, the transactions from the example application were designated with the tag WZL9GCX9IOTA9POC9IIOT999999. However, because the data itself is stored in a DynamoDB, and only its signatures or hashes are stored in the Tangle, Tangle explorers are not suitable for decrypting and visualizing data.

For this reason, a node.js-based front end was implemented with the help of the vue.js library. For the configuration of the front end, the following scenario was defined: The scenario assumes a manufacturing economy where a machine, a B2B customer, or a B2C customer needs access to the production data of one or more safety-critical components. It is also assumed that the manufacturer is globally active—i.e., owns multiple production sites and produces a variety of parts. Based on these constraints, the aim was to visualize production capacity utilization with the help of a diagram. Even if this is not the actual motivation behind IOTA, it was possible to demonstrate implementation (see Fig. 13).

A further diagram visualizes how production capacity utilization is distributed among various sites. Furthermore, it is possible to find all components via browser using a combination of Tangle and AWS DynamoDB. The employment of an additional database such as AWS DynamoDB is envisaged, since a database entry

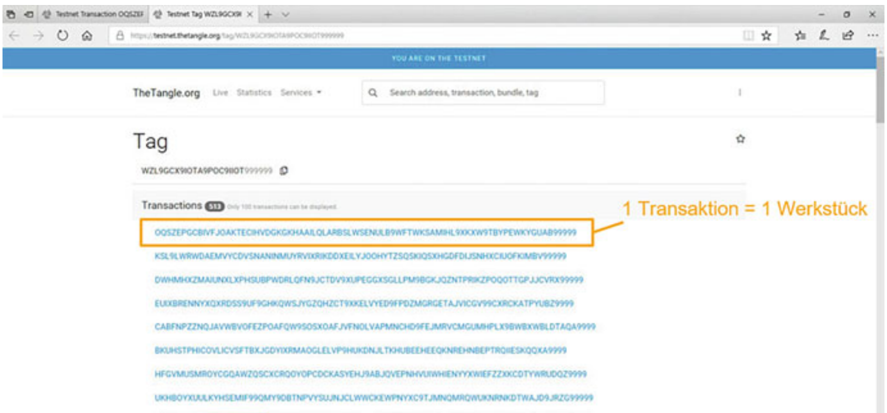


Fig. 12 Transactions in the tangle explorer

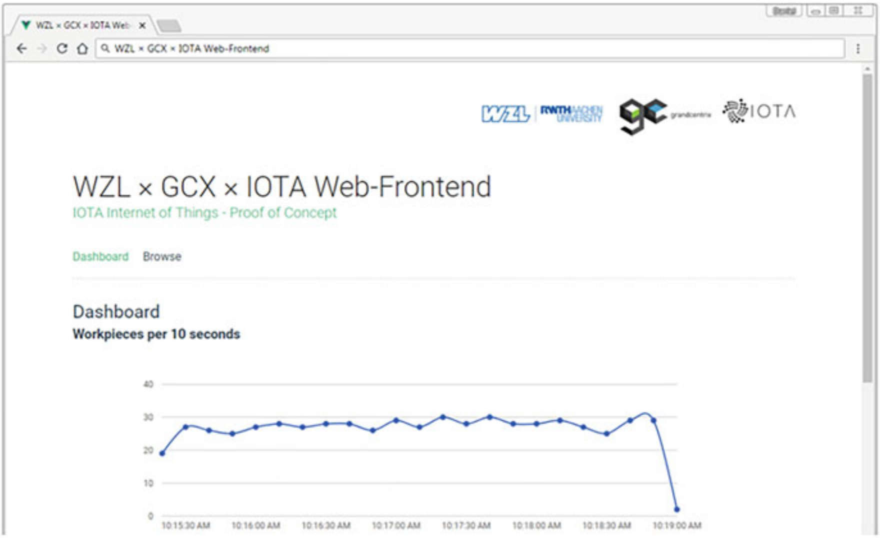


Fig. 13 Production utilization in the IOTA frontend

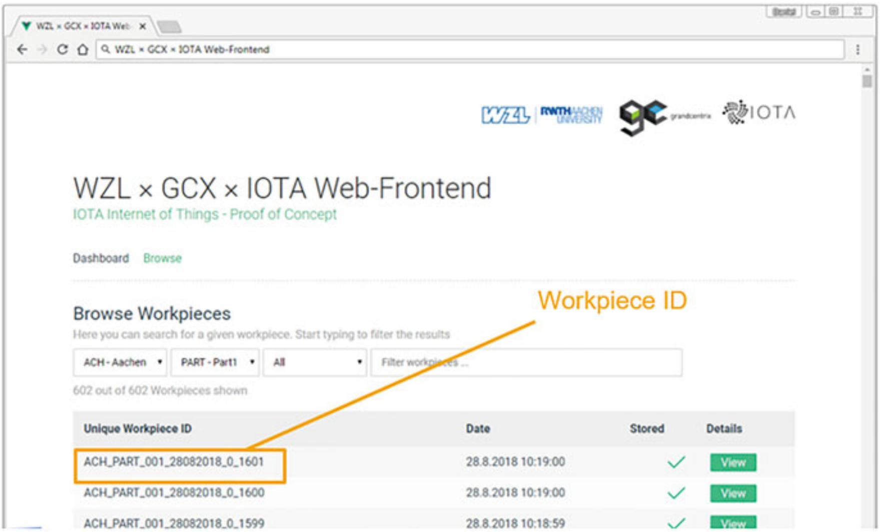


Fig. 14 Search for transactions in the IOTA frontend

of more than 1.5 kB per workpiece is planned for the future. For every workpiece ID, the corresponding data set must be extracted from the Tangle and the database (see Fig. 14).

The integrity of the database entries must be guaranteed by signature and tamper-proof storage of this signature in the Tangle must also be possible (see Fig. 15).

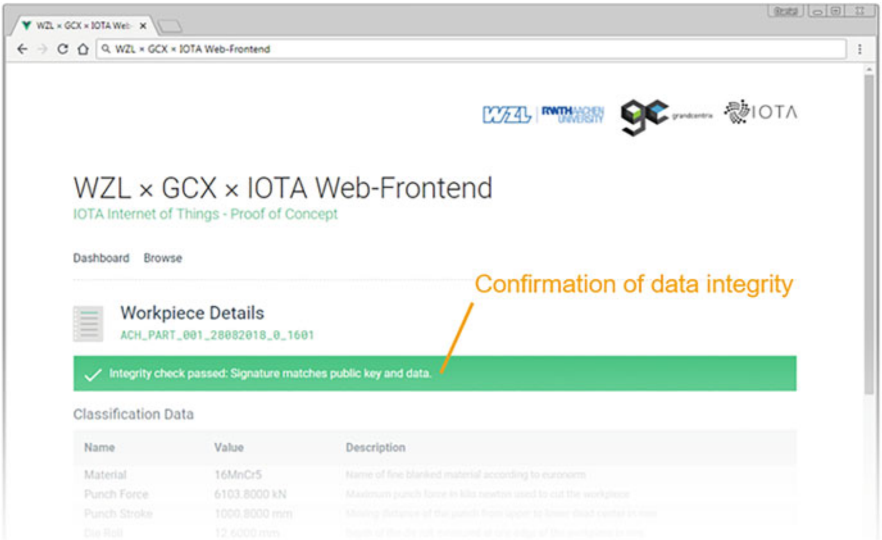


Fig. 15 Verification of the data integrity

4.4 Future Outlook

The previous segments discussed how component data and information can be used and traded in the framework of the IoP by establishing a production or manufacturing economy based on distributed ledger technologies. In the form of a future outlook, the following paragraphs will outline brief examples that illustrate which component data exhibits high relevance with regard to manufacturing technology and possesses high use potential in a manufacturing economy.

In the course of this, the generation of digital twins is of particular importance. Because the concept presented above allows for feeless transactions within a DLT system, the payment of minuscule amounts (known as micropayments) becomes possible. If, for example, a digital twin has been created based on a finite element analysis for component information, it can be coupled to an IOTA data set and may be viewed on transfer of a micropayment (see Fig. 16). To achieve this, an additional M2M connection to an FE server is established when the user makes this request and provides payment.

4.4.1 Simulation-Supported Design of Forming Processes Using the Example of the Precision Molding of Glass

High-precision optical components made of glass can be found in a wide variety of applications. Laser and medical technology, digital cameras and camera systems for automotive assistance systems are only a few examples. Common to all of them is

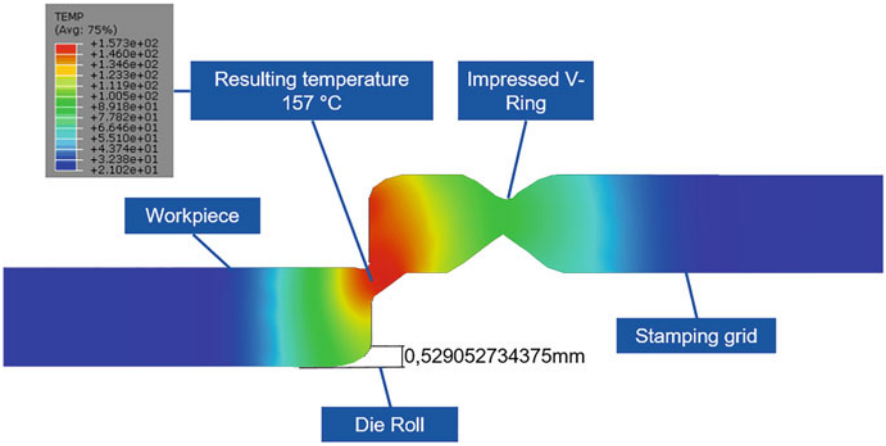


Fig. 16 Outlook: Coupling of an IOTA data set with a finite element simulation

that the optical components exhibit accuracies in the submicron range. This places tremendous demands on the manufacturing processes. One technology currently being researched at the Fraunhofer Institute for Production Technology (IPT) that has the potential to achieve both the production of large quantities as well as the required precision is precision molding. In precision molding, the glass component is heated above the so-called "transition temperature" of the glass before being pressed into the desired shape between two molds. Although this procedure is capable of drastically reducing manufacturing costs compared to conventional grinding and polishing processes, these costs remain too high for some future applications (e.g., smartphone camera modules). In this case, the IoP approaches involving the smart collection and use of process data might represent a further breakthrough in cost reduction. A key role is played here by efficient process models that, in combination with corresponding finite element (FE)-based simulation software, would permit an a priori design of the processes. This also makes it possible to observe the condition of the glass and of the molding tools (temperature, stresses) during the forming process and to intervene in case of a potentially adverse development. The forming process can thus be made "visual."

4.4.2 Model- and Data-Based Analysis and Visualization of Process Information—Digital Twins for Manufacturing

This section has already illustrated the relevance of digital twins for secure audit trails based on DLT technology. One obvious conclusion is that the creation of digital twins tailored to the respective use case harbors great potential. When it comes to Industry 4.0, "digital twin" is a popular buzzword that is often defined and understood differently. Generally speaking, "digital twin" refers to the notion of transferring existing physical products and states into a consolidated digital model.

However, there is no uniform understanding of what information is digitally transferred at what precision and at what time and for what purpose. One of the reasons for this is that every field of engineering has its own intentions when placing requirements on a digital twin. An engineer who optimizes supply chains and logistics will, for example, formulate different requirements for the digital twin than, say, one who optimizes the manufacturing process. For logistics processes, it is of greater interest at what point in time a component is located at what location in the factory, whereas the manufacturing process is more interested in the many microphysical interactions that effect the component during processing. This is the reason why the creation of digital twins for manufacturing is being researched in depth at the Fraunhofer Institute for Production Technology (IPT). Here the aim is both to recognize possible errors at an early stage and to learn from the interactions so that the manufacturing process can be further optimized. This, however, requires different forms of information processing and exploitation, which can be divided into four steps. The first and simplest step is the description of the actual process status. This can be achieved using select key figures and the contextualized, visual processing of this data. In the course of this, the collected process data or information can be preprocessed using models and algorithms to improve its informative value. The second step is analysis. Here the data is examined for previously undetected correlations. In the third step, prediction, the identified correlations can be used to predict the occurrence of future process events. In the fourth and final step, implementation, the derivation and possibly initiation of an optimal recommended action occurs based on the results of the previous three steps.

4.4.3 Planning of Adaptive Process Chains With Technology Models and Historical Process Data Using the Example of Tool Engineering

The planning of process chains in single-item and small-batch production has historically been designed to achieve the maximum capacity utilization of the existing machines. Here process design is based on the experiential knowledge of the staff involved and is supported, among others, by the use of CAM simulations.

Thanks to the research of the past few years, however, it is now possible to describe the individual technologies and their results in more detail through the use of technology models. Alongside the capacity planning, it is thus also possible to predict target values such as component quality or tool wear. As a result, these technology models support the optimization of the predictive capability of manufacturing technologies, so that process chain design can be based on objective data. Moreover, the evaluation of historical process data guarantees that the process chain is adapted to company-specific constraints and can therefore be adaptively configured at any time. The planning of adaptive process chains with technology models and historical process data is a further subject of in-depth research at the Fraunhofer Institute for Production Technology (IPT).

5 Summary

The Fourth Industrial Revolution is leading to the emergence of smart, interconnected, and autonomous cyber-physical systems. The possibility of machine-to-machine transactions between these CPS is creating a new production economy where data is the most important resource. A production economy is characterized by the independent, economic operation of machines in an economic system where data can be traded as an economic resource.

Distributed Ledger Technologies represent a novel approach to the recording, storage and management of this data. These technologies are based on the use of cryptography, peer-to-peer network architectures and new types of data structures. DLT makes the integrity and verifiability of data in decentralized systems possible, thus serving as the foundation of a production economy. IOTA refers to a distributed ledger that was designed specifically for a production economy and the IoT. DLT holds great technological and commercial promise particularly for use cases that require the shared use of a data set that must be updated and verified by multiple participants.

Examples include the industrial data market and audit trails in decentralized supply chains. This section demonstrated the fundamental feasibility of this kind of DLT-based secure audit trail using the example of fine blanking. Highly relevant in the context of manufacturing technology is the generation of digital twins of the manufactured components whose data sets are stored, managed and traded within a manufacturing economy based on DLT.

References

Bashir I (2018) Mastering Blockchain. Distributed ledger technology, decentralization and smart contracts explained, 2nd edn. Packt Publishing, Birmingham

Bauerle N (2018a) What is a Distributed Ledger? https://www.coindesk.com/information/what-is-a-distributed-ledger. Accessed 16 Jan 2019

Bauerle N (2018b) What Are the Applications and Use Cases of Blockchains? https://www.coindesk.com/information/applications-use-cases-blockchains. Accessed 16 Jan 2019

Bauerle N (2018c) What Can a Blockchain Do? https://www.coindesk.com/information/what-can-a-blockchain-do. Accessed 16 Jan 2019

Bischoff J (ed) (2015) Study "Erschließen der Potenziale der Anwendung von Industrie 4.0 im Mittelstand."

Cisco (2017) Cisco Visual Networking Index: Forecast and Methodology, 2016–2021

Gehrke L et al (2015) A discussion of qualifications and skills in the factory of the future: a German and American perspective. VDI and ASME, Düsseldorf

Pritam R (2018) Extending the Tangle network to a distributed cloud. https://medium.com/swlh/extending-the-tangle- network-to-a-distributed-cloud-2f5857db02f5. Accessed 16 Jan 2019

PwC/Long Finance (2016) Chain reaction: how blockchain technology might transform wholesale insurance. Z/Yen Group Limited and PricewaterhouseCoopers LLP, London

Rajasingham D (2017) Welcome to the machine-to-machine economy. Opportunities and challenges in a connected world.

Rüther M (2017) Project twin of things - digital twin
Schuh G (ed) (2017) Digital Connected Production, Aachen
Sønstebø D (2018) IOTA Data Marketplace. https://blog.iota.org/iota-data-marketplace-cb6be463ac7f. Accessed 16 Jan 2019
Statista (2019) Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025 (in billions). https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide. Accessed 16 Jan 2019
Tam KC (2018) Blockchain and the Challengers. https://medium.com/@kctheservant/blockchain-and-the-challengers-74e22cf0cf4e. Accessed 16 Jan 2019
Waschbusch L (2018) Fujitsu: Authentifizierung von Komponenten mit IOTA. https://www.industry-of-things.de/fujitsu-authentifizierung-von-komponenten-mit-iota-a-710722/. Accessed 16 Jan 2019
Yaffe L (2017) Using the Blockchain to Bridge Data Silos. https://medium.com/@lyaffe/using-the-blockchain-to-bridge-data-silos-852aa82d84eb. Accessed 16 Jan 2019

Networked Production Through Digital Shadows: Machine Tool 4.0

Christian Brecher and Mathias Brockmann

Contents

1 Industrie 4.0 Requirements for Production Technology

All over the world, the term Internet of Things (IoT) has gained traction as a concept describing the horizontal Internet-based networking of cyber-physical systems (Gubbi et al. 2013).

Within the IoT approach, field data (e.g., customer behavior) is used effectively, enabling the implementation of new types of digital business models (Porter and Heppelmann 2014). In most application scenarios within the industry, large amounts of field data are generated, and subsequently described by a relatively small number of parameters. However, the full potential of this approach cannot be fully exploited yet because access to production data is not consistently available and the data is very heterogeneous due to the complex physical interrelationships. These interdependencies and lack of availability lead to completely new challenges, especially concerning data analysis (Gandomi and Haider 2015). Using the IoT as an outline for

C. Brecher · M. Brockmann (✉)
WZL – Laboratory for Machine Tools and Production Engineering, RWTH Aachen University, Aachen, Germany
e-mail: C.Brecher@wzl.rwth-aachen.de; M.Brockmann@wzl.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_22

how such large capacities of data can be used to improve the production process is a major chance within the industry. Especially in Germany, this transfer of the IoT approach to production is called **Industrie 4.0.**

Modern production technology is characterized by numerous, so-called domain silos which contain complex, mainly empirical, and domain-specific models and data, (e.g., design and manufacturing data). The continuous specialization in each of these areas leads to a high degree of data heterogeneity, which on the one hand limits the accessibility of data and knowledge across domains. On the other hand, direct access to data from neighboring domains is hardly possible, which means that engineers often work with outdated or incorrect information from other domains. This domain-specific isolation led to the development of mostly consecutive approaches within product development often based on milestones. Production and use of technical products within their unique use phases is also affected by domain-specific isolation (Allwood et al. 2016).

A great opportunity for Industrie 4.0-technology is given within the synchronized use of domain knowledge. Access to models and data from all relevant areas could significantly increase the productivity and flexibility of production systems. Within the shortest possible time, innovations could be converted into products ready for series production, making the overall process faster and more cost-effective for producers and consumers.

Cross-domain data access will open up completely new possibilities for manufacturing companies by closing the gap between vertical integration within a domain and the entire **product life cycle**.

2 Recommendations for the Needs-Based Implementation of Data-Driven Approaches in Production

The main benefit of Industrie 4.0 is reaching an entirely new level of cross-domain collaboration. This goal will be achieved by providing semantically adequate and context-related data gathered from various phases of the production process and its results. With Industrie 4.0-techniques, data from the production, development and use phase of products gathered at an appropriate level of granularity, in real-time.

Production technology is characterized by highly parameterized models that are validated for a certain working point and therefore very ingrained in a specific context. As an example, typical models for machining require several parameters to consider the influences of workpiece material, tools, and the cooling lubricant. For other domains, however, this high number of specific parameters hinders model usability across domains. Therefore, these models must be incorporated with respect to the appropriate semantics required for the final use, the data cannot be gathered as isolated chunks.

By increasing validation data, significant parameters can be identified, and the model parameters can be reduced to the required core. In a networked production world, like the IoT aspires to provide, the increasing amount of available field data can be used by considering every process of real production as a potential experiment. The following aspects can be considered practical advice for manufacturing companies:

i. Further development of existing reference infrastructures that connect the main areas of the company to each other (e.g., product development, manufacturing, production planning and other areas)
ii. Development of principles for data modeling and aggregation based on existing knowledge (manuals, standards, expert knowledge)
iii. Integration of reduced models as a basic requirement for the use of data analytics and for cross-domain knowledge formation
iv. Increase access to data from all domains through suitable sensors, software systems and needs-based solutions
v. Further development of existing engineering tools, methods and processes based on the data provided in this way

3 The Concept of the Digital Shadow

Different approaches to reduced model parameters in production technology have already been developed in various research projects (Brecher and Özdemir 2017). However, the underlying hypotheses of these models are mostly implicit and based on physical relationships (e.g., the energy conservation law). In contrast to this form of abstract physical theories, purely data-driven models are also useful at certain operating points. However, a transfer to other operating points is usually not possible without additional validation. Nevertheless, physical models can be deduced based on the underlying physical relationships and assumptions (Gao et al. 2015). The combination of data-driven and physical-causal approaches offers the greatest potential for production technology through the networking of the IoT.

Therefore, a concept is required that enables the merging of the large amounts of heterogeneous data with the specialized production engineering models

This concept is named the **Digital Shadow**. Digital Shadows are data records meeting the following criteria:

i. specifically aggregated, multi-perspective and persistent data sets
ii. which are generated by semantically correct selection and data cleansing and
iii. can be used for reporting, diagnosis, prediction, and recommendations in domain-specific real-time

Digital Shadows can be seen as reduced data sets, which provide semantically correct statements despite being collected from such a large source.

These shadows are generated with a specific question or assignment in mind, this approach is called "pay-as-you-go principle." The results can be reused for subsequent tasks. Digital Shadows as a concept are shown in Fig. 1. The Digital Shadows are constantly optimized with each use, since the underlying hypotheses for

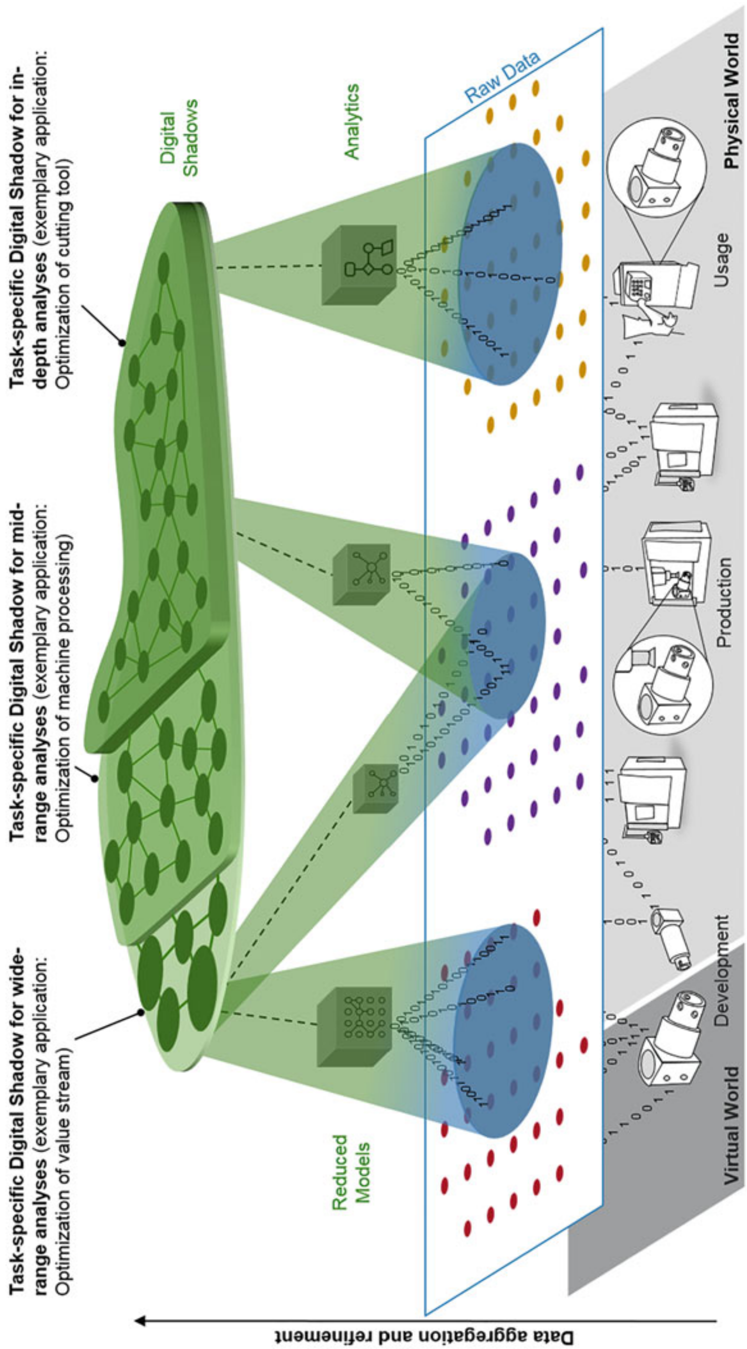


Fig. 1 Concept of the digital shadow

reduction are validated and expanded with every further experiment, similarly to a search engine on the Internet. Search engines manage the huge variety of data they are supposed to analyze through a reduced model of the Internet. The "Digital Shadow" of the search engine consists of linguistic models and sophisticated data-driven techniques. Comparable to a search engine, the Digital Shadow of production technology creates reports, diagnoses, predictions, and recommendations in domain-specific real time (instead of "just" searching) by combining reduced models with data analyses.

4 Requirements for a Reference Infrastructure for Digital Shadows

To ensure that generating, conditioning and the optimization as well as application of processes which rely on Digital Shadows is successful, a **new cross-domain reference infrastructure** has to be created.

This includes and interacts with data, models and people from development, production, and use (see Fig. 2). The core of the infrastructure is an expanded **Smart Data 1 level**.

An appropriate infrastructure provides tools for fast cross-domain decision support at various resolution levels. To reach this stage, methods of **artificial intelligence** and **machine learning** can be incorporated into the system. The people engaging with the system receive the freedom to delegate their tasks to virtual agents, these agents controlling event-based decisions, autonomous actions, and adaptive processes. They are also able to learn and make decisions themselves.

The Smart Data Layer provides multimodal access to data from the different production and life cycle domains. In this fashion, it serves as a container for Digital Shadows.

Because of these beneficial qualities, the Smart Data Layer is the next required step based on existing analytical platforms[1] and powerful Data Mining-Toolkits.[2]

The concept of Smart Data offers data integration of the relevant relationships within development, production, and use. This kind of integration is based on comprehensive data models with storage options and temporary storage options for the data, benefitting context-dependent real-time processing. This interaction in real time creates minimal latency in relation to the interaction with the Smart Expert layer. Digital Shadows are generated by using advanced analysis methods of raw and process data. These methods, such as correlation analysis and cluster algorithms, recognize patterns in data, models, and processes. At a higher level, learning

[1] Namely, Siemens Mindsphere, GE Predix, PTC ThingWorx, SAP HANA.

[2] Machine Learning Frameworks (z.B. TensorFlow, Apache Spark) or "Machine Learning-as-a-Service" providers like Amazon Web Services, Google Cloud Platform, Microsoft Azure, etc.

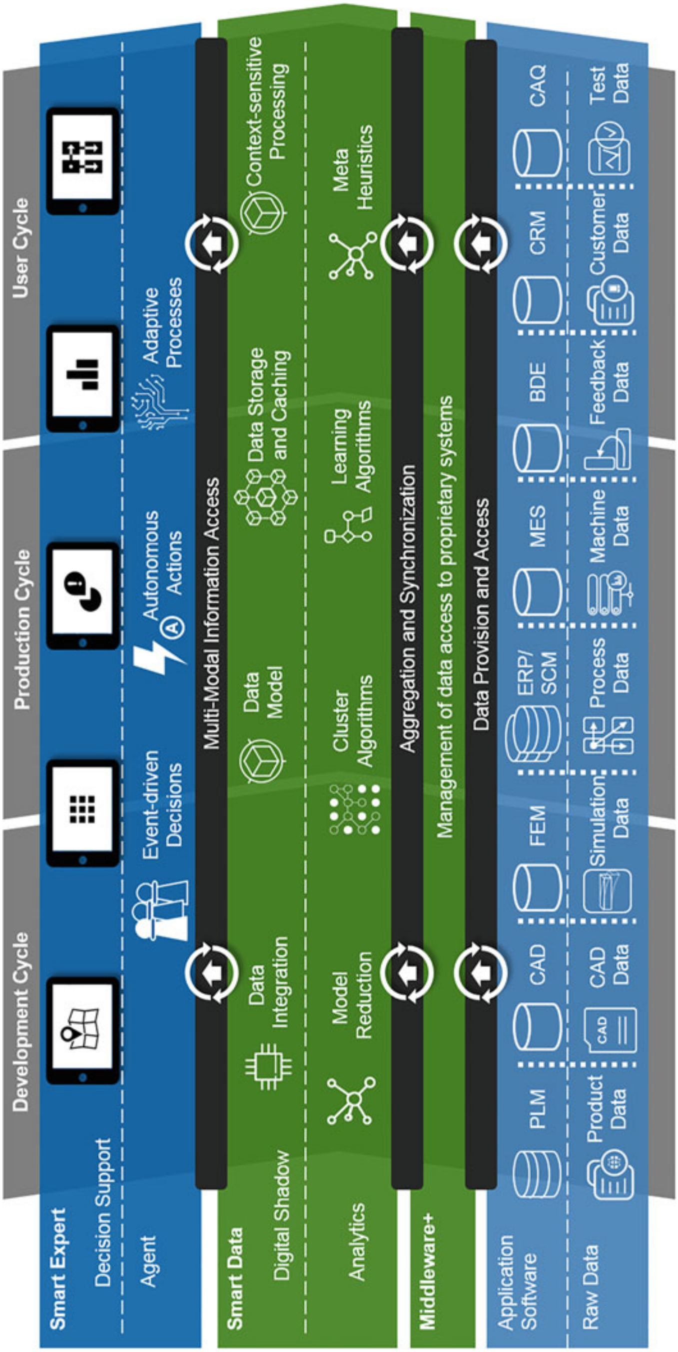


Fig. 2 Infrastructure for digital shadows

algorithms automatically search for optimized programs that lead to meta-heuristics, making Smart Data a necessity for successful Industrie 4.0-techniques.

The most important requirements are:

i. Metadata structures for **Smart Data** that are freely addressable, generic, expandable, and flexible to meet the increasing demands of complexity in the different areas.
ii. Cross-domain combinations of data and models as Digital Shadows with task-specific granularities.
iii. Analysis algorithms that use the theoretical foundations of production technology as prior knowledge for abstraction and structuring ("best of both worlds").

Most manufacturing companies operate in a similar way, using **various application software systems** to generate and manage the amounts of raw data produced by the machines as the basis of Smart Data. This raw data is mainly managed in Big Data silos. Therefore, data access, cleansing, aggregation, filtering, contextualization, and synchronization are linked to a lot of manual effort. Because of this approach, typical problems in cross-domain questions arise. A large amount of data combined with a lack of context information, low information density, restricted accessibility and the extensive domain knowledge required to understand the data are such of these problems. The use and processing of essential information in the raw data (CAD, FEM, ERP, MES, etc.) can often only be carried out by specialists who have access to the original systems.

A potential solution lies within a semantic interoperability layer, which helps gaining access to the raw data. It is a crucial challenge to manage the large amounts of data from decentralized sources and to make interoperation possible. Interoperation can be enabled through model mapping between proprietary application systems. In this context, methods for operation and cross-domain accessibility are needed, while naturally also focusing on privacy and data protection.

5 Use of the Digital Shadow for Production Technology

The optimized infrastructure leads to a significant increase in collaboration productivity in the domains. Similar to agile methods of software development (e.g., SCRUM), the interaction as a central part of Digital Shadow effectiveness enables continuous and integrated further development of specifications and benefits the processing of domain-specific tasks and cross-domain validation results. Three scenarios from production will serve as examples of this approach:

6 Scenario 1: Mastering Manufacturing Complexity

Production is generally characterized by a broad spectrum of highly specialized technologies. As reflected on before, in a complex manufacturing context, however, some details of the **interrelationships** are often not understood in their entirety. This is especially prominent when **long process chains** are involved. In addition, in many facets of the industry a large contingent of components is involved in production: A typical civil aircraft, for example, consists of more than three million parts that are exposed to high thermomechanical loads during operation. Identifying manufacturing errors is a great challenge as the entire spectrum of parts ranges from lightweight structural components for airframes (made from milled aluminum and titanium alloys) all the way to fuselage shells that combine molded sheets and carbon fiber composites. Each of those parts has an individual process chain history in which it is subjected to several manufacturing technologies. Stochastic influencing factors from material generation also have their specific impacts on the assembled system and must be observed.

To date, the prediction and control of the production results, ranging from intermediate to end product properties is hardly possible due to complex physical interactions. Many other factors such as energy consumption must be considered for a precise result.

While there are specialized physical models for individual aspects (Mayr et al. 2012), the totality of all models required to represent the complexity of the real world just cannot be calculated precisely while maintaining economic conditions. For example, quality measurements during milling are often limited to geometrical shape and surface tolerances. In reality, the quality of the parts depends on many influencing factors such as the thermal, static, and dynamic behavior of the manufacturing system as well as the tool wear status. A reasonably accurate prediction becomes even more difficult when residual stresses lead to additional workpiece deformation, which is extremely hard to visualize beforehand. A strong link to parameters of the previous manufacturing process chain is required, as the parameter space grows exponentially.

Machines that are used in manufacturing process chains already generate data from the NC (Numerical Control) and the PLC (programmable logic controller) or from additional sensors. Here, data-driven identification approaches have proven to be reliable at certain, but constant operating points. However, when it comes to identifying stochastic production results the information required is often not directly included. Adding sensors that would theoretically capture every influencing factor leads to very large amounts of data, which not only makes the overall problem unpredictable, but may also lead to the identification of false correlations. To make the existing batches of data analyzable, the history and context of the individual products must be present. In addition, hypotheses and domain knowledge are required so false alarms can be recognized, and causalities can be reinforced (Bayes Theorem). In the simplest case, this could be the extraction of characteristic values from the time signal (e.g., moving average), which are compared with a tolerance band.One of the major challenges will be to refine machine learning methods while also taking engineering knowledge into account. Specifically, the interaction between physical models across manufacturing process chains must be represented by the ideal combination. The real-time efficiency of these combined models requires a significant model reduction, e.g., by choice of granularity or numerical reduction methods such as Dual Craig-Bampton (Rixen 2004).

7 Scenario 2: Process-Accompanying Real-Time Quality Control in Production

Optimally, one would be able to predict the component quality during the manufacturing process already, seeing that structural components of an aircraft are specified by shape tolerances of a few micrometers. The relevant context information (e.g., CAD model of the workpiece, shape tolerances) is extracted and combined with models of reduced order (kinematics, statics, and dynamics of the machine tool) so the **Digital Shadow** is a real-time machining simulation validated by streamed NC signals (e.g., actual positions, accelerations, torque-generating drive currents).

In the first step, a machining simulation carried out on a graphics processing unit (GPU) and fast calculation models where tolerance violations are shown in an intuitive 3D view. The simulation makes the current process forces and the resulting deformations predictable. Accordingly, the machine operator can manually decide to adjust manipulated variables, initiate post-processing measures, or declare the part as a reject part.

As a next step, the Digital Shadow makes scalability across machines, production technologies and entire process chains possible. Relevant streaming data, context information and simulation models are recorded, reduced, and compiled automatically. In addition, algorithms for cause-effect relationship modeling create an assignment of main influencing factors. Similarly, decision-making is supported on a role-specific basis and ranges from machine operators and maintenance personnel to quality and plant managers, so people become users and trainers of the system at the same time. The simple Digital Shadow, a simulated real workpiece geometry (CAD format), can then be expanded with process information in every step. By transferring to subsequent processing steps such as deburring or grinding, optimized process parameters can be better predicted. Importantly, knowledge from previous steps of the lengthy process is transferred, e.g., internal stresses through hot rolling or freeform forging, enabling a highly adaptive production process.

8 Scenario 3: New Business Models in Production

In general, manufacturing companies are not only interested in purchasing machinery, but aim for production capacity. However, they rarely exploit the full potential of their facilities from new types of production equipment. Another example:

Due to the increasing demand for its products, a manufacturer wants to build up more production capacity. There is a high diversity of products in his portfolio, so the manufacturer provides the machine supplier with a list of requirements derived from the most demanding products. Based on the requirements defined in this way, the provider cannot assess the real need situation and proposes a standard specification. After the start-up and usage phase, a sample analysis shows that the new production facilities have an overall equipment effectiveness (OEE) that is far below the required values. This happened because non-optimal configurations and parameters of the standard specification have been set. In most situations, the provider neither has access to the manufacturer's production data to determine optimal operating points, nor does the manufacturer recognize suboptimal operating processes.

The Overall Equipment Effectiveness (OEE) is 100% if only good parts are produced as quickly as possible and without downtimes.

Usage data should be useable across companies or even departments to optimize operational processes, investment decisions or innovation processes. However, in most cases this is not yet established. Learning and analyzing could be much faster and more efficient if manufacturers could not only use their own data, but also access data from similar contexts in other industries. Therefore, incentives, process control and new methods of user integration are required (Kortmann and Piller 2016). At the same time, the ability to implement the integration of Industrie 4.0 is a question of the right incentives to harmonize the different interests and priorities of the partners involved. Given the recent developments in platform-based business models the next step will be to create incentives for sharing production knowledge and data.

In a networked production world, the manufacturer could use technological knowledge and usage data from other manufacturers to gain better knowledge of its own requirements. OEE could be improved, for example, by learning from best practices from similar processes and building experience. The machine manufacturer could in turn use the knowledge to provide improved systems. For example, the data could be used by the manufacturer to provide a customized initial configuration that allows the new machines to be used directly with a much higher OEE. By accessing extensive usage data, which not only covers machine data, but also data on the production context, material properties and the behavior of the operator, the user could continuously receive input on how the OEE can be improved. "Analytics apps" used during the usage cycle provide data of both the products (outputs) that were generated on the connected machines and their users (customers). That way, continuous feedback of this data for the development cycle of the machine manufacturer is made accessible. Finally, the OEE can be increased even further and requirements for the development of hardware and software of the next generation become known.

References

Allwood JM, Duncan SR, Cao J et al (2016) Closed-loop control of product properties in metal forming. CIRP Ann-Manuf Technol 65(2):573–596

Brecher C, Özdemir D (eds) (2017) Integrative production technology: theory and applications. Springer, Cham

Gandomi A, Haider M (2015) Beyond the hype: big data concepts, methods, and analytics. Int J Inf Manage 35(2):137–144

Gao R, Wang L, Teti R et al (2015) Cloud-enabled prognosis for manufacturing. CIRP Ann Manuf Technol 64(2):749–772

Gubbi J, Buyya R, Marusic S et al (2013) Internet of things (IoT): a vision, architectural elements, and future directions. Future Generation Comput Syst 29(7):1645–1660

Kortmann S, Piller F (2016) Open business models and closed-loop value chains: redefining the firm-consumer relationship. Calif Manage Rev 58(3):88–108 18

Mayr J, Jedrzejewski J, Uhlmann E et al (2012) Thermal issues in machine tools. CIRP Ann Manuf Technol 61(2):771–791

Porter ME, Heppelmann JE (2014) How smart, connected products are transforming competition. Harv Bus Rev 92(11):64–88

Rixen D (2004) A dual Craig-Bampton method for dynamic substructuring. J Comput Appl Math 168(1–2):383–391

Mechanics 4.0

Artificial Intelligence for the Analysis of Mechanical Systems

Arnd Koeppe, Daniel Hesser, Marion Mundt, Franz Bamer, Michael Selzer, and Bernd Markert

Contents

A. Koeppe (✉) · M. Selzer
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
e-mail: arnd.koeppe@kit.edu; michael.selzer@kit.edu

D. Hesser · F. Bamer · B. Markert
Chair and Institute of General Mechanics, RWTH Aachen University, Aachen, Germany
e-mail: hesser@iam.rwth-aachen.de; bamer@iam.rwth-aachen.de; markert@iam.rwth-aachen.de

M. Mundt
Minderoo Tech and Policy Lab, UWA Law School, The University of Western Australia, Perth, WA, Australia
e-mail: marion.mundt@uwa.edu.au

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_23

1 Introduction

The digitization of industry in the context of “Industry 4.0” comprises four pillars: networking, information transparency, decentralized decision-making, and technical assistance (Hermann et al. 2016). In particular, the ability to make decisions decentrally—based on relevant information—and the provision of informed digital assistance systems require methods for analyzing physical systems. The **analysis of physical systems** in industry is a classical application area of mechanics and requires accurate measurements, descriptions, and interpretations of the states of mechanical systems. A recent survey underscores the need for necessary engineering skills, particularly technical mechanics when working on Industry 4.0 projects (Heidling et al. 2019).

Current mechanics are characterized by a combination of time-consuming, highly specialized experiments, well-founded theoretical approaches, and deterministic simulation methods, such as the **Finite Element Method (FEM)** (see, e.g., Bathe 1996). The FEM has the advantage that arbitrary structures can be composed of smaller elements, but it is computationally demanding when nonlinear material behavior and strong discontinuities are present. Thus, simulations of manufacturing processes become exceptionally costly, such as milling simulations that consider the wear of the milling head.

However, in the context of digitized production chains, fast and nevertheless accurate statements about such highly nonlinear relationships are necessary so that a specially designed experiment or a complete FE simulation of the entire previous production chain cannot be performed every time for each production piece. The large amount of time and expert knowledge required for classical mechanics methods must be adapted and modernized to keep pace with a digital Industry 4.0. Precisely fitting and fast analysis methods for mechanical systems in Industry 4.0, e.g., analysis methods based on **Artificial Intelligence (AI),** are thus the subject of current research to meet the increasing demand in connection with **cyber-physical systems** and **digital twins** in production technology.

The objective of such a fit **Mechanics 4.0 is** thus the training and effective use of **AI** methods to describe and interpret the behavior of mechanical systems. For this purpose, data from different sources, such as experiments and simulations, are combined and generalized via learning algorithms. Based on this generalized experience, **AI** systems can make fast and accurate predictions about the behavior of the underlying mechanical system.

In recent years, **AI** methods, especially **(artificial) neural networks, have** revolutionized diverse application fields ranging from image processing to speech recognition and autonomous driving (LeCun et al. 2015). This incredible leap in performance to learn even highly nonlinear and complex relationships has been enabled by increased (parallel) computing capacities, **deep neural network architectures**, efficient learning algorithms, and the availability of **large datasets (Big Data)** (Goodfellow et al. 2016). Notable applications include the neural network-based AIs Alpha Go and Alpha Go Zero, which achieve superhuman performance in the traditional Chinese board game Go (more possible positions than atoms in the observed universe) (Silver et al. 2016, 2017). In less than two days of training, Alpha Go Zero managed to surpass more than a thousand years of human experience (Silver et al. 2017).

In mechanics, **AI** is used as an active computational method primarily in three areas: **Acceleration of simulations**, **structural health monitoring,** and **dynamic behavior prediction**.

Simulations can be accelerated using intelligent material models, elements, or surrogate models. Intelligent **material models** (Ghaboussi and Sidarta 1998; Javadi et al. 2009; Oeser and Freitag 2009; Stoffel et al. 2018) in **FEM** used the neural network capabilities to represent arbitrary functions but only reduced the required computation time of simulations to a limited extent. **Intelligent surrogate models** (Graf et al. 2010; Freitag et al. 2011; Cao et al. 2016; Koeppe et al. 2016; Bamer et al. 2017; Freitag et al. 2017; Koeppe et al. 2017, 2018b, 2019) provided significant speed benefits, but are limited to one model at a time, which must be trained individually. **Intelligent meta elements** (Koeppe et al. 2018a, 2020) compensate for this disadvantage by training parts of the structure over general load cases and using them as elements in the **FEM**.

Structural health and condition monitoring offer the advantage that the assessment of structures and processes can occur during operation (Hesser and Markert 2019). As a result, maintenance intervals can be adjusted accordingly, and spare parts can be ordered in advance in the event of diagnosed damage. This form of condition-based maintenance ensures an efficient and reliable operation. System downtime can thus be kept to a minimum, service costs can be saved, and maintenance schedule can be optimized through predictive maintenance systems.

In biomechanics, e.g., in motion analysis applications (Mundt et al. 2019, 2020a, b, c), neural networks offer both a speed advantage and the possibility to **predict parameters** that cannot be measured directly or calculated analytically. To date, the gold standard in motion analysis has been optical measurement techniques for capturing the **kinematics of** motion in combination with force plates for capturing ground reaction forces. **Inverse-dynamic calculations** can be used to determine joint kinetics. However, since this approach also offers several disadvantages, inertial sensors have become increasingly popular. However, these can only measure kinematic parameters. With the aid of neural networks, however, the **kinetics** can also be predicted.

In the following, some methods and application examples of **AI** learning behavior of mechanical systems are explained. Further examples from biomechanics and motion analysis are described in chapter Life Sciences 4.0.

2 Methods

2.1 Data of Mechanical Systems

In mechanical systems, data can be classified according to two criteria: Space dependence and time dependence. Spatially and temporally independent data exhibit constant values that do not change for the whole system at any time considered. A classic example from mechanics is Young's modulus (the origin tangent of the material stiffness) in a system consisting of a homogeneous material at a constant temperature. **Space-dependent data, in** addition to their value, have a topological relationship, i.e., a relative position to each other (see Fig. 1). The collection of space-dependent data, such as displacements of different structure points, are considerably more complicated since many sampling points (e.g., via optical systems) have to be defined. **Time-dependent data** change their value during the period under consideration. Since the data can only be collected at fixed points in time, the sampling rate determines the effects captured and the amount of data generated.

In general, data are both time and space-dependent. Based on mechanics expert knowledge, idealizations can be introduced for the considered space and time intervals, which simplify the problem. In the following, some methods for typical applications in mechanics and their data are described.

Simulation methods in structural mechanics are used to calculate the mechanical behavior of structures. In general, the mechanical behavior is defined by partial

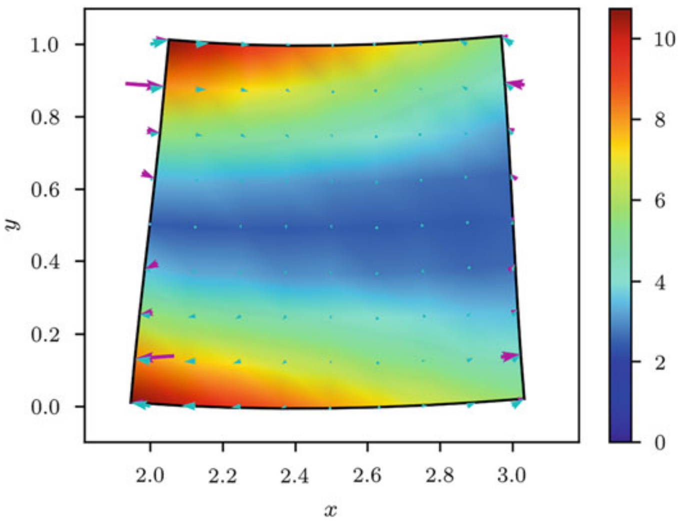


Fig. 1 A deformed square 2D model under mixed bending and shear loading with linear elastic material behavior. The external loads (magenta) are balanced by internal stresses (color gradient of Von Mises equivalent stresses), which cause deformation of the structure (cyan). All the above quantities are space-dependent. In the case of inelastic material behavior, stresses and deformations are also time-dependent

differential equations. To solve these equations, the structure is discretized in the **FEM** with small elements, over which the courses of mechanical field variables (usually dependent on space and time) are assumed. Material models with numerical integration over the elements define the interrelationships of the field quantities. The equilibrium conditions are then solved on the discretized structure, usually using **iterative solution methods** such as the Newton-Raphson algorithm (see Fig. 1). The **FEM** has been widely used in industry since its invention in the 1960s. Despite considerable progress in also calculating **complex material behavior** (such as inelasticity, fracture processes) and **heterogeneous multiphase or nanomaterials** (such as fluid-saturated, porous media or hierarchically structured nanomaterials) (see, e.g., Ehlers and Markert 2001; Markert 2005, 2007; Ehlers et al. 2009; Xiao et al. 2009), these calculation methods have not yet been established in the industry. Here, the Pareto principle often rules, i.e., 80% of the solution in 20% of the time is preferred, which corresponds to fast solutions at the expense of simulation quality.

Experiments are time-consuming, expensive, require expert knowledge, and are also often destructive. For example, it is not efficient to test the safety of every single car produced in a crash test after it has been manufactured, although this would allow the most accurate statement to be made about how safe the car used to be. Thus, experiments are often used in combination with simulations by determining material parameters, validating simulation models, and verifying simulation results before the product goes into production. Efficient and **non-destructive experiments** with inexpensive sensors, e.g., with piezo elements and inertial measurement systems, allow applications ranging from **structural health monitoring** to **gait analysis under real conditions** (see chapter Life Sciences 4.0).

Piezoelectric transducers can be used for sensory and actuator applications due to their interaction between mechanical deformation and electrical charge change. Here, the direct piezoelectric effect describes a potential voltage generation when a mechanical load is applied. In turn, mechanical deformations can be caused by the application of an electrical voltage, which is called the inverse piezoelectric effect (Curie and Curie 1880, 1881). This effect is used in non-destructive testing to generate elastic waves in solids and measure their wave response. The individual wave packets provide information about the structural condition so that damage in a structure can be detected and localized (Hesser and Markert 2018).

Inertial sensor units usually consist of three sensors: an accelerometer to determine the linear acceleration in all three spatial directions, a gyroscope to determine the rotation rate in all three spatial planes, and a magnetometer to determine the magnetic flux density in all three directions. Based on this information, a sensor's orientation can be calculated by using a Kalman filter (Sabatini 2006). However, this approach has a drawback: the magnetometer data are perturbed by external, inhomogeneous magnetic fields so that the calculated orientations are erroneous (de Vries et al. 2009). Therefore, it is often useful to use only the data from the linear accelerometers and the gyroscopes.

This section shows that collecting, generating, and interpreting data are tasks that requires expertise in a particular area of mechanics. For example, any learning algorithm can only map the probabilities that are present in the data. For example, the impact of data on the outcome can only be learned if there is enough variance in the training data. The challenge is to ensure that the training data accurately represent the mechanical system's real behavior and that the largest possible amount of data are available.

2.2 Data Preparation and Data Management

Data preparation determines whether the **AI** algorithm is even able to learn effectively from the training data and generalize, i.e., apply what it has learned to new input data not used during training (cf. Bishop 1995; Goodfellow et al. 2016).

For this purpose, all available data are first divided into three **datasets.** Most of the data are used in the training dataset to train the **AI** algorithm. In addition, a smaller test dataset is used to show how the network generalizes after the learning phase. The third dataset, the validation or development dataset, is used during learning but not to train the **AI**. By comparing the errors on the training and validation datasets, we check for **overfitting** to the training data and thus poor **generalization**. This comparison explicitly does not involve the test dataset because treating overfitting adjusts parameters of the **AI** algorithm, which in turn can cause overfitting to the dataset compared to the training set (cf. Bishop 1995; Goodfellow et al. 2016).

Following the **dataset split,** the training data are usually standardized (centered, normalized, and rotated if necessary) and transformed (e.g., logarithmically compressed) to remove influences of absolute values and obtain a standardized distribution of the training data. Details about this so-called **feature scaling** and other common methods and theoretical background can be found, e.g., in Bishop (1995) or Goodfellow et al. (2016).

The best methods for normalizing and unifying a dataset depend heavily on the data, the task, and the **AI** algorithm being used. For example, it is often useful to obtain the order and relative magnitude of temporally or spatially dependent variables when learning these relationships.

To handle the large amounts of data during generation and processing in an automated workflow, a **data management infrastructure** is necessary (cf. Selzer et al. 2020). Besides the raw data, workflows to create AIs need well-structured meta data, which—in the best case—contain the relations between different kinds of information stored in the datasets. Using the meta data, the raw data can be converted, structured, and translated to specific requirements of the AI algorithm. In Sandfeld et al. (2018), this process is referred to as **digital transformation**. To transform datasets into useful structures, it is recommended to start collecting all possibly required meta data as early as possible in the data generation process. Thus, digital data transformation helps to increase the generalization capabilities of trained AI and data-driven algorithms and results in better predictions of mechanical system behavior by the **AI**.

Subsequently formalized into a bottom-up **data ontology**, the semantic rules that describe, connect, and represents the processes of specific research domains, such as mechanics, become apparent and eligible to analysis. By comparing the results of the **AI** workflows with conventional workflows, the data ontologies valid for the individual datasets and the entire research domain can evolve, intertwine, and unify with more abstract top-down **data ontologies**. In turn, this incremental evolution of **data ontologies** improves the prediction of **AI** algorithms to better accord with reality.

2.3 Artificial Intelligence for Mechanical Systems

The tasks for **AI** algorithms for mechanical systems cover all three typical application areas of **AI**: **Regression of** expected values (e.g., predicting stresses in a structure), **classification of** system states (e.g., telling whether a structure is damaged), and, less commonly, **clustering** (e.g., discovering dependencies between state variables). The **AI** methods used for this are interchangeable and have individual strengths and weaknesses depending on the data and the task.

The following subsections provide an overview of two exemplary learning algorithms for mechanics. For the general case, Goodfellow et al. (2016) provide a comprehensive compilation of current theories, methods, and algorithms.

2.3.1 Support Vector Machines

The **Support Vector Machine (SVM)** represents a nonparametric machine learning method that solves **classification and regression problems** (Russell et al. 2010). It is mainly applied in areas where the mathematical and physical descriptions of the dependencies are highly complex, e.g., in using elastic wave propagation in solids for damage localization (Hesser and Markert 2018). The core of any SVM is the determination of a boundary between the different classes given by the training data. The partitioning is done using the **maximum margin classifier** so that the class boundary is the farthest possible from the test data, maximizing generalizability (Bishop 1995; Russell et al. 2010). Figure 2 shows a two-dimensional problem that is separated using a linear decision boundary. The points with the minimum distance to the class boundary are also called **Support Vectors** and are marked with an extra frame in Fig. 2. The support vectors completely define the boundary's optimal location so that the other data points do not need to be considered.

If the data cannot be linearly separated, the **kernel trick** can map the input data into a higher dimension space. This extension allows some classes to be linearly separated in high-dimensional space, resulting in a nonlinear decision boundary in the original dimension (Russell et al. 2010; Schölkopf et al. 2002). In both cases, linear and nonlinear, finding the SVM solution represents a **convex quadratic optimization problem**, where each local minimum is a global minimum. Thus,

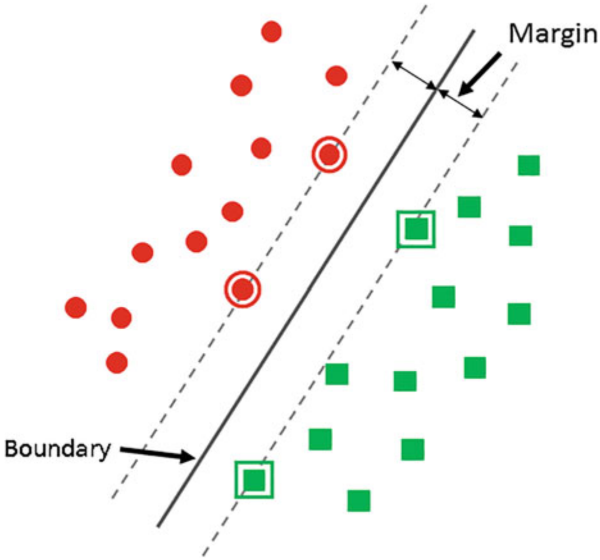


Fig. 2 Illustration of the linear decision boundary for a two-dimensional problem

SVMs are easy to train but limited in their generalizability compared to artificial neural networks (Bishop 1995; Schölkopf et al. 2002; Goodfellow et al. 2016).

2.3.2 Artificial Neural Networks

Artificial neural networks consist of layers of artificial neurons that weight and sum their input values and compute an output value via a nonlinear activation function (Fig. 3). Having multiple layers and multiple neurons per layer increases the capacity of the neural network to map **arbitrary nonlinear functions** (Hornik et al. 1989). Neural networks can thus **solve regression and classification problems** that remain nonlinear even in higher-dimensional spaces. The weights of individual neurons are adjusted during training to reduce an error function. Training these weights presents a **non-convex optimization problem** and requires a large amount of data. The major advantage of neural networks trained with **large amounts of data (Big Data)** is a better generalization and lower error on new data (cf. Goodfellow et al. 2016).

The simplest neural network architecture—**dense feedforward neural networks**—as shown in Fig. 3 requires many variables to store all the weights. Thus, they require lengthy training, use vast amounts of memory, and are difficult to interpret (cf. Goodfellow et al. 2016). **Recurrent neural networks** (Gers et al. 1999; Hochreiter and Schmidhuber 1997; Williams 1992) and **Convolutional neural networks** (LeCun et al. 1998) use temporal or spatial dependencies of the data to share weights between different locations and time points. They leverage the spatial position and time dependencies within the data to create more effective architectures that can be trained efficiently. For example, Freitag et al. (2017) and Koeppe et al.

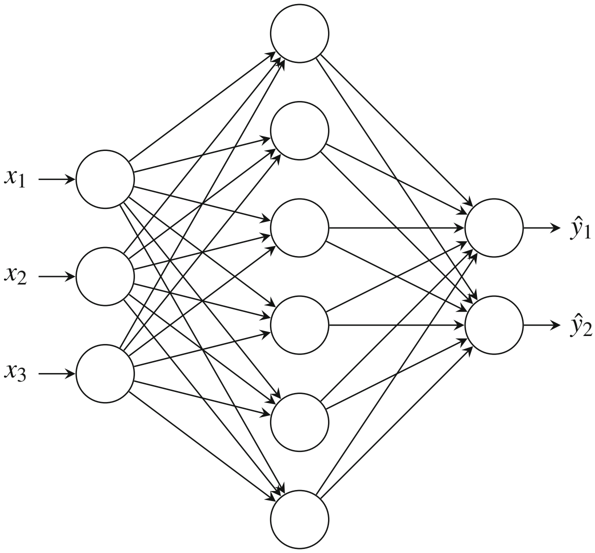


Fig. 3 An artificial neural network consisting of four layers with dense connections between the layers

(2017) describe examples of applications of such efficient neural network architectures in mechanics.

A high evaluation speed counterbalances the initial extra effort to collect data and train the neural networks towards even high-dimensional, nonlinear problems, benefiting, e.g., the following use cases.

3 Application Examples

3.1 Real-Time Monitoring of Milling Heads

The first application example illustrates the use of **AI** in the **condition monitoring of** milling tools of a CNC machine (Hesser and Markert 2019). The focus here is on the mechanical interaction between the tool's cutting force and the induced vibration in the machine structure. Through suitable sensor selection and positioning, even old CNC machines that have not been designed with digital interfaces may be outfitted toward digitization since only a fraction of today's production landscape employs only industry-4.0-capable infrastructure (Bosch Media Service 2019). This **retrofitting** offers companies the opportunity to achieve production networking by upgrading existing machines and systems and eliminates the need for costly and time-consuming new construction of production units and their infrastructure (Bosch Media Service 2019).

In this example, the milling tool of a Deckel Maho DMU 35M CNC machine is monitored during the work process, and increasing wear of the cutting surfaces is detected (Hesser and Markert 2019). The CNC machine cannot monitor process

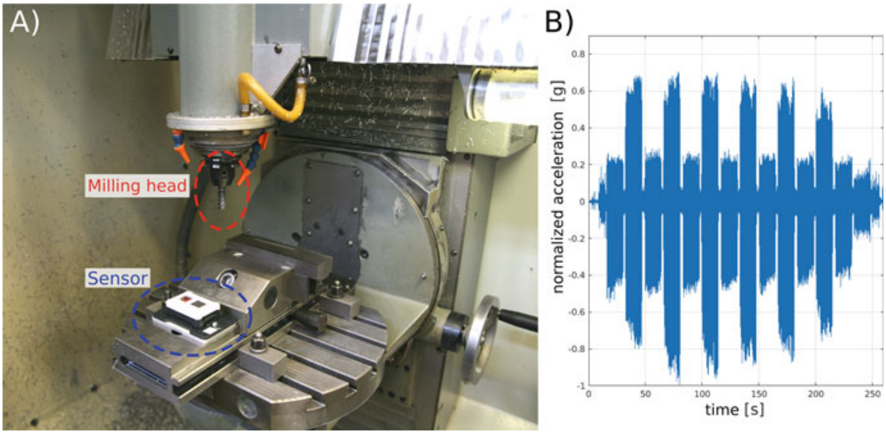


Fig. 4 CNC machine with sensor (**a**) and raw data of a measurement series (**b**)

variables such as spindle speed or feed rate. For this reason, the Bosch XDK development platform is integrated into the machine (see Fig. 4(a)). The platform offers the possibility to acquire different physical measurands via built-in sensors. In this case, acceleration data are collected on the machine's workbench so that a correlation between the cutting force of the tool and the induced vibration in the machine can be derived. Optionally, the data can be transmitted via a WLAN or Bluetooth module of the development platform and is thus available for further analysis. Figure 4(b) represents the collected raw data for a work step. Here, a steel block was machined with a zigzag path, which makes the force effect during co-rotation and counter-rotation of the milling cutter apparent. By machining several steel blocks, a database is built up that makes it possible to reflect the tool's wear characteristics. Before the data can be processed, normalization, and filtering of the raw data must take place. In the following step, the complexity of the data series, which consists of several hundreds of time points, is reduced by introducing statistical features. Using statistical features offers the advantage that the training must deal only with reduced input dimensionality, which dramatically reduces the training time.

Consequently, a neural network can classify the tool's health state, as shown in Fig. 5. For this purpose, the test data was collected with a new tool and an already worn tool. In addition, the progressive wear of the tools can be detected when the measurement data are recorded. This information can be used in **predictive maintenance to estimate** the remaining serviceability of a tool, an advantage that can significantly improve the quality and reliability of products. The decisive factor in **AI**-supported predictions, e.g., by neural networks, remains the expert in production, who in this case must determine the quality limit for replacing the tool based on empirical values and standards. Based on this classification, the remaining service life of a tool can be predicted.

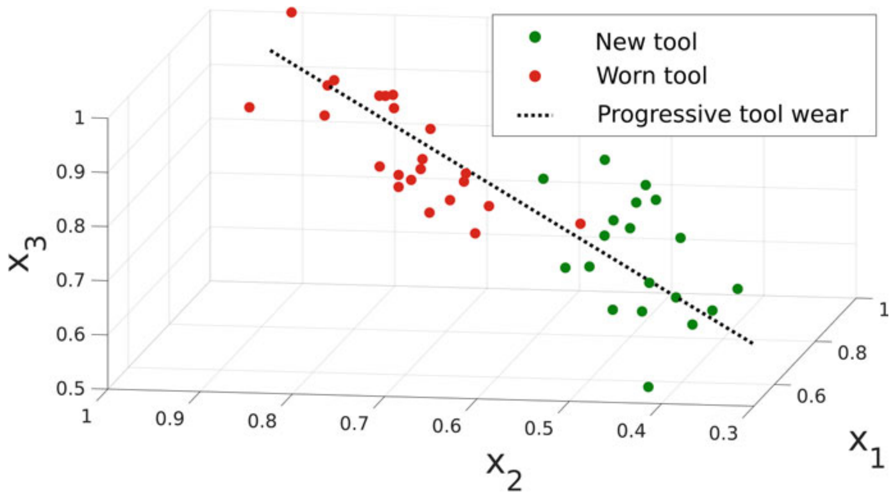


Fig. 5 Classification of new and old tools based on a neural network. The input data has been reduced by normalization and filtering to statistical features X1, X2, and X3, which allow a clear distinction between new and old tools. The wear of the tool is recognizable as a regression line (dashed)

3.2 Intelligent Surrogate Models and Intelligent Meta Elements for Structural Mechanics Simulations

In the second example, structural mechanics **finite element simulations** are accelerated using neural networks. The dimensionality of the problem, which corresponds to the number of degrees of freedom of the finite element model or the number of data points to be computed, affects the accuracy and speed of the computation. The straightforward approach of **intelligent material models** (Ghaboussi and Sidarta 1998; Javadi et al. 2009; Oeser and Freitag 2009) speeds up the computation per data point, but due to the small number of variables per data point, it is only useful for small neural network architectures. The ability of neural networks to efficiently process high-dimensional data remains unused. This lever is addressed by **intelligent surrogate models** (Graf et al. 2010; Freitag et al. 2011; Cao et al. 2016; Freitag et al. 2017; Koeppe et al. 2017, 2018b) and **intelligent meta elements** (Koeppe et al. 2018a, 2020). While intelligent surrogate models replace a single model with a specially trained neural network that speeds up computations, intelligent meta elements reduce the number of degrees of freedom within finite element models. Thus, they can predict the response to a wide variety of load cases.

In Fig. 6, an FE model of a cantilever beam is compared with a model with intelligent meta elements of the same geometry but with a reduced number of degrees of freedom. Both models assume linear-elastic material behavior (Young's modulus: 1e4 N/m2, lateral contraction: 0.25), plane stress states, and small deformations. The finite elements are P2 Lagrangian triangles. The **intelligent meta**

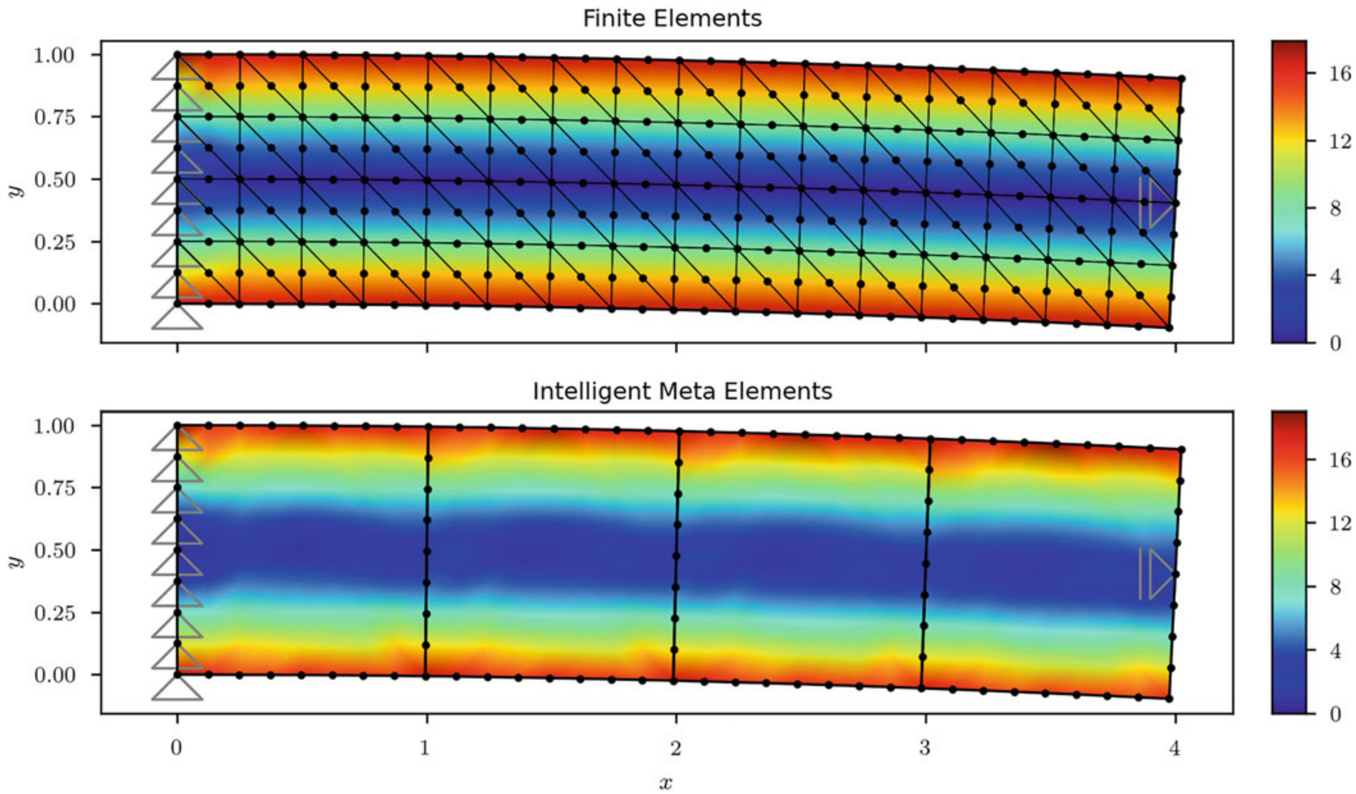


Fig. 6 A cantilever beam under a moment load at the right end with linear material behavior. The finite element model (top) has 594 degrees of freedom, while the intelligent meta element model (bottom) has only 202 degrees of freedom. The upscaled deformations and the curves of the von Mises equivalent stresses agree well

elements share a **Fully-connected Feedforward Neural Network** that predicts displacements, stresses, and internal forces of the individual elements from their boundary deformations. The neural network was trained with 35,000 examples from **general randomized load cases** (cf. Fig. 1), computed with the **FEM**, with 7500 equally generated examples each retained for validation and testing. The normalized error (**range-normalized root mean square error**) on the independent test dataset is 1.12%.

The use case shown in Fig. 6 is an example of using intelligent meta elements in the **FEM** for general load cases. Thereby, they compensate for the main disadvantage of the intelligent surrogate models, e.g., by Graf et al. (2010), Freitag et al. (2011), or Cao et al. (2016). By offloading the computations to the neural network, the number of degrees of freedom can be significantly reduced, which significantly speeds up the simulation.

4 Summary and Outlook

This chapter reviewed some methods and relevant applications in a **Mechanics 4.0** setting. Recommendations have been made for the selection, handling, and preparation of mechanical systems data and the selection of learning algorithms to be used. Both areas benefit from mechanical expertise to optimally prepare data and interpret the results of the learning algorithms. The collection of data and the **AI** methods' training mean an additional effort at first. However, they show a higher evaluation speed for high-dimensional, nonlinear mechanical problems in the long run.

Thus, **AI** methods are increasingly finding their way into the field of mechanics. The established methods, such as experiments or established simulation processes such as the **FEM**, are not replaced here but supplemented by the new methods' advantages. The accelerated, hybrid calculation methods harmonize with the other pillars of Industry 4.0, particularly **Big Data** and networking, whereby production and design processes can be accelerated through the efficient use of **cyber-physical systems** and **digital twins.**

Acknowledgments The authors gratefully acknowledge the financial support of the German Federal Ministry of Education and Research (BMBF) within the "FESTBATT" consortium (Grant No. 03XP0174E).

References

Bamer F, Koeppe A, Markert B (2017) An efficient Monte Carlo simulation strategy based on model order reduction and artificial neural networks. Proc Appl Math Mech 17:287–288. https://doi.org/10.1002/pamm.201710113

Bathe K-J (1996) Finite element procedures. Prentice Hall

Bishop CM (1995) Neural networks for pattern recognition. Clarendon Press
Bosch Media Service (2019) Old machine + rapid connectivity = new benefit [WWW Document], n.d. URL https://www.bosch-presse.de/pressportal/de/en/old-machine-rapid-connectivity-new-benefit-69632.html (accessed 1.30.19)
Cao B-T, Freitag S, Meschke G (2016) A hybrid RNN-GPOD surrogate model for real-time settlement predictions in mechanised tunnelling. Adv Model Simul Eng Sci 3:5. https://doi.org/10.1186/s40323-016-0057-9
Curie J, Curie P (1880) Development, via compression, of electric polarization in hemihedral crystals with inclined faces. Bulletin de la Societe de Minerologique de France 3:90–93
Curie J, Curie P (1881) Contractions and expansions produced by voltages in hemihedral crystals with inclined faces. Comptes Rendus 93:1137–1140
de Vries WHK, Veeger HEJ, Baten CTM, van der Helm FCT (2009) Magnetic distortion in motion labs, implications for validating inertial magnetic sensors. Gait Posture 29:535–541. https://doi.org/10.1016/j.gaitpost.2008.12.004
Ehlers W, Markert B (2001) A linear viscoelastic biphasic model for soft tissues based on the theory of porous media. J Biomech Eng 123:418–424
Ehlers W, Karajan N, Markert B (2009) An extended biphasic model for charged hydrated tissues with application to the intervertebral disc. Biomech Model Mechanobiol 8:233–251
Freitag S, Graf W, Kaliske M, Sickert J-U (2011) Prediction of time-dependent structural behaviour with recurrent neural networks for fuzzy data. Comput Struct Civil-Comp 89:1971–1981. https://doi.org/10.1016/j.compstruc.2011.05.013
Freitag S, Cao BT, Ninić J, Meschke G (2017) Recurrent neural networks and proper orthogonal decomposition with interval data for real-time predictions of mechanised tunnelling processes. Comput Struct. https://doi.org/10.1016/j.compstruc.2017.03.020
Gers FA, Schmidhuber J, Cummins F (1999) Learning to forget: continual prediction with LSTM 850–855. https://doi.org/10.1049/cp:19991218
Ghaboussi J, Sidarta DE (1998) New nested adaptive neural networks (NANN) for constitutive modeling. Comput Geotech 22:29–52. https://doi.org/10.1016/S0266-352X(97)00034-7
Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press
Graf W, Freitag S, Kaliske M, Sickert J-U (2010) Recurrent neural networks for uncertain time-dependent structural behavior. Comput-Aided Civ Infrastruct Eng 25:322–323. https://doi.org/10.1111/j.1467-8667.2009.00645.x
Heidling E, Meil P, Neumer J, Porschen-Hueck S, Schmierl K, Sopp P, Wagner A (2019) Engineers for Industry 4.0, Impuls. IMPULS Foundation for Mechanical Engineering, Plant Engineering and Information Technology, Munich
Hermann M, Pentek T, Otto B (2016) Design Principles for Industrie 4.0 Scenarios. In: 2016 49th Hawaii International Conference on System Sciences (HICSS). Presented at the 2016 49th Hawaii International Conference on System Sciences (HICSS), pp 3928–3937. https://doi.org/10.1109/HICSS.2016.488
Hesser DF, Markert B (2018) Guided wave-based damage detection in solids using computational intelligence. PAMM 18:e201800296. https://doi.org/10.1002/pamm.201800296
Hesser DF, Markert B (2019) Tool wear monitoring of a retrofitted CNC milling machine using artificial neural networks. Manuf Lett 19:1–4. https://doi.org/10.1016/j.mfglet.2018.11.001
Hochreiter S, Schmidhuber J (1997) Long short term memory. Neural Comput 9:1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. Neural Netw 2:359–366
Javadi AA, Tan TP, Elkassas ASI (2009) Intelligent finite element method and application to simulation of behavior of soils under cyclic loading. In: Abraham A, Hassanien A-E, Snášel V (eds) Foundations of computational intelligence volume 5, studies in computational intelligence. Springer, Berlin, Heidelberg, pp 317–338
Koeppe A, Bamer F, Markert B (2016) Model reduction and submodelling using neural networks. Proc Appl Math Mech 16:537–538. https://doi.org/10.1002/pamm.201610257

Koeppe A, Bamer F, Hernandez Padilla CA, Markert B (2017) Neural network representation of a phase-field model for brittle fracture. Proc Appl Math Mech 17:253–254. https://doi.org/10.1002/pamm.201710096

Koeppe A, Bamer F, Markert B (2018a) An intelligent meta-element for linear elastic continua. PAMM 18:e201800283. https://doi.org/10.1002/pamm.201800283

Koeppe A, Hernandez Padilla CA, Voshage M, Schleifenbaum JH, Markert B (2018b) Efficient numerical modeling of 3D-printed lattice-cell structures using neural networks. Manuf Lett. https://doi.org/10.1016/j.mfglet.2018.01.002

Koeppe A, Bamer F, Markert B (2019) An efficient Monte Carlo strategy for elasto-plastic structures based on recurrent neural networks. Acta Mech 230:3279–3293. https://doi.org/10.1007/s00707-019-02436-5

Koeppe A, Bamer F, Markert B (2020) An intelligent nonlinear meta element for elastoplastic continua: deep learning using a new time-distributed residual U-net architecture. Comput Methods Appl Mech Eng 366:113088. https://doi.org/10.1016/j.cma.2020.113088

LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86:2278–2324. https://doi.org/10.1109/5.726791

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444. https://doi.org/10.1038/nature14539

Markert B (2005) Porous media viscoelasticity with application to polymeric foams, Report / Universität Stuttgart, Institut für Mechanik (Bauwesen), Lehrstuhl II. Verl. Glückauf, Essen

Markert B (2007) A constitutive approach to 3-d nonlinear fluid flow through finite deformable porous continua. Transp Porous Media 70:427–450

Mundt M, David S, Koeppe A, Bamer F, Markert B, Potthast W (2019) Intelligent prediction of kinetic parameters during cutting manoeuvres. Med Biol Eng Comput 57:1833–1841. https://doi.org/10.1007/s11517-019-02000-2

Mundt M, Koeppe A, Bamer F, David S, Markert B (2020a) Artificial neural networks in motion analysis—applications of unsupervised and heuristic feature selection techniques. Sensors 20:4581. https://doi.org/10.3390/s20164581

Mundt M, Koeppe A, David S, Witter T, Bamer F, Potthast W, Markert B (2020b) Estimation of gait mechanics based on simulated and measured IMU data using an artificial neural network. Front Bioeng Biotechnol 8. https://doi.org/10.3389/fbioe.2020.00041

Mundt M, Thomsen W, Witter T, Koeppe A, David S, Bamer F, Potthast W, Markert B (2020c) Prediction of lower limb joint angles and moments during gait using artificial neural networks. Med Biol Eng Comput 58:211–225. https://doi.org/10.1007/s11517-019-02061-3

Oeser M, Freitag S (2009) Modeling of materials with fading memory using neural networks. Int J Numer Meth Eng 78:843–862. https://doi.org/10.1002/nme.2518

Russell SJ, Norvig P, Davis E (2010) Artificial intelligence: a modern approach, 3rd edn. Prentice Hall series in artificial intelligence. Prentice Hall, Upper Saddle River

Sabatini AM (2006) Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing. IEEE Trans Biomed Eng 53:1346–1356. https://doi.org/10.1109/TBME.2006.875664

Sandfeld S, Dahmen T, Fischer F, Eberl C, Klein S, Selzer M, Nestler B, Möller J, Mücklich F, Engstler M, Diebels S, Tschuncky R, Prakash A, Steinberger D, Kübel C, Herrmann H-G, Schubotz R (2018) Digitale Transformation in der Materialwissenschaft und Werkstofftechnik (Electronic Resource). Deutsche Gesellschaft für Materialkunde e.V., Frankfurt

Schölkopf B, Smola AJ, Scholkopf MD of the M.P.I. for B.C. in T.G.P.B., Bach F (2002) Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT Press

Selzer M, Brandt N, Ephraim S, Zschumme P, Altschuh P, Herrmann C, Zhao Y, Tosato G, Nestler B (2020) Kadi4Mat [WWW Document]. Kadi4Mat. URL https://kadi.iam-cms.kit.edu/ (accessed 2020-12-28)

Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D

(2016) Mastering the game of Go with deep neural networks and tree search. Nature 529:484–489. https://doi.org/10.1038/nature16961

Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A, Chen Y, Lillicrap T, Hui F, Sifre L, van den Driessche G, Graepel T, Hassabis D (2017) Mastering the game of Go without human knowledge. Nature 550:354–359. https://doi.org/10.1038/nature24270

Stoffel M, Bamer F, Markert B (2018) Artificial neural networks and intelligent finite elements in nonlinear structural mechanics. Thin-Walled Struct 131:102–106. https://doi.org/10.1016/j.tws.2018.06.035

Williams RJ (1992) Training recurrent networks using the extended Kalman filter. In: [Proceedings 1992] IJCNN International Joint Conference on Neural Networks. Presented at the [Proceedings 1992] IJCNN International Joint Conference on Neural Networks, vol 4, pp 241–246. https://doi.org/10.1109/IJCNN.1992.227335

Xiao S, Stacklies W, Cetinkaya M, Markert B, Gräter F (2009) Mechanical response of silk crystalline units from force-distribution analysis. Biophys J 96:3997–4005

Robotic 4.0

Burkhard Corves, Mathias Hüsing, Stefan Bezrucav, Michael Lorenz, Nils Mandischer, Markus Schmitz, Amirreza Shahidi, Tobias Haschke, Carlo Weidemann, and Tobias Huhn

Contents

B. Corves (✉) · M. Hüsing · S. Bezrucav · M. Lorenz · N. Mandischer · M. Schmitz · A. Shahidi · T. Haschke · C. Weidemann · T. Huhn
Institute of Mechanism Theory, Machine Dynamics and Robotics, RWTH Aachen University, Aachen, Germany
e-mail: corves@igmr.rwth-aachen.de; huesing@igmr.rwth-aachen.de; bezrucav@igmr.rwth-aachen.de; lorenz@igmr.rwth-aachen.de; mandischer@igmr.rwth-aachen.de; schmitzm@igmr.rwth-aachen.de; shahidi@igmr.rwth-aachen.de; haschke@igmr.rwth-aachen.de; weidemann@igmr.rwth-aachen.de; huhn@igmr.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_24

1 Robotics in Industry 4.0

Modern robot systems are a core element of Industry 4.0 to implement **flexibility and individualization of production up to batch size one with high resource efficiency**. A goal that can only be achieved with comprehensive automation and digitalization of the entire value-added chain. Within this value-added chain, it is essential to have tailored robots with the required capabilities available at the right time and in the right place.

In addition to conventional industrial robots in production cells, driverless, automated transport systems, collaborating robots, comprehensive sensor systems and a corresponding information and communication architecture are becoming increasingly important. Thus, current technical developments are characterized by the need to enable robot systems for use in Industry 4.0, as recently shown at the Hanover Trade Fair, among others. For example by the development of multimodal capacitive tactile sensors for human-robot collaboration (HRC) applications or the integration of robots into modern, self-learning workstations 4.0 (Kim et al. 2019).

This chapter gives an overview of relevant core topics for robots in Industry 4.0: From the Internet of Robotic Things, including the necessary enabling technologies as well as stationary, mobile and collaborative systems. Finally, some projects representing the current Robotics 4.0 research initiatives at the Institute of Mechanism Theory, Machine Dynamics and Robotics (IGMR) at RWTH Aachen University are presented.

2 Internet of Robotic Things

The concept of **the Internet of Robotic Things** (IoRT) combines the Internet of Things (IoT) and Cloud Robotics. In IoT, things are physically connected to each other via the Internet and mapped virtually, which enables a comprehensive exchange of data and information as well as interaction with people via offered services. In Cloud Robotics, a network of robots is extended by a cloud, which provides additional computing capacity and data sources and thus significantly expands the possibilities for cooperation. The IoRT is a concept that makes robot services accessible to other robots, things and users (Ray 2016). With this functionality, applications can be offered that exceed the potential of an IoT, as robots are added to the network that can physically manipulate their environment, interact with it actively and act partially autonomously (Simoens et al. 2018). Any robot connected to IoRT can benefit from cloud services such as cloud computing, communication resources, large amounts of data and collective learning. At the same time, the robots provide services and applications for all connected objects (Ray 2016).

IoT focuses on the provision, exchange and monitoring of data and information (Simoens et al. 2018), whereby exchanged information enables reactions in the sense

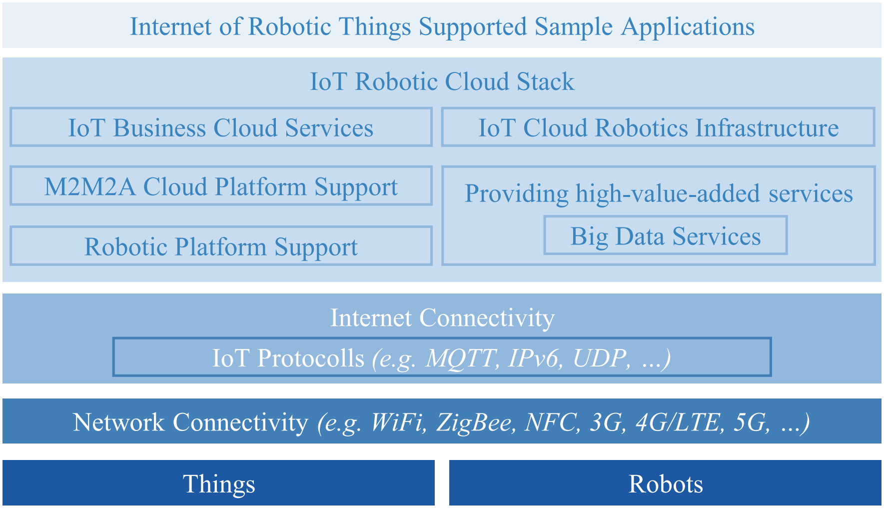


Fig. 1 Architecture for the Internet of Robotic Things (Ray 2016)

of data services. Robotic objects play an active role that goes beyond data services to enable robots to interact with their environment.

In the architecture of IoRT, services, similar to IoT, play an important role in most considerations (Ray 2016), but in the top layer there are applications that can only be provided by IoRT, not IoT (see Fig. 1), and which go beyond data services alone through interaction with the environment. Therefore, the IoRT architecture extends known IoT architectures mainly by robots and robot specific infrastructure (see Fig. 1).

3 Enabling Robotic Applications for Industry 4.0

To enable robotic applications for Industry 4.0, they must exhibit the characteristics of reactivity, flexibility, intuition, and predictability. The enabling technologies to be applied are described below.

3.1 Sensor Systems

Many different sensor types are used in the Industry 4.0 sector. Most measured variables are permanently available for all process units via the network. Sensors that are integrated in the robot are called internal sensors, while external sensors are either additionally located in the end effector or in the robot environment. Both types of sensors provide data on a wide range of physical parameters, such as motion (position, speed, and acceleration), temperature, current, pressure, force, etc. Only by suitable data processing—which can be realized by cloud services, especially in the field of IoRT—information can be obtained, which describes either the state of the robot or the state of the robot environment. In the first case one refers to proprioceptive sensor technology, in the second case one refers to exteroceptive sensor technology. Internal sensors tend to be more suitable for proprioceptive sensing and external sensors for exteroceptive sensing (Christensen and Hager 2016).

3.1.1 Proprioceptive Sensors

The encoders or resolvers of the servo drives are particularly available for recording the robot motion state. This enables the motion of each individual robot link to be recorded and the robot's pose to be determined by kinematic forward calculation. In addition, the joint and end effector velocities can be determined.

Alternatively, the current signals of the servo drives can be used to infer the state of motion of the robot with the help of a dynamic model. Another possibility for the acquisition of robot motion, especially for mobile platforms, is the use of inertial measurement units (IMU) for the detection of accelerations. Furthermore, the internal temperature sensors in the servo drives of the robot can be used for motor monitoring.

3.1.2 Exteroceptive Sensors

Exteroceptive sensor technology enables the robot to react to environmental influences. This is done, for example, with tactile sensors by measuring the acting forces from the environment. Often their function is electrically capacitive or resistive, like a smartphone display. Solutions for the measurement of a force are, besides indirect force measurement, force-torque sensors. These can absorb even very small forces and thus enable a sensitive reaction to environmental influences (Christensen and Hager 2016).

Exteroceptive sensor technology enables the robot to react to environmental influences. However, if only these sensors were used, the robot would only be able to react when contact is made. In many cases, this can already cause damage to humans or the environment. Therefore it is important for the robot to perceive its environment more comprehensively, e.g., by means of distance measurement methods. Most distance measurement methods are based on the time of flight (TOF) principle. A directed energy pulse is emitted into the environment and the time until the detection of a reflection is measured. Based on the speed of the energy carrier (e.g., light or sound), the distance to the detected object can be calculated. One of the most variable optical distance measuring methods are laser sensors. The most commonly used design is the Light Detection and Ranging (LiDAR) system.

Depending on the design, these sensors are suitable for collecting 2D as well as 3D data (Bosch 2001).

In addition to the distance of an object to the robot, its shape and figure are also of importance. Especially in collaborative manufacturing and assembly, the robot must be able to distinguish between humans and the environment. Camera systems such as RGB cameras (Red, Green, Blue), depth cameras or TOF cameras help the robot to perceive its environment. These systems are also often used for safety and work area monitoring, such as the Pilz Safety Eye System (Müller and Lotter 2018).

3.2 Online Safety of Networked Robots

A robot offers extensive possibilities to physically interact with the environment. However, the interaction between robot and environment (including humans) has an increased risk potential: High forces, masses, and speeds—which are often necessary for efficiently acting robot systems—endanger humans and machines in the environment. This risk potential must be considered especially in networking, and thus in the basic accessibility and control from the outside, because against the background of collaborative Robotics 4.0, classic protective devices are increasingly being replaced by more intelligent options for human-robot collaboration without separating protective devices. Therefore, access rights and possibilities of influence from outside accordingly have to be designed secure and sage. Inherent safety functions of robotic systems must not be bridgeable. However, this poses an increased challenge for complex technical systems, such as image-based collision detection. It must not be possible to circumvent the security of the algorithms, for example by changing the parameterization. At the same time, the flexible adaptation and further development of the systems requires appropriate external influence.

Against this background, digital security systems for managing access rights and protecting against misuse must be designed and implemented at the highest possible technical level. In comparison to other Industry 4.0 systems, robots do not have fundamentally different technical requirements for security technology, but the potential danger goes far beyond material and economic aspects in case of malfunction or misuse.

3.3 Simulation, Virtuelle Systeme, Virtuelle Abbilder

The simulation of robotic systems offers substantial advantages for the development, the production and the integration of robot in Industry 4.0 (Rüßmann et al. 2015, p. 5).

In Robotics 4.0, robots are used in a wide variety of scenarios where they work with humans and interact with other robots (Rüßmann et al. 2015, p. 5). They possess their own "intelligence," learn from their environment and make decisions on their own. Before such a Robotics 4.0 system is used in production, it can be put into operation in a virtual world. This involves modelling not only the robots but also the environment with machines, humans, sensors, and their interactions.

A corresponding modelling approach provides insights into the necessary data provision, data structuring and data feedback, which enables the generation of a so-called Digital Shadow. The Digital Shadow brings a dual capability to the system: On the one hand, ready-made data in the simulation can facilitate data mining capabilities (adaptive systems, artificial intelligence, etc.); on the other hand, the real-time data flow from the physical environment will create an adaptive and self-optimizing system based on the simulation of reality.

This enables the developer to test and optimize the new production process in the virtual world before the physical realization and thus to virtually program, test and optimize possible system adjustments or new scenarios with less effort. This greatly reduces integration and further development costs and increases quality (Rüßmann et al. 2015, p. 3). In the field of research and development, new ideas and concepts can be tested and validated. It is also easier to check borderline situations without the risk of destroying the system. The final sensor selection and the corresponding data analysis can be based on the results of such a simulation.

Another important aspect is the data exchange between the system components, which can also be tested within the virtual environment. Step by step, the simulated hardware and software components can then be replaced by software-in-the-loop and hardware-in-the-loop procedures with the real components.

There is a wide range of Simulation Software on the market. From the programs of robot manufacturers, such as KUKA-Sim (KUKA Germany GmbH 2020) and FANUC-ROBOGUIDE (FANUC Germany GmbH 2020), to manufacturer-independent software such as Webots (Webots 2018) or open source programs such as Gazebo (Koenig and Howard 2004). They all allow the physical simulation of mobile and stationary robots (serial or parallel) including sensors. Furthermore the user can model the real environment with complex CAD models or with simple building blocks.

3.4 Networking and Interfaces

Industry 4.0 means the need for systems and machines to communicate safely and seamlessly. Robust communication between different robot and sensor systems, with adaptable integration into the information network, requires a robust networking system. Manufacturer-specific systems can be used here, although the networking of different systems can encounter difficulties. An extended approach is offered by a **ROS (Robot Operating System)-environment**, which ensures and enables flexible communication between robots, sensors and other tools (Quigley et al. 2009).

The ROS environment provides libraries and tools that help software developers create robotic applications. The framework, licensed under an open-source BSD (Berkeley Software Distribution) license, provides hardware abstraction, device drivers, libraries, visualization environments, message processing, packet management, and more that enable communication of hardware units and algorithms in a network of so-called nodes. The nodes used in a robot application communicate via messages and services. In addition to the design of drivers for the corresponding robots or sensors, data are made available (Publisher) or received (Subscriber) via the communication network of the ROS environment. The corresponding interfaces are standardized in such a way that a change or addition of new units (e.g., a robot, a gripper, a sensor) or an algorithm can be carried out trouble-free and flexibly (Quigley et al. 2009).

To ensure the exchange of information, different interfaces on different communication levels (e.g., from machine to machine or machine to systems) are required. There are different Protocols, for this purpose, such as I/O, Serial, CAN, etc. In this context, the OPC-Foundation has published a number of protocols already accepted in the industry to ensure compatibility in industrial automation. Several reasons motivated the OPC Foundation to develop its new **OPC Unified Architecture specification (OPC UA),** such as unified data access, technology migration, etc. (Lange et al. 2010, pp. 95–261). The OPC UA specification is divided into several parts, which are based on the security model, the abstract services, the address space model, the information model, and the mapping of the abstract services to a concrete technology. Furthermore, different profiles for OPC UA clients and servers are specified and specializations for data access as well as alarms, conditions and programs are discussed (Mahnke et al. 2009, pp. 283–292). However, provider usually require a platform-independent specification that allows OPC applications to run on non-Microsoft systems (Lange et al. 2010, pp. 95–261), (OPC UA1). This interface offers a promising basis for Robotics 4.0 to meet the requirements not only in terms of networking and interfaces, but also in the area of simulation and virtual systems and virtual images.

3.5 Intuitive Programming

One of the objectives of current research and industry is to achieve a higher degree of automation, especially in small and medium-sized enterprises (SME). The usually time-consuming and **complex programming of robotic systems**, especially for small batch sizes, cannot relate to the required high flexibility in SMEs. Thus an intuitive possibility for programming robots even for non-experts is mandatory.

In particular, uncertainties due to human, robotic and machine collaboration make adaptability through human creativity, flexibility, learning and improvement essential for the design and configuration of systems, processes and products, especially in the context of Industry 4.0 (Jäger and Ranz 2014).

The intuitive programming of robots offers the potential to use the sensory and cognitive abilities of humans for the increased requirements. Intuition is characterized by the independent solution of open problems in the process, even if not all details have been specified (Rickert and Gaschler 2017).

Conventional robot programming is often performed by long and complex text-based code in the language of the respective robot manufacturer. These programs are difficult to understand and review in later stages. A CAD programming environment provides a higher level of reference to the application, since paths can be created based on CAD data. To establish an intuitive, explicit reference to the process, “Teaching by Showing” (Siciliano et al. 2010, pp. 215ff.) is a possibility to save robot poses during robot programming by manually moving the robot.

An extension of this procedure is programming by demonstration, which is essentially linked to imitation learning (Billard et al. 2016). A special form is the direct procedure (Orendt et al. 2016), which results in the principle of intuitive programming “demonstration.” The robot is guided by the user through a path and the stored data are made available for post-processing in a 3D programming interface (Fraunhofer IPA 2016b; Meyer et al. 2007). The “leading” or “guiding” of the user can also be implemented without contact by a control element that senses the position and direction and triggers robot movements by moving and pointing of the user (Keba 2018). This can also be done by gestures or voice control. For example, a gesture-based, intuitive robot programming can be implemented by pointing a finger (Heimann and Hügle 2018). The use of natural language also enables the intuitive programming of industrial robots (Stenmark and Nugues).

Intuitive programming makes use of a wide range of tools. To influence the pose of the interacting robot, e.g., force-moment sensors at the end-effector or at the robot base are used, see also Sect. 3.3. For contactless interference with the robot configuration, camera systems or position sensors are used (Keba 2018). During the demonstration by the user, additional sensors can be integrated to meet or control process-specific requirements. Furthermore, the user can be supported in intuitive programming by additional systems from the field of virtual or augmented reality, e.g., by displaying process-specific information (Schmidt et al. 2016).

4 Stationary Robots

Stationary robot systems are usually classified according to their mechanical structure. Serial **articulated robots** usually address various industrial applications, such as complex welding or handling processes, and therefore have the highest market share of 61%. In addition, **gantry robots** with linear actuators and **SCARA robots** (short for Selective Compliance Assembly Robot Arm) are used for manufacturing or packaging as well as pick-and-place tasks. The last class of stationary robots is represented by **parallel robots**, which are particularly used for high-speed pick-and-place applications as well as for flight and driving simulators (International Federation of Robotics (IFR) Statistical Department 2019).

Stationary robot systems are main components of modern automation concepts. The advantages of an increasing use of robotic systems are an increasing output, an improved product quality, an increased flexibility as well as lower production costs. Further factors driving the ongoing trend towards automation with robotic systems are the demographic change, the preservation of production in high-wage countries and increasing wage structures in low-wage countries.

In the context of Industry 4.0, stationary robots are further developed into intelligent systems, enabling the interaction of man and machine using ergonomic and safe interfaces. In addition to an improved energy efficiency, further objectives are realized by an increased flexibility of the systems (e.g., through a better reconfigurability and increasing adaptability) and the simplified implementation of such systems (euRobotics AISBL 2016). High implementation costs not only reduce the flexibility of the robot systems, but also represent a major challenge, especially for small and medium-sized companies beyond the high-automated series production. In this context, current research approaches especially focus on collaborative robots (Cobots). For example, a cooperation between ABB and Kawasaki is investigating uniform intuitive user interfaces for collaborative two-arm robots, which, among other things, enable an improved usability (ABB 2018).

Stationary robots are an integral part of modern cyber-physical production systems and IoRT platforms, providing comprehensive analysis and diagnostic tools. Accordingly, the mechanical health, maintenance status and related processes can be monitored in real time. Thus, on the one hand the system performance can be dynamically optimized, on the other hand maintenance and costs of idleness can be reduced by damage prediction models. In addition, there are numerous concepts and extensive industrial capabilities for an intelligent networking of production systems with stationary robots. Accordingly, almost all major robot manufacturers provide comprehensive solutions, such as Kuka SmartProduction, FANUC Zero Downtime or Yaskawa Cock-pit.

5 Mobile Robotic

Besides stationary robotics, mobile robotics is an important component of IoRT. Within **mobile robotics**, a distinction can be made between spatial robots and mobile terrain robots. Spatial robots can be further subdivided into Air and Underwater Robots, while Terrain Robots can be categorized as walking, wheeled, or tracked robots. Walking robots have an increased mobility in difficult terrain and at the same time a demanding dynamic behavior. In contrast, Wheeled Robots are characterized by stability, speed, and simplicity in kinematics and kinetics. Tracked robots are an intermediate position in terms of speed and increased mobility on difficult terrain. The use of mobile robots as wheeled robots has largely increased in the past years in industrial applications.

The term mobile robot is usually used to describe a platform equipped with wheels, which can both move along given paths and autonomously locate itself in an environment, determining and following paths and trajectories according to the task to be performed. Wheel-based mobile robots exist in a multitude of variants, which initially differ kinematically in terms of the number, type, and position of the wheels. Each individual wheel is subject to restrictions regarding the direction of motion (in the direction of the wheel axis: standard wheel or additionally normal to the wheel axis: Mecanum Wheel (Ilon 1975) or Omnidirectional Wheel (Blumrich 1974), the steerability and the direction of the drive effect. The combination of the wheels on a mobile robot thus determines the degree of mobility of the entire unit and partially enables a holonomic motion of the mobile robot, i.e., the robot pose in the workspace can be continuously changed at any time in any direction.

The development of today's mobile robotics is historically based mainly on work in the three areas of cybernetics, artificial intelligence, and robotics. The developments in the control architecture of (Freeman 2001) and (Braitenberg 1986), but also the work within the Shakey project at the Stanford Research Institute (Nilsson 1984) are to be mentioned here. Equipped with limited ability to perceive and model its environment, Shakey was a mobile robot that could perform experiments requiring planning, routing, and rearrangement of simple objects.

In the context of Robotics 4.0, mobile robotics enables flexible, situation-adapted interaction and collaboration between robots, machines, and humans. Thereby, mobile robots with knowledge of other process participants can efficiently solve tasks, perceive unknown robots, machines or humans in the environment and adapt corresponding plans. An example scenario is presented in Sect. 8.2.

6 Human-Robot Collaboration

Concerning the interaction between humans and robots, the so-called human-robot collaboration (HRC) offers a wide range of applications in the environment of Industry 4.0. This is due to the advanced development of manipulators, sensor systems and safety devices. HRC combines human qualifications, such as experience, **improvisation and learning abilities**, with the precision and high load-bearing capacity of robots. Therefore, HRC systems are especially suitable for assistance as well as for the automation of risky or monotonous activities. From this, the future scenario of an optimal, supportive division of labour between humans and robots under exclusion of hazards and risks can be developed. These systems increasingly compete with conventional automation despite inevitable safety-related restrictions and reduced performance of the HRC capable robot. In these scenarios, humans and robots work interactively and without a separating safety fence in a shared working area. Figure 2 shows the classification of HRC into different forms of interaction between humans and robots. While known forms of interaction separate manual and automated activities spatially or temporally from each other, HRC enables the simultaneous working of human and robot at the same object (Thomas 2017, pp. 32ff).

Sensor-based protection systems and intelligent control electronics are guaranteeing the physical integrity of the human. This minimizes the risk potential of the robot by active collision avoidance or by force and power limitations. Human-robot contact is therefore possible, but is accompanied by significantly reduced speeds and force and torque monitoring to minimize risks posed by the

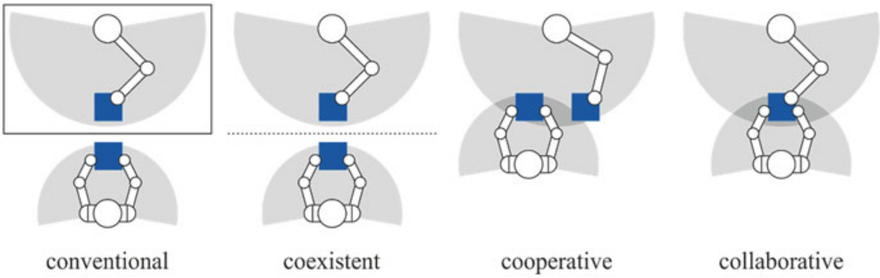


Fig. 2 Forms of interaction between humans and robots

robot and prevent injuries to humans according to standardized biomechanical limits. To minimize risks emanating from the robot and prevent injuries to humans, standardized biomechanical limits are applied (DIN ISO/TS 15066, 2016, pp. 24ff). DIN 10218-1 (2016, pp. 12ff) regulates the compliance with these limits regarding speed and force or torque as well as the equipment with appropriate protective devices. In this context, it distinguishes between speed and distance monitoring systems, which prevent physical contact between humans and robots, and systems with power and force limitation that limit the contact forces occurring during a collision. In the former case, for example, the robot is equipped with a capacitive sensor skin or workspace monitoring by means of visual sensors. Joint torque sensors or a tactile skin can actively implement the force and power limitation needed. In addition, HRC capable robots should have smooth surfaces as well as rounded edges and optional passive safety mechanisms such as compliant drives or a padded enclosure.

Currently the number of HRC capable robots is growing. These systems cover a wide range of sensor technology, performance, and costs. Thereby, it has to be distinguished between lightweight robots especially developed for HRC applications and systems consisting of classical industrial robots that are adapted for these applications.

7 Exoskeletons for Humans in Industry 4.0

Besides HRI-applications, exoskeletons become increasingly popular to directly reduce the physical strain of the workers during manual activities and other production scenarios (Schick 2018, p. 6). Accordingly, the human musculoskeletal system is supported by direct, external force to prevent stress injuries and increase the physical performance. This provides new opportunities in the work environment for both, heavy-duty jobs and older or physically handicapped employees, and thus reduction of production downtime.

In this context, exoskeletons are classified into active and passive systems. Passive systems guide the motion of certain body parts or support them with springs, e.g., as a mobile seat like the “chairless chair” currently used in car production (Farin 2017). Thus, there is no external energy supply, motion and potential energy is temporarily stored dependent on the system. In active exoskeletons, support is usually provided by electrical actuators and energy stored in accumulators (Schick 2018, p. 14ff), as in the Cray-X-System. With this regard important drivers are lightweight design, high power density and advanced control systems, which in some systems, such as the “Hybrid Assistive Limb®” (Cyberdyne 2018), are directly based on muscular signals.

In addition to technological developments, safety aspects and risk factors must also be considered during product design. Increased performance must not result in increased payloads and new developments should always focus on safety aspects in case of malfunctions and accidents (Schick 2018, p. 24). Up to now, there are no safety guidelines for exoskeletons, thus necessitating appropriate guidelines depending on the specific application (BGHW 2018, pp. 1–6).

8 Case Studies and Current Projects

In the research group “Robotics and Mechatronics” of the Institute of Mechanism Theory, Machine Dynamics and Robotics (IGMR) at RWTH Aachen University, different aspects of the use of robots in Industry 4.0 scenarios are investigated. The following sections provide an insight into current research projects.

8.1 Center of Advanced Robotics (COAR)

As part of the Center of Advanced Robotics (COAR) at IGMR, an environment for the interaction of robotic units including the necessary tools, sensors and other measurement systems was created inside a complex framework. The vision of the framework is the combination of multiple systems, whose communication with each other takes place in a flexible application scenario and who share a global absolute coordinate system.

The current development particularly focuses on the stabilization, standardization and organization of source code. Based on the version management system Git (Haenel and Plenz 2016), there are sub-groups for all elements of a robot application (algorithms, robots, sensors, grippers, etc.), under which the so-called repositories of hardware and software components are organized. The unified structure enables accelerated development through fixed workflows. Individual software packages are tested across project boundaries, which enables future combination and compatibility of all units and algorithms. An essential component is the automated quality inspection.

The proven framework at IGMR consists of robots (industrial robots, mobile manipulators, combinations of serial and mobile manipulators), sensors (indoor GPS

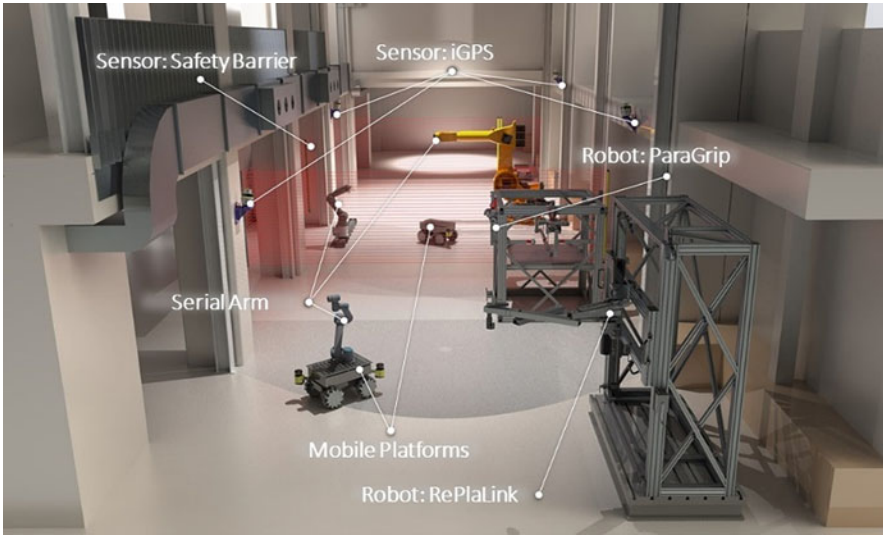


Fig. 3 Center of Advanced Robotics (COAR) at IGMR

(iGPS), lidar and radar sensors, camera systems and security devices), grippers and tools as well as corresponding algorithms for processing of the acquired data. The individual elements are brought together and used for a wide variety of projects and are constantly being expanded. For specific projects, the hardware and software components are integrated into the corresponding project repository in the form of submodules.

In the context of Industry 4.0 COAR is a flexible framework for the implementation of robot scenarios in an IoRT (Internet of Robotic Things) environment. In the sense of an IoRT network, the network and the infrastructure enable to collect relevant information of the real robot scenario, to link it with each other and to make it available in the entire framework. The additional capability of the automation system and thus the increase in intelligence enables the flexible interaction of several robot units under uncertain environmental influences. The agile organization of the COAR framework also offers increased potential to meet changing conditions in industrial robot applications even faster.

As a Proof-of-Concept Scenario, all systems of the IGMR have been installed in a new test field. For this purpose, simulation-based robots, and sensors, which have previously been virtually integrated into the new premises and thus be tested for optimal positioning, checking for collision-free operation, or setting up safety systems (see Fig. 3) have now been physically integrated. Thus, COAR has now reached a working state that allows open-source access.

8.2 *Robotic Network for Factory of Future*

Agile, freely networked assembly systems, which are characterized by the sensor-supported cooperation of several mobile and stationary robots, represent a part of the Internet of Production (IoP) through the unification of robot and sensor networks (Brecher et al. 2018; Lu 2017).

At IGMR such an assembly scenario is illustrated, which is characterized by cooperation between several mobile and stationary robot and sensor systems. Three robots cooperate in the process: a mobile manipulator as robotic worker, a mobile platform as logistics unit and a gantry robot as stationary workstation. The local referencing between these agents was realized by an iGPS, the communication takes place via a common information network. To enable the integration of this process in the context of Industry 4.0, the process is modelled as a combination of four scopes (cf. Fig. 4) (Shahidi et al. 2020a). A metamodeling structure implemented in a cloud structure supports the communication between these scopes.

8.2.1 Digital Shadow

This scope concerns the modelling and simulation of the entire system. This enables the development of a comprehensive communication system or an adaptive control system. The application of augmented reality (AR) or virtual reality (VR) in the physical environment of the robot network can also be integrated here. The data stream from the physical world can be used to optimize the model-based feedforward control and simulation results.

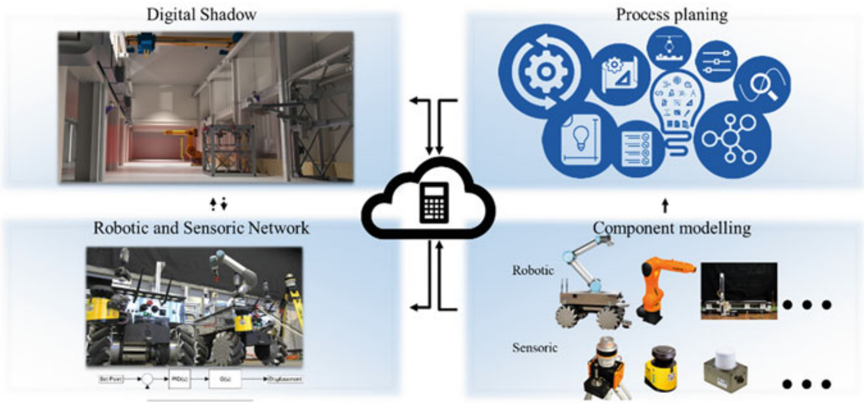


Fig. 4 Framework for an agile robotic assembly network

8.2.2 Process Planning

This scope is responsible for the optimal planning of the overall system. The system-specific information from the component modelling (e.g., robot workstation, payload, gripper information and information of the sensor network) is collected and subsequently comprehensive process parameters (e.g., station and work positions as well as the time sequence) are planned based on models.

8.2.3 Component Modelling

The independent modelling of all system components is carried out in this area. The kinematics and dynamics model of the robots, the models of the sensors, as well as the models of other system components are mapped in this scope. Model reduction and sensor data integration shall also be realized in this scope.

8.2.4 Robotic and Sensoric Network

A local control (e.g., adaptive control for disturbance compensation) as a supplement to the comprehensive control system in the cloud is processed in this scope. Real-time-capable exchange of information with the cloud enables possible process replanning (Shahidi et al. 2020b).

8.3 Production Through Robotic Multidirectional Additive Manufacturing

Additive manufacturing processes offer new possibilities to meet **constantly growing demands for product individualization and smaller batch sizes**, as production is largely independent of complex component geometries and is possible for more and more materials (Fastermann 2016). Solid bodies are created by adding material layer by layer, which represents an added value compared to conventional methods, especially when creating complex geometries with cavities and undercuts (Gebhardt 2016).

In Fused Layer Manufacturing (FLM) the material (usually thermoplastics) is successively deposited in a flat layer through a nozzle. This enables the construction of cost-effective manufacturing systems (3D-Drucker, Fig. 5a). However, the component properties are limited by inhomogeneity due to the layered structure (Siebrecht et al. 2016). In addition, there is the need for non-functional support structures for overhangs and inclined surfaces (Allen and Trask 2015).

The limitations of conventional 3D printers mentioned above can be reduced by adjusting the layer thicknesses and material quantities (Hope et al. 1997) or curved

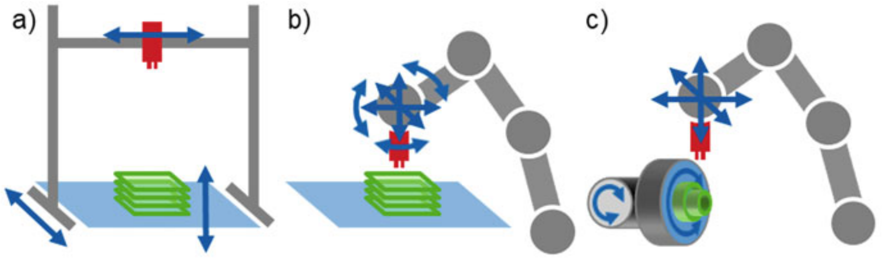


Fig. 5 (**a**) classic 3D printers, (**b**) and (**c**) concepts for the realization of additional degrees of freedom in additive manufacturing

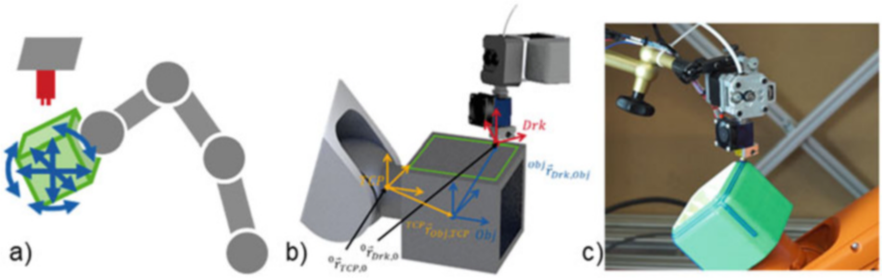


Fig. 6 5D-Printing process at IGMR

layers in near-surface areas (Jin et al. 2017), which not only improves the printing speed but also the mechanical component properties (Spoerk et al. 2017). The radii of curvature are limited by the fixed orientation of the nozzle. The addition of further degrees of freedom is therefore a logical further development to achieve a higher curvature of the layers. By using robot-guided nozzles, three rotational degrees of freedom can be realized in addition to three translational ones (Fraunhofer IPA 2016a), Fig. 5b. The combination of object rotation and translational movement (Knabel 2016) of the nozzle creates five degrees of freedom, Fig. 5c.

Within the research framework of IGMR, an ongoing project (Detert et al. 2017; Schmitz et al. 2021) is developing and implementing a 5D printing technology in which the object to be printed is continuously guided by a jointed-arm robot under a stationary 3D print head (see Fig. 6). The shown addition of degrees of freedom on the component reduces the need for support structures, since the overhang can be influenced within certain limits. The strength of the components can also be increased, since the layers can be arranged along the loads in the component.

One additional focus is the optimization of the trajectory using a redundant robot degree of freedom. Significant increases in printing speed have been demonstrated (Detert 2018).

8.4 Inclusive Workplaces of the Next Generation

Staffing policy in industry as well as the crafts and the service sector is increasingly characterized by aging population in connection with the demographic change. In this context, the Industry 4.0 initiative offers outstanding opportunities to create new jobs, particularly by opening up new technologies and through efficient communication between men and innovative machines and systems. Within the project **"Next Generation - Developing inclusive jobs by flexible robot solutions"** IGMR of RWTH Aachen University addresses this concept by developing innovative and inclusive workplaces especially for physically and cognitively disabled people who have disadvantages in the general labour market. In this respect, human-robot collaboration serves as a tool for participation in working life and at the same time creates a new type of application in the context of Industry 4.0.

The central element of the project is the use of lightweight robots as flexible assistance for the individual enhancement of physical and cognitive abilities of severely multiple handicapped persons (Fig. 7). In this way, the conflict between economic productivity on the one hand and the inclusion of people with disabilities on the other can be resolved: This enables companies to access a new group of employees for technically supported and demanding activities. In return, handicapped persons have the opportunity of equal participation in all areas of life—and a more satisfied and independent life through the performed work.

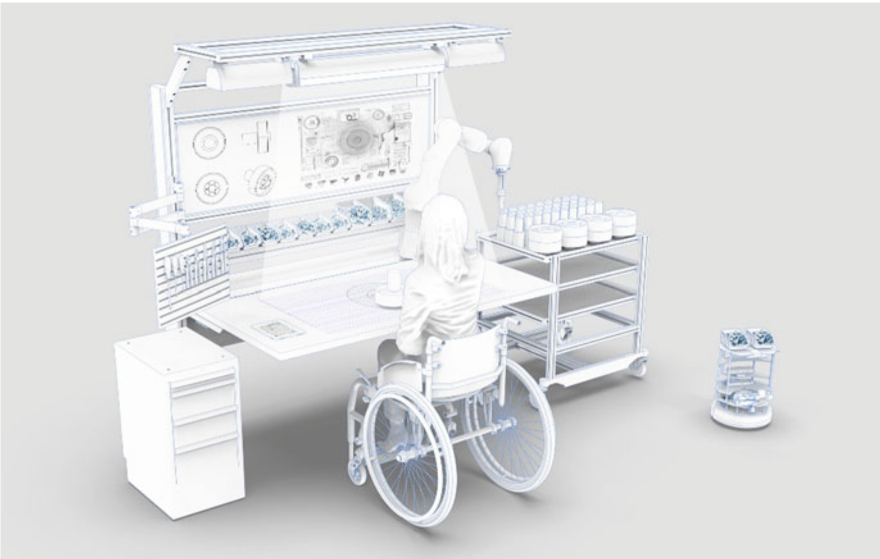


Fig. 7 Inclusive workplaces based on human-robot collaboration

References

ABB (2018) Architektur: Ein Roboter verbindet digitale und physische Welt: ABB-Technologie im Mittelpunkt des Projekts Communication Landscapes. ABB about - Das Kundenmagazin von ABB Deutschland:26–27

Allen RJA, Trask RS (2015) An experimental demonstration of effective Curved Layer Fused Filament Fabrication utilising a parallel deposition robot. Addit Manuf 2015:78–87

BGHW (2018) Einsatz von Exoskeletten an gewerblichen Arbeitsplätzen. Einsatz von Exoskeletten an gewerblichen. https://www.dguv.de/medien/fb-handelundlogistik/pdf-dokumente/exoskelette.pdf

Billard AG, Calinon S, Dillmann R (2016) Learning from humans. In: Springer handbook of robotics. Springer, pp 1995–2014

Blumrich J (1974) Omnidirectional wheel(US3789947A)

Bosch T (2001) Laser ranging: a critical review of usual techniques for distance measurement. Opt Eng:10. https://doi.org/10.1117/1.1330700

Braitenberg V (1986) Vehicles: experiments in synthetic psychology. MIT Press

Brecher C, Epple A, Knape S, Schmidt S (2018) Netzwerkarchitekturkomponenten auf Werkstattebene. ZWF 113:342–345. https://doi.org/10.3139/104.111904

Christensen HI, Hager GD (2016) Sensing and estimation. In: Springer handbook of robotics. Springer, pp 91–112

Cyberdyne (2018) What's HAL®: the world's first cyborg-type robot HAL®. https://www.cyberdyne.jp/english/products/HAL/

Detert T (2018) Restriktionsfreie Additive Fertigung: Promotionsvortrag von Tim Detert. IGMR, Aachen

Detert T, Lorenz M, Schmitz M, Hüsing M, Corves B (2017) Robotergeführte Objektmanipulation für die generative Fertigung. In: Beitelschmidt M (ed) Tagungsband 12. Kolloquium Getriebetechnik: Technische Universität Dresden, 20.-22. September 2017. TUDpress, Dresden

DIN 10218-1 (2016) Industrieroboter – Sicherheitsanforderungen – Teil 1: Roboter. Internationale Organisation für Normung

DIN ISO/TS 15066 (2016) Roboter und Robotikgeräte – Kollaborierende Roboter. Deutsches Institut für Normung

euRobotics AISBL (2016) Multi-Annual Roadmap (MAR) for Horizon 2020: SPARC Robotics. https://www.eu-robotics.net/cms/upload/topic_groups/H2020_Robotics_Multi-Annual_Roadmap_ICT-2017B.pdf

FANUC Germany GmbH (2020) Simulation Software ROBOGUIDE. https://www.fanuc.eu/de/en/robots/accessories/roboguide. Accessed 2 Nov 2020

Farin T (2017) Ein Roboter für den Körper: Exoskelette. http://www.faz.net/aktuell/beruf-chance/beruf/exoskelette-maschine-zum-anziehen-die-koerperliche-arbeit-erleichtern-15341414.html. Accessed 3 Dec 2018

Fastermann P (2016) 3D-Drucken: Wie die generative Fertigungstechnik funktioniert, 2., aktualisierte Auflage. Technik im Fokus. Springer Vieweg, Berlin, Heidelberg

Fraunhofer IPA (2016a) 3D-gedruckte Leichtbauteile für den Mittelstand (Endlosfaserdrucken). https://www.ipa.fraunhofer.de/de/presse/presseinformationen/2016-07-21_3d-gedruckte-leichtbauteile-fuer-den-mittelstand.html. Accessed 9 Apr 2018

Fraunhofer IPA (2016b) INTEACH - Intuitive Roboterprogrammierung. https://www.ipa.fraunhofer.de/content/dam/ipa/de/documents/Kompetenzen/Roboter%2D%2Dund-Assistenzsysteme/Produktblatt_Inteach.pdf. Accessed 17 Jan 2019

Freeman WJ (2001) Biographical essay on W. Grey Walter. Encyclopedia of Cognitive Science, Macmillan, Basingstoke

Gebhardt A (2016) Additive Fertigungsverfahren: Additive Manufacturing und 3D-Drucken für Prototyping - Tooling - Produktion, 5th edn. Hanser, München

Haenel V, Plenz J (2016) Git: Verteilte Versionskontrolle für Code und Dokumente

Heimann O, Hügle J (2018) Intuitive Roboterprogrammierung per Fingerzeig. ATZextra 23:24–27

Hope RL, Roth RN, Jacobs PA (1997) Adaptive slicing with sloping layer surfaces. Rapid Prototyp J 3:89–98. https://doi.org/10.1108/13552549710185662
Ilon BE (1975) Wheels for a course stable selfpropelling vehicle movable in any desired direction on the ground or some other base(US3876255A)
International Federation of Robotics (IFR) Statistical Department (2019) World Robotics – Industrial Robots 2019. VDMA Verlag
Jäger A, Ranz F (2014) Industry 4.0: challenges for the human factor in future production scenarios. In: 4th Conference on Learning Factors
Jin Y, Du J, He Y, Fu G (2017) Modeling and process planning for curved layer fused deposition. Int J Adv Manuf Technol 91:273–285. https://doi.org/10.1007/s00170-016-9743-5
Keba (2018) KeTop T10 directMove. https://www.keba.com/web/downloads/industrial-automation/data-sheets/KeTop_T10_Zusatzbroschuere_D.pdf. Accessed 17 Jan 2019
Kim W et al (2019) Adaptable workstations for human-robot collaboration: a reconfigurable framework for improving worker ergonomics and productivity. IEEE Robot Autom Mag 26 (3):14–26. https://doi.org/10.1109/MRA.2018.2890460
Knabel J (2016) Stratasys zeigt neue FDM Fertigungsanlage mit 8-Axen. https://3druck.com/drucker-und-produkte/stratasys-zeigt-neue-fdm-fertigungsanlage-mit-8-axen-1448149/. Accessed 28 Mar 2018
Koenig N, Howard A (2004) Design and use paradigms for gazebo, an open-source multi-robot simulator. In: 2004 IEEE/RSJ International Conference, pp 2149–2154
KUKA Germany GmbH (2020) Software KUKA.Sim. https://www.kuka.com/de-de/produkte-leistungen/robotersysteme/software/planung-projektierung-service-sicherheit/kuka_sim. Accessed 2 Nov 2020
Lange J, Iwanitz F, Burke TJ (2010) OPC: from data access to unified architecture, 4., rev. ed. VDE-Verl., Berlin
Lu Y (2017) Industry 4.0: a survey on technologies, applications and open research issues. J Ind Inf Integr 6:1–10. https://doi.org/10.1016/j.jii.2017.04.005
Mahnke W, Leitner S-H, Damm M (2009) OPC unified architecture. Springer, Berlin
Meyer C, Hollmann R, Parlitz C, Hägele M (2007) Programmieren durch Vormachen für Assistenzsysteme–Schweiß-und Klebebahnen intuitiv programmieren (Programming by Demonstration for Assistive Systems–Intuitive Programming of Welding and Gluing Trajectories). it-Information Technology 49:238–246
Müller U, Lotter E (2018) Intelligente (smarte) Robotik. In: Wagner RM (ed) Industrie 4.0 für die Praxis: Mit realen Fallbeispielen aus mittelständischen Unternehmen und vielen umsetzbaren Tipps. Springer Gabler, Wiesbaden, pp 115–127
Nilsson NJ (1984) Shakey the robot: Technical Note 323. SRI International
OPC UA1 OPC UA Specification: Part 1: Overview and Concepts
Orendt EM, Fichter M, Henrich D (2016) Robot programming by non-experts: intuitiveness and robustness of one-shot robot programming. In: 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 2016, pp 192–199
Quigley M, Gerkey B, Conley K, Faust J, Foote T, Leibs J, Berger E, Wheeler R, Ng A (2009) ROS: an open-source Robot Operating System:5
Ray PP (2016) Internet of robotic things: concept, technologies, and challenges. IEEE Access 4: 9489–9500. https://doi.org/10.1109/ACCESS.2017.2647747
Rickert M, Gaschler A (2017) Robotics library: an object-oriented approach to robot applications. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017, pp 733–740
Rüßmann M, Lorenz M, Gerbert P, Waldner M, Justus J, Engel P, Harnisch M (2015) Industry 4.0: the future of productivity and growth in manufacturing: The Boston Consulting Group. https://www.bcg.com/de-de/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries.aspx. Accessed 3 Dec 2018
Schick R (2018) Einsatz von Exoskeletten in Arbeitssystemen: Stand der Technik - Entwicklungen - Erfahrungen. Sicherheitswissenschaftliches Kolloquium, Wuppertal-Vohwinkel

Schmidt E, Schweizer S, Hendrich D (2016) Augmented reality robot operation interface with Google Tango. In: ISR 2018; 50th International Symposium on Robotics, pp 1–8
Schmitz M, Wiartalla J, Gelfgren M, Mann S, Corves B, Hüsing M (2021) A robot-centered path-planning algorithm for multidirectional additive manufacturing for WAAM Processes and pure object manipulation. Appl Sci 11:5759. https://doi.org/10.3390/app11135759
Shahidi A, Hüsing M, Corves B (2020a) A decentralized structure for the digital shadows of internet of production. In: Venture G, Solis J, Takeda Y, Konno A (eds) ROMANSY 23 - Robot Design, Dynamics and Control: Proceedings of the 23rd CISM IFToMM Symposium, vol 601, pp 179–186
Shahidi A, Peitsch K, Klimmek S, Hüsing M, Corves B (2020b) Real-time motion-planning in dynamic environments via enhanced velocity obstacle. In: Lenarčič J, Siciliano B (eds) Advances in Robot Kinematics 2020. ARK 2020.: Springer Proceedings in Advanced Robotics, vol 15. Springer, Cham, pp 158–165
Siciliano B, Sciavicco L, Villani L, Oriolo G (2010) Robotics: modelling, planning and control. Springer Science & Business Media
Siebrecht J, Jacobs G, Beckers A, Löwer M, Schmidt W, Körber M (2016) Einfluss von Prozessparametern auf die Funktionserfüllung von Bauteilen beim Fused Deposition Modeling. In: Krause D, Paetzold K, Wartzack S (eds) Design for X : Beiträge zum 27. DfX-Symposium. TuTech Verlag, Hamburg, pp 123–134
Simoens P, Dragone M, Saffiotti A (2018) The Internet of Robotic Things. Int J Adv Robot Syst 15. https://doi.org/10.1177/1729881418759424
Spoerk M, Arbeiter F, Cajner H, Sapkota J, Holzer C (2017) Parametric optimization of intra- and inter-layer strengths in parts produced by extrusion-based additive manufacturing of poly(lactic acid). J Appl Polym Sci 134:1. https://doi.org/10.1002/app.45401
Stenmark M, Nugues P Natural language programming of industrial robots. In: 44th International Symposium on Robotics (ISR), pp 1–5
Thomas C (2017) Entwicklung einer Bewertungssystematik für die Mensch-Roboter-Kollaboration. Dissertation, Shaker Verlag GmbH
Webots (2018) Webots. http://www.cyberbotics.com. Accessed 31 Jan 2019

Digital Twins in the Product Life Cycle of Additively Manufactured Components

Talu Ünal-Saewe, Christian Vedder, Simon Vervoort, and Johannes Henrich Schleifenbaum

Contents

1 Introduction

Production technologies in all areas of the value chain must adapt quickly to short innovation and development cycles, growing product diversity, small batches, rapidly changing material flows, and increased energy and resource costs. To meet these requirements, the industry must be able to design processes and production chains quickly and faultlessly, regardless of the applications, such as the manufacture, coating, or repair of new components.

These requirements can be addressed by using technologies enabling the industry to utilize broad production processes without sacrificing the economic benefits of economies of scale in production. As there is an increasing demand for a high degree of flexibility and geometric freedom, laser-based Additive Manufacturing

T. Ünal-Saewe (✉) · C. Vedder · S. Vervoort
Fraunhofer-Institute for Laser Technology ILT, Department of Laser Material Processing, Aachen, Germany
e-mail: talu.uenal-saewe@ilt.fraunhofer.de; christian.vedder@ilt.fraunhofer.de; simon.vervoort@ilt.fraunhofer.de

J. H. Schleifenbaum
RWTH Aachen University, Chair for Digital Additive Production, Aachen, Germany
e-mail: johannes.henrich.schleifenbaum@dap.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_25

(AM) shows the greatest potential for such applications. In this context, the most important AM processes are powder-based Laser Powder Bed Fusion (LPBF) and Laser Material Deposition (LMD).

However, today's AM production facilities do not yet have the required functions to integrate these two processes into an Industry 4.0 environment. One of the main deficits is that currently available plants are not equipped with sufficient sensor technology and interfaces for comprehensive process analysis. Starting from downstream product evaluation, iterative processes (trial & error) can only be used to determine which adaptation of the process and machine parameters will lead to the desired product quality. This leads to expensive, lengthy development processes. In addition, the available data hardly allow any conclusions to be drawn about the machine condition within the framework of predictive maintenance. Furthermore, the effects of adjusting individual parameters on the condition of the machine cannot be adequately described. Due to these limitations, no complete process data-based quality management and maintenance management can be integrated into the production process.

For Industry 4.0 to be successfully implemented along the entire product life cycle of a component, intelligent components must be developed and used, components that can absorb and communicate data and stimuli from their environment. The data obtained, e.g., on the thermal or mechanical load of the component over the entire product life cycle, generate tangible added value for the digital twin. In addition to improving how well the component design can be evaluated, the data enable predictive maintenance to make cost-intensive systems and components more readily available. In conventional, subtractive component manufacturing, sensors are attached externally to the component. Accordingly, measured values can only be recorded on the component shell, but temperatures and mechanical stresses on the component shell differ in many cases from the respective values inside the component. Integrating sensors, e.g., thermocouples, in the component through holes is possible, but the contacting them in the holes is complex and the contact between sensor and component is not ideal. If necessary, the desired measuring points can be placed in the component only by modifying it, via drilling holes, or not at all. The geometric freedom afforded by additive manufacturing processes can be used, as shown below, to circumvent these problems in conventional manufacturing and place sensors in components in a functionally optimized manner.

2 Metal-Based Additive Manufacturing Processes

2.1 Laser Powder Bed Fusion

Laser Powder Bed Fusion (LPBF) is a powder-based generative manufacturing process. As with the LMD process, three-dimensional components can be produced layer by layer by locally remelting a powder material.

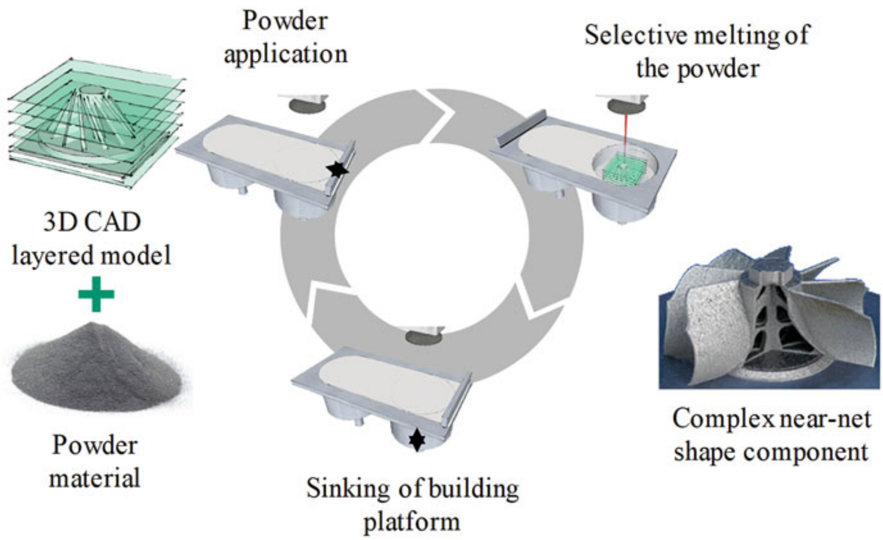


Fig. 1 LPBF process steps (source: Fraunhofer ILT)

After the geometric data of the component to be created is provided in the form of a CAD model, it is sliced into layers with a defined layer thickness using suitable software. In addition, an exposure strategy is defined for each layer. Then, in an iterative manner, a powder layer is first applied to the substrate plate with a coater and is selectively exposed. Basic process parameters of the LPBF process are laser power, scan speed, beam diameter, track distance and layer thickness. For process-related reasons, the applied powder layer is fused with the underlying layers of the component (cf. Fig. 3). The build platform is then lowered by the height of a layer thickness. The process is repeated until the component has been completely built. Then it can finally be removed (see Fig. 1); Poprawe 2005).

2.2 *Laser Material Deposition (LMD)*

Laser Material Deposition (LMD, also known as laser-based Directed Energy Deposition) is suitable for coating, additive manufacturing and repairing components.

During the LMD process, the surface of the workpiece is melted by a laser beam. The powder is transported into the melt pool by using argon or helium as the conveying gas (Hoffmann 1997).

LMD processes have long been used for the repair and surface protection of components in various industries. Multi-axis machines enable the generation of three-dimensional, almost net-like geometrical shapes (see Fig. 2). LMD differs from the LPBF process as it is able to produce larger components, apply on 3D

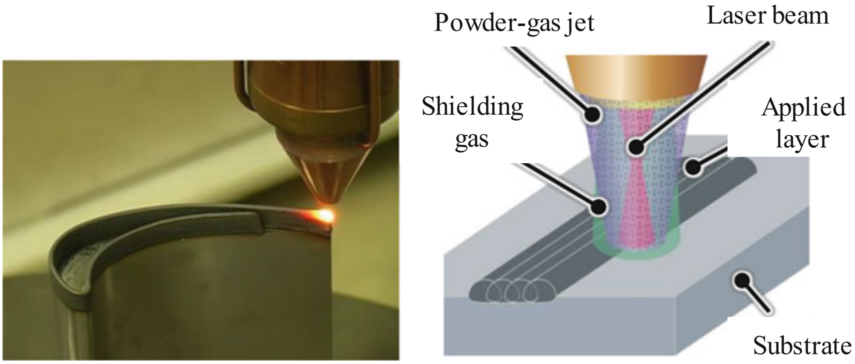


Fig. 2 Repair of a turbine blade using LMD (left), (source: Fraunhofer ILT); process principle of the LMD process (right), (source: Fraunhofer ILT)

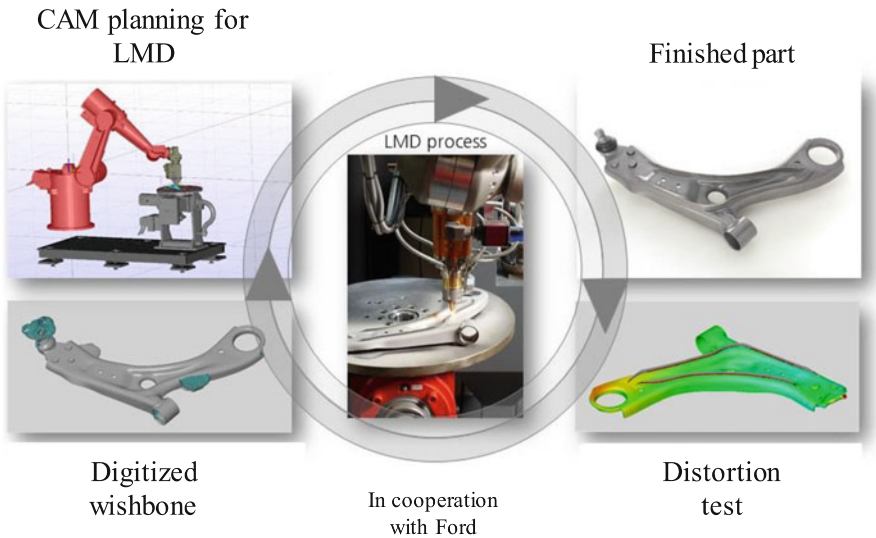


Fig. 3 Digital and physical process chain for local reinforcement of a wishbone using hybrid additive LMD process (source: Fraunhofer ILT)

surfaces, and produce multi-material components. Figure 3 shows the digital and physical process chain for a hybrid-additive LMD process (Hoffmann 1997).

The local reinforcement of a wishbone starts with digitizing the component (e.g., with a GOM system). In the second step, offline programming of the LMD machine is performed using the captured data. Suitable CAD/CAM tools are required for this step. The LMD process can then be started, resulting in a new, individualized component (Poprawe et al. 2019).

3 Digital Twin in Additive Manufacturing

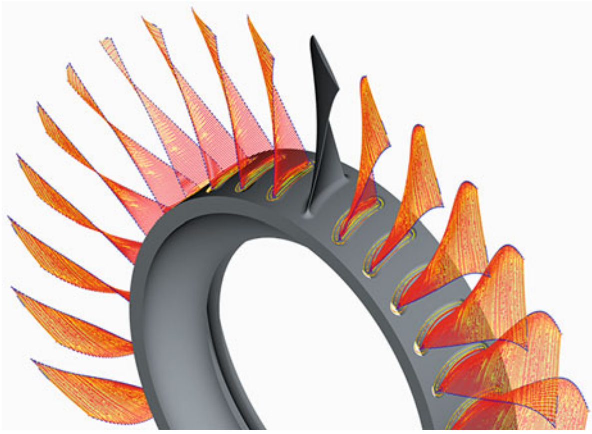
The basis for developing a digital twin is the selection of suitable variables and signals that present information about the process and that are measurable. For laser-based manufacturing processes, sensors are selected in particular that capture properties of the process signatures representing the result of how process influence variables interact. This is the thermal emission of the molten bath, in which the process-related dynamic properties of heating, melting and solidification of the powder material are combined. In this process, the melt pool grows or shrinks, for example, depending on the amount of energy introduced. Irregularities in the melt-pool behavior can be detected based on the melt pool radiation (Spears and Gold 2016; Tapia and Elvany 2014).

To gain a profound understanding of the process, however, merely observing process signatures is not sufficient since the origin of process irregularities or defects cannot be deduced from process signatures alone. Therefore, alongside acquiring process signatures, additional data of the machine control (machine axis states, peripheral data, etc.) should be acquired to generate a data-based integrated image of the process. A suitable sensor network and potential data acquisition systems must be selected and evaluated based on investigations regarding how the machine is configured and how the measurement data to be recorded are defined. Particular consideration must be given to the connection and communication options of the sensor types with different characteristics, as well as the machine control with different fieldbus systems and communication protocols. Thus, on the one hand, sensors can be integrated in a safe and stable manner and, on the other, proprietary machine control can be connected for synchronous and latency-free transmission of the machine data. Fully synchronized machine data and process monitoring data open up new potential of documenting the process for quality management of the manufactured components.

Developing the digital twin of the additive manufacturing process and the underlying data analysis pose a particular challenge since data structures are so complex. This results from the separate acquisition of the process data and the machine data as well as from the large number of processes influencing variables. The aim is to automate data cleaning, combination, and evaluation in such a way that manual sifting of the various data sources can be dispensed with, and the process digitized. To do this, the data extracted from the machine is transferred to a suitable data acquisition system, where it is sorted and, if necessary, aggregated. Invalid and erroneous data must be detected and discarded.

Traditionally, process data is documented in a time-resolved manner. Now, by coupling both the data from the controls of the production plants and the sensor data, the process data can be linked with the respective axis position in the component. Thus, a link between the quality of discrete volume elements of a manufactured component and the process and machine condition can be established while the components are examined. All process and machine data will be used in the long term to predict machine and process defects in advance by means of suitable

Fig. 4 Prototype of a digital process twin for the hybrid manufacture of turbine blades for a BLISK using LMD (source: Fraunhofer ILT)



analyses—as predictive maintenance—and to take measures to prevent them in good time.

In Fig. 4 such a type of process data visualization using the example of a process for the hybrid manufacture of turbine blades on a BLISK (**Bl**ade **I**ntegrated Di**sk**) is shown. Melt pool emission data (such as melt pool temperatures) are synchronized with the online acquired axis coordinates, so that a spatially resolved representation of the relevant process signature is achieved by false color representation. This methodology enables process irregularities to be mapped with spatial accuracy and critical component areas to be delimited during quality inspection.

4 Sensor Integration for the Production of Components Enabled for Industry 4.0

Additive manufacturing makes it possible to integrate sensors into components in new ways. Every point inside and outside the component is temporarily accessible during the manufacturing process. Thanks to this, sensors can be freely positioned in the component, even in places that are not accessible with conventional integration methods. Whereas holes for supply lines for signal transmission and energy supply can only run in a straight line, in additively manufactured components it is possible to freely design the shape and position of the corresponding channels in space. The sensors are fully integrated into the component and, thus, protected from environmental influences, provided that the channels for the supply lines are sealed, e.g., with resin (see Fig. 5).

Current research is examining the use of wireless systems for signal and energy transmission for sensors and the LPBF process is particularly suitable for integrating them into components. The heat input into already manufactured component

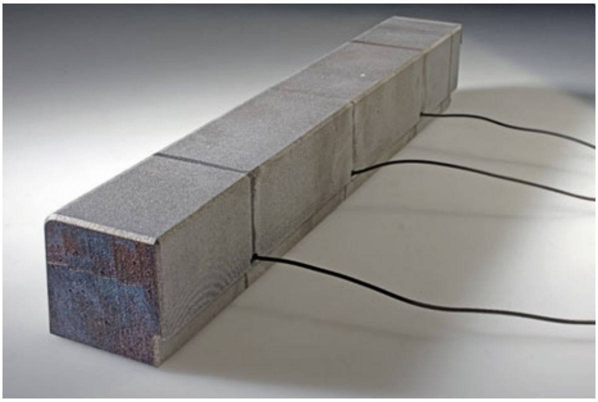


Fig. 5 Mold insert with three pressure sensors integrated close to the contour (source: Fraunhofer ILT)

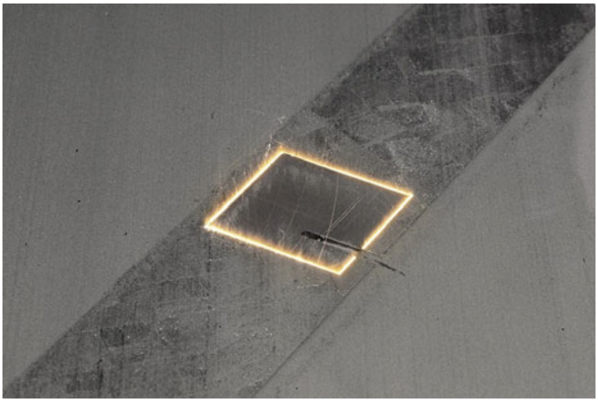


Fig. 6 Connection of a pressure sensor using a process laser-beam source (source: Fraunhofer ILT)

segments is generally smaller in LPBF than in the LMD process. This means that sensors with lower thermal destruction thresholds can be integrated.

To insert the sensors, the manufacturing process is paused at the desired layer. The sensor can be inserted either in-situ, i.e., inside the process chamber with the substrate plate mounted, or outside the process chamber by dismantling the substrate plate. A corresponding cavity for the sensor is provided in the CAD model. If the sensor is a wired sensor, a channel is also designed for the cabling. The sensor must be located at the highest point of the channel in the build-up direction so as not to obstruct the coating application in the further course of the manufacturing process.

In addition to clamping or bonding for metallic sensor components, the sensor can also be physically bonded to the component using the process laser beam source. In the case of temperature sensors, this method offers an optimized response time due to improved heat transfer, and in the case of pressure sensors, stronger measurement signals due to improved force transmission. The connection of a pressure sensor on a rectangular substrate is shown in Fig. 6.

The building process is then continued, if necessary, integrating further sensors, until the component is fully manufactured. In addition to embedding discrete sensors and cables inside the component, these components can also be additively manufactured by applying layers to the surface of conventionally or additively manufactured components. Digital printing and laser processes, among others, are used for this purpose, enabling production of up to batch size 1 without retooling and thus the individualization of mass products. The process chain for manufacturing the layers is inline-capable and can be automated:

In addition to examining how pressure and temperature sensors can be integrated in components manufactured with LPBF, current research is also looking at producing AM-based strain gauges on prefabricated components for use in Industry 4.0 applications. For this purpose, after a possible laser-based cleaning of the surface from lubricants or similar, a necessary electrical insulation layer is applied in the first step by means of a printing process: This layer can be applied over a wide area via dip or spray, or locally and selectively via inkjet, aerosol, dispensing, pad, screen printing processes, etc. The materials used are, for example, polymers, glass, or the like, which are available in nano- or microparticulate as well as sol-gel form or monomer solutions. After this initial layer is applied, a thermal treatment is carried out to create an electrical functional layer from the wet printed layer: Initially, the component is dried to remove the liquid (solvent) as well as solid, organic components (binder) necessary for printing. Then, the layer is functionalized, which involves sintering, melting, crosslinking, etc. of the functional components such as nano/microparticles, monomers or precursors.

Thermal post-treatment can be performed by conventional oven processes, etc., or by modern laser processes. Some material systems can be processed by means of UV radiation; here, too, the conventional UV lamp can often be replaced by a laser.

The advantages of laser processes are the short interaction times and fast heating and cooling rates in small, locally limited volumes, all of which enable a short-term coating temperature above the destruction threshold of the substrate without damaging it. Thus, temperature-sensitive components can now be coated, something previously impossible. This also includes, for example, processing hardened steels.

After the insulation layer is applied, the other functional layers such as conductive tracks, sensor measurement structures, encapsulation, etc. follow, whereby the sensor system is assembled layer by layer. Depending on the application, the individual layer thicknesses vary between 1 and 30 μm, in exceptional cases also up to 300 μm (Fig. 7).

Currently, the conventional application of foil strain gauges still mostly consists of manually gluing prefabricated foil strain gauges onto components, a process that can lead to non-reproducible results despite great skill on the part of the person carrying out the work. The application is also limited to easily accessible areas of the component; from there the energy and signal lines run via soldered-on cables to the evaluation electronics. In the long term, additive manufacturing of strain gauges could replace the manual application of conventional strain gauges in many areas: The gauge can be integrated into the manufacturing process of the component and, together with the deposition of component-connected conductor paths, can enable application in areas that are later difficult to access. Conceivably, it will become possible to retrofit strain gauges (subsequent equipping) on components in

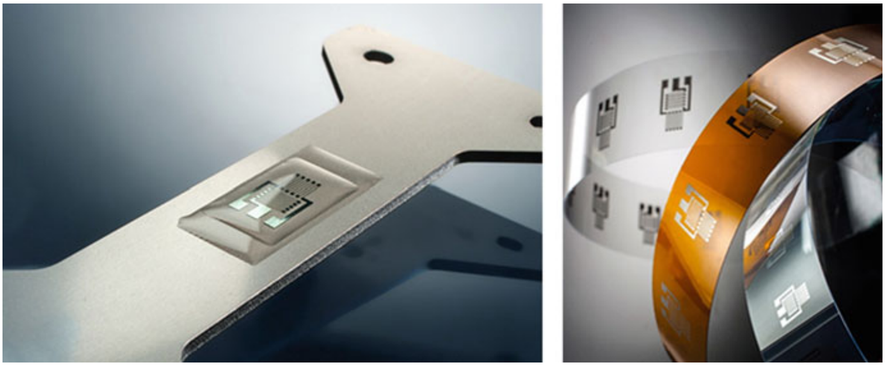


Fig. 7 Encapsulated strain gauge (half bridge) on a metal component (left) and on temperature-sensitive polymer films (right), manufactured using additive printing and laser processes (source: Fraunhofer ILT)

Fig. 8 Strain gauges (full bridge) and other sensors in various stages of completion on a rolling bearing outer ring with operating temperatures of up to 500 °C, manufactured using additive printing and laser processes (source: Fraunhofer ILT, IKTS)

manufacturing systems, such as on rollers installed in production lines (Figs. 8 and 9).

5 Summary

In this article, the application of additive manufacturing processes is introduced in the context of Industry 4.0. The LMD process is used to demonstrate how digital twins of manufacturing processes are derived and developed. The goal of this twin is a digital image of the process for testing and ensuring the desired component quality without the need for subsequent expensive, time-consuming testing procedures. As a

Fig. 9 Additive manufacturing of electrical functional layers on a 3D component using printing and laser processes (source: Fraunhofer ILT)

basis for developing these digital twins, suitable variables are selected, which present information about the process and are measurable. Important points in the approach used are as follows:

– Measurement of process signatures by looking at the melt-pool issues in laser-based welding processes
– Extraction of process data from the control system of the production machine and the periphery
– Synchronization of this data with the axis positions during production to detect process defects in the component with spatial resolution

The object of subsequent work will be not only to detect process irregularities, but also to establish a correlation between all the measurement data obtained and the quality of the manufactured components. Challenges regarding this task are the following:

• The non-deterministic nature of laser-based manufacturing processes and the resulting unpredictability of the process flow
• Lack of models that reliably represent a link between input and output variables of the process
• Larger amount of input parameters compared to conventional machining processes and the high interaction of these parameters
• Difficulty in capturing process outputs and quality measures in-situ

For this purpose, future work will compare and evaluate the results of simulation models with machine learning methods.

In addition, it will be shown how the advantages of geometric freedoms in additive manufacturing processes can be used to produce components for Industry 4.0 applications and to integrate sensors in them. Current research work is investigating the potential of this approach and the quality of the measurement data from the integrated sensors.

References

Hoffmann E (1997) Laserstrahlgenerieren – Ein Verfahren zur Herstellung metallischer Bauteile. Dissertation, Aachen

Poprawe R (2005) Lasertechnik für die Fertigung. Grundlagen, Perspektiven und Beispiele für den innovativen Ingenieur, Berlin

Poprawe R, Hinke C, Meiners W, Eibl F, Zarei O, Voshage M, Ziegler S, Schleifenbaum JH, Gasser A, Schopphoven T, Willenborg E, Flemmer J, Weingarten C, Finger J, Reininghaus M (2019) Digital photonic production along the lines of industry 4.0. SPIE LASE, 2018, San Francisco, California, United States, 2–7 February 2019, Proc. SPIE 10519, Laser Applications in Microelectronic and Optoelectronic Manufacturing (LAMOM) XXIII, 051907

Spears TG, Gold SA (2016) In process sensing in selective laser melting (SLM) additive manufacturing. Integr Mater Manuf Innov 5(1)

Tapia G, Elvany A (2014) A review on process monitoring and control in metal-based additive manufacturing. J Manuf Sci Eng 136(6)

Textile 4.0

Thomas Gries, Frederik Cloppenburg, Fabian Schreiber, and Maximilian Kemper

Contents

1 Introduction

In many status reports, conference proceedings, and books on Industry 4.0 and the digital transformation of German industry, it is repeatedly mentioned (e.g., Jacobs et al. 2017) that people are a central component in this transformation process. This is described by the term **Work 4.0** and topics such as **human-technology**

T. Gries (✉) · F. Cloppenburg · M. Kemper
Institut für Textiltechnik (ITA) of RWTH Aachen University, Aachen, Germany
e-mail: thomas.gries@ita.rwth-aachen.de; frederik.cloppenburg@ita.rwth-aachen.de; Maximilian.Kemper@rwth-aachen.de

F. Schreiber
Textilforschungsinstitut Thüringen-Vogtland e.V., Managing Director, Greiz, Germany
e-mail: f.schreiber@titv-greiz.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_26

interaction, usability, assistance systems and working environments of the future are mentioned. In the German textile industry, employee-friendly design is also a major challenge due to factors such as the **demographic structure** of the workforce and the degree of automation in production. On the following pages, this book chapter will give a rough selection of past and current efforts at the Institute of Textile Technology at RWTH Aachen University in the field of digitalization and especially **Work 4.0. Fields of research in the digitization of the textile industry are shown in** (Fig. 1).

2 Research Landscape of Technical Digitalization Projects in the Textile Industry

For more than 80 years, the Institute of Textile Technology (ITA) at RWTH Aachen University has, among other topics, been conducting research, on implementation projects to increase productivity and quality in textile production and the automation of textile machines. In recent years, projects have increasingly focused on topics related to Industry 4.0, such as **artificial intelligence** through self-learning **neural networks** and the networking of production processes and chains.

2.1 Self-Learning Textile Machine Processes

The research on neural networks and **fuzzy logic** for the optimization of textile machines up to self-learning control processes at ITA has already started in the 1990s (Veit 2012, p. 9 ff). Nowadays, neural networks and **algorithms** are mainly used for the self-optimization of control processes to increase quality or productivity of textile machines in order to keep the production in the textile industry in high-wage countries. An important field of application is the reduction of manual production effort. Especially for the process set-up in textile production a high manual labor demand is necessary (Auerbach et al. 2011, p. 747 ff). In addition, finding the right process parameters for the article to be produced is a considerable time factor. Today, even despite process databases, this is often done according to the “trial & error” principle and the experience of the machine operator is necessary. Therefore, ITA is continuously researching a model-based self-optimization of the weaving machine. This has already made it possible to reduce scrap production and to determine automated trial plans. Additional sensor technology provides the machine operator with information on quality criteria, yarn tension forces and set-up aids (Fig. 2). In other projects, such as the “RegelTuft” project funded by the AiF, these approaches have also been transferred to the tufting process and other surface technologies such as flat knitting.

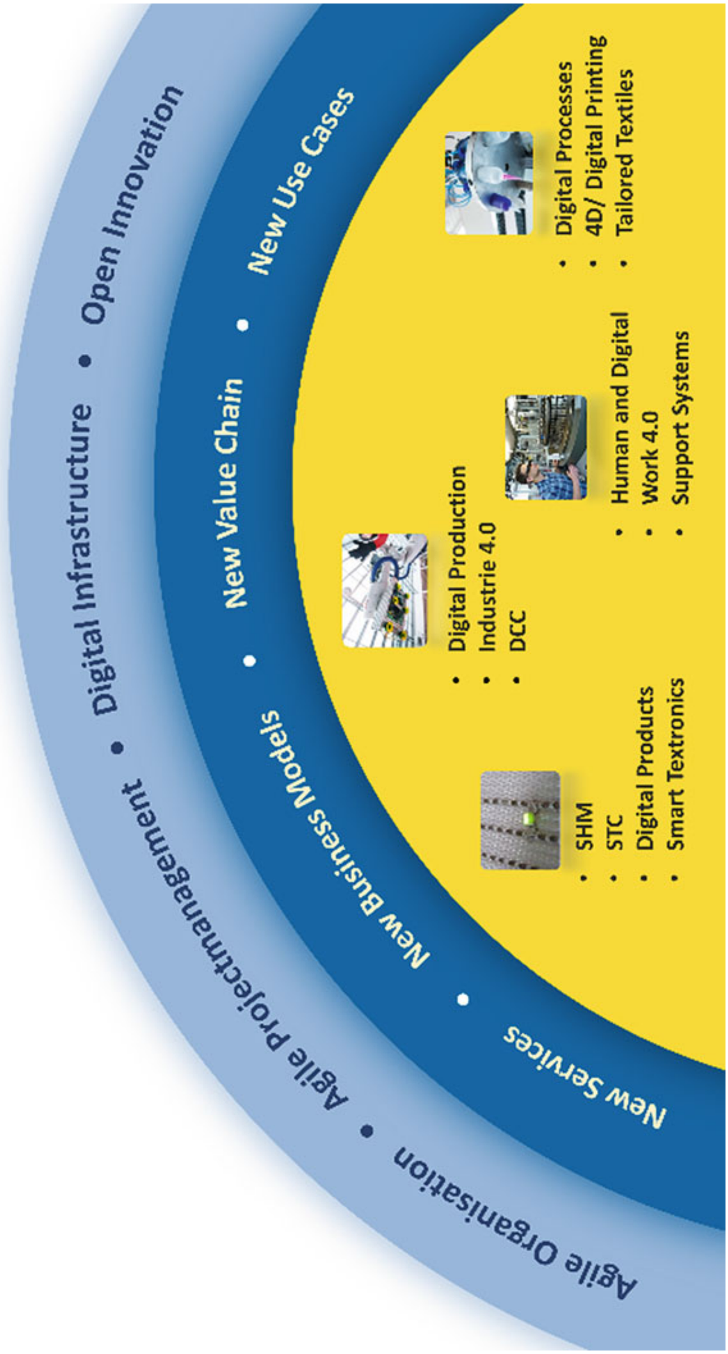


Fig. 1 Digitization at ITA

Fig. 2 Exploring digitization of the warping process

The setting and operation of nonwovens lines is also mainly based on experience, despite numerous measuring and automation technologies. The rejects produced in the nonwovens industry in Germany, worth around 50 million euros per year, have to be recycled at great expense. In the "Easy Nonwoven 4.0" project, among other things, research is being carried out to measure the subjective **quality** of the nonwoven fabric as perceived by the machine operator. In addition, ITA is developing a cost-benefit-based setting aid and control of machine parameters for nonwoven cards. The setting aid considers not only the qualitative effects of settings but also the economic consequences. This prevents the selection of an operating point at which the system is operated uneconomically in favor of **product quality**. The adjustment aid will be based on a model of the nonwoven line. The product quality will be simulated in advance at different settings and a comprehensive production cost model will be used to calculate the respective production costs. Subsequently, an **optimization algorithm** is used to autonomously determine the operating point at which the desired product quality is produced at the lowest possible cost (Cloppenburg, 2019).

2.2 Quality Enhancement Through Digital Setting Aids and Image Recognition Algorithms

The Speedfactory project, which was funded by the German government within the "Future of Manufacturing" program, focused on innovative products and new production technologies, while at the same time considering consumer needs and **sustainability**. In addition to the further development of flat knitting technology, the ITA conducted research on new sewing automation to produce demand-oriented seam geometries for batch size 1 textiles. For this purpose, with the help of camera technologies and in collaboration with industry partners, image recognition algorithms have been used to customize and control seams (Simonis and Lutz 2017).

In the Korean-German ZIM project "DTP Alarm," an automatic quality assurance system based on a **self-learning defect detection algorithm** was developed for digital printing. Here, the print image was determined directly at the print head for rapid detection of clogged nozzles of the ink print head. The aim was to minimize the waste of cost-intensive ink (Park et al. 2017). In the AiF project "CNCPrecisionSew," near-infrared camera technology was being used in collaboration with the Institute for Image Processing at RWTH Aachen University to further develop the detection of seam geometry and the necessary algorithms behind it.

In the field of woven, braided and non crimp fabric production for technical applications such as fiber composite technology, camera technologies were used in the projects "OnLoom Imaging" and "AutoBraid," among others, to detect defects promptly during the production process and to be able to draw conclusions about possible sources of defects in the production process based on the computer-aided interpretation of data (Saggiomo et al. 2014; Reimer et al. 2016).

2.3 Networked Process Chains

Digitization of fiber and textile production makes it possible for the first time to network the entire process chain in the plants and thus to communicate the individual process steps with each other. One focus of research at ITA is on spinning mills and production plants for carded and "airlay" nonwovens. Here, staple fibers are processed into slivers and nonwovens. The fibers are transported between the individual machines of the large plants by pneumatic fiber transport systems. Despite the low energetic efficiency of pneumatic fiber transport systems, this type of transport is the only option so far (Cloppenburg et al. 2018, p. 71 ff). The transport systems are operated with a volume flow up to 30% higher than necessary, because if the output is too low, the lines or machines could become clogged. In the "Ziel2NRW" project "Effect," a test rig was therefore developed at ITA on which the pneumatic transport of fibers is represented on a pilot scale. In the "DynAir" project, this test rig was used with the help of industry partners to develop an innovative and demand-oriented control of filter systems throughout the entire process chain. The aim of the project was to achieve energy savings of at least 20% on conveying fans and 40% on filter systems by networking the individual process steps and using innovative control.

In the "SmartFactory" project, the aim was to look at the specific requirements of selected approaches to **networked process chains** in weaving plants. Concrete guidelines for the implementation of these approaches in the textile industry have been developed. In particular, the current status in the field of woven home textiles and technical textiles has been analyzed first. Two parallel research objectives were pursued here. The first goal was to develop a "smart" weaving mill through process chain networking. In this smart factory, the weaving machines are upgraded to **cyber-physical systems** and the focus is on automated monitoring of process and product quality. Built-in sensors provide quality-relevant data, and gateways and control systems convert and process this data into information. The second research objective was to develop a guideline for determining the strategy of machine communication to network the textile production chain. The guide is a methodical approach to determine the machine communication strategy and was applied to the case study of woven **textile** production. By applying the guide, it is determined which parameters have to be taken from which production process. In addition, it is determined which previous and subsequent processes should be changed to ensure and optimize the overall product and process quality.

2.4 Cross-Company Networking of Production Processes

Few companies in the textile industry in Germany still manufacture on a fully integrated basis. Instead, the textile production chain is fragmented among many small and medium-sized companies that specialize in individual process stages. For this reason, the AiF project "DigiTextil" at ITA and FIR, for example, is investigating which conditions must apply for cross-company networking of production processes, so that a strategic alliance of many small companies could map an integrated production chain in staple fiber and nonwoven production. In addition to the further development of production and automation processes, research has been conducted in the necessary networking of machines within the companies and on encryption technologies for secure data transmission between the companies.

3 People in the Digital Transformation Process in Companies in the Textile Industry

In addition to the obvious process technology opportunities and automation potential, a key research focus is the transformation of companies at the organizational level through the transformation of business models, sales structures, and the

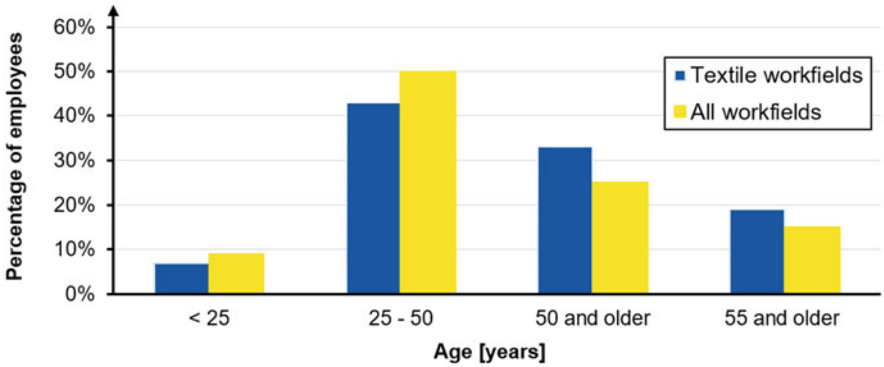


Fig. 3 Age structure in the textile industry (Illustration according to Löhrer et al. 2018)

workplace, both behind the desk and on the production floor (Schreiber and Felk 2017, p. 436). The 2018 year of science also focused on the work of the future. The most central component of **Work 4.0** in the future is the human being itself and the challenges that arise in the work environment for the worker. Triggered by the future topic of Industry 4.0, various technical and guideline committees have been formed around the topic of **Work 4.0**. The VDI expert committee "Arbeitswelt Industrie 4.0" (Working World Industry 4.0) is concerned with work design in companies that is conducive to learning, and the VDI guideline committee 404 on the topic of the digital transformation process of companies is developing a guideline for the implementation of a transformation process in small and medium-sized companies on the way to a networked production and working world (Gronauer et al. 2017).

In Germany, like many other industrialized countries, **demographic change** and an aging workforce additionally represent a factor to be considered when developing work (support) systems for a Work 4.0. The textile industry is affected by demographic change to an above-average extent. The proportion of employees over the age of 45 in Germany's textile companies is above average, and the proportion of employees under the age of 25 is also below average (Löhrer et al. 2018, p. 75) (Fig. 3).

One approach that is currently being investigated is the **participatory design** of digital systems to consider the various needs of the very heterogeneous workforce and to reduce and counter fears (e.g., data protection and privacy) and reservations against new technologies such as Big Data, **artificial intelligence,** and augmented and virtual reality at an early stage.

Software and machine are also growing closer together in the textile industry. The challenges for humans in terms of training, further education and understanding of complex interrelationships are increasing rapidly. Especially in the development process of innovative products and machines, new forms of requirements analysis such as **requirements engineering** or approaches such as design thinking are playing an increasingly important role in the development process. This is due to the increasing importance of software in production processes and machines.

Aspects from the product design are found increasingly often in the development process, since requirements for a process/product arise during the development or even in the use phase, when using open innovation approaches. In the software sector, therefore, creativity techniques and design methods have been increasingly used in the development process for years. In textile machinery construction, this approach is not yet used. Thus, proven design methods based on the waterfall principle, such as VDI2221, continue to be used in the development process. Agile development methods such as SCRUM are also not yet widespread (Lauenroth et al. 2016, p. 192 ff).

However, it is indisputable that the importance of software will increase even more with Big Data analyses of production data or artificial intelligence of machines and systems. In addition, in the course of stronger competition in the area of digital business models in the textile and clothing industry, companies must prepare themselves to serve these new markets to be able to play a shaping role themselves. Due to the age structure of the textile industry, it is faced with a particularly difficult task, as many employees either have to be trained or retrained at great expense, or young people increasingly have to be sought on the labor market. At this point, however, a rather conservative textile industry competes with innovative and future-oriented industries such as the software and entertainment industry, not only in Germany but worldwide. In addition to new training centers such as the one in Mönchengladbach, the industry is also involved in various research projects at the Institute for Textile Technology at RWTH Aachen University and other textile research institutes in Germany. A selection of these projects related to the topic of Work 4.0, **socio-technical system design** and **demographic employee structure** are presented below.

3.1 WissProKMU

The WissProKMU project was launched as part of the BMBF's "Future of Work: SMEs - Innovative and Social" research program to support SMEs in digitization and to promote not only their technical innovation potential but also their social innovation potential. The project is based at the Institute of Sociology at RWTH Aachen University and the ITA. The aim of the project is an integrative scientific utilization of results and analysis of potential in business and science. For this purpose, the project groups of the research focus are networked with each other and the comprehensive transfer of results for small and medium-sized enterprises is supported. The procedures, methods, and results of all the projects in the network are analyzed and brought together within the framework of an accompanying scientific study. Through events and public relations work, the exchange between science and industry is specifically promoted and synergies for cross-project cooperation are created. The transfer of results involves networking the players in the collaborative projects with interested parties from business, science, the media, and society. In addition, the research results are made accessible to a broad public and thus also to small and medium-sized enterprises in particular.

3.2 WisoTex

The aim of the DFG-funded interdisciplinary research project Wisotex 4.0 at RWTH Aachen University is to create recommendations for action for companies and associations based on participatively determined future scenarios. The project takes an **interdisciplinary approach** (from economic geography, textile technology or engineering, economic, social, and technological history) to investigate which spatial implications could be associated with Industry 4.0 in the future (Marshall et al. 2018, pp. 115–116).

Various research questions are answered in the field of textile engineering. Among others, the fields of innovation associated with the paradigm shift 4.0 are analyzed. From a company perspective, it will be determined which 4.0 innovations increase technology efficiency and which organizational and training structures are advantageous for the efficient use of Industry 4.0 solutions.

For the project, current and historical (company) data will be combined with 50 expert interviews and 10 in-depth company reviews from the three economic areas under consideration: Aachen/Lower Rhine, Münsterland/East Westphalia-Lippe and Southwest Saxony/Hof-Münchberg. The future scenarios are developed from a workshop together with associations, entrepreneurs, and researchers. The questions of what need there is for the implementation of digitization concepts, what benefits are anticipated and what requirements are placed on “good implementation” will also be examined.

3.3 SozioTex

Assistance systems as technical tools to support personnel in certain situations and actions are considered one of the components of the production workplace of the future (Jacobs et al.). The assisted learning of work processes by these systems is an important component to meet the challenges of the above-described demographic change in the textile industry. The junior research group “SozioTex - New Socio-Technical Systems in the Textile Industry,” funded by the Federal Ministry of Education and Research for five years is concerned with the development of **socio-technical assistance systems** for heterogeneously growing workforces. The **inter-disciplinary** group of young researchers, consisting of sociologists, educational scientists, and engineers, has started 2014 against the background of the research area of textile production at the Institute of Textile Technology (ITA) in cooperation with the Institute of Sociology at RWTH Aachen University (Altepost et al. 2017, p. 153 ff) (Fig. 4).

Fig. 4 User tests of the assistance system on a weaving machine

The research team developed and implemented **participatory methods** to enable an intelligent human-machine interface in the form of an assistance system using the example of weaving machine operation. For this purpose, the system was implemented in reference weaving mills (Fohn and Altepost 2018, p. 371).

The current research focus is on the participatory introduction of digital technologies in weaving mills to enable the subsequent use of newly introduced technologies in the company with the early involvement of the entire workforce. In this context, the **human-machine interaction** between machine operator, weaving machine and assistance system is understood as socio-technical systems in which technologies and people are equally involved. Technographic analyses describe in detail the transformation of the assistance system during the whole development time together with the included components. By means of technographic use case analyses, design consequences are derived, and the technical system is adapted in iterative loops to the respective target groups: shift supervisors, managers, machine operators and trainees. The SozioTex **assistance system** (Fig. 5) consists of the components: Production planning tool, workflow-controlled work assistance, learning tool for trainees, operating data visualization for shift managers and a knowledge management tool. For example, a not-so-experienced machine operator is guided through each step of a warp beam change, and a learning tool built into the system helps the operator through each step as needed and provides background material to enhance knowledge. As the operator's experience increases, they can skip over

Fig. 5 Graphical user interface of the SozioTex assistance system

individual explanations and instructions for action. In addition, an **augmented reality** assistant is available for operators to insert and detect a weft break. The production planning tool provides a three-dimensional floor plan of the weaving hall as well as real-time data for the shift supervisor to schedule machine operators according to demand. Initial studies show that by using the digital assistant and a **blended learning** concept, a reduction in stretching and the duration of warp beam changes can be achieved. Further management and efficiency of production are increased.

3.4 Competence Center 4.0: Textile Networked

The Aachen showcase of the BMWi Mittelstand 4.0 competence center for textiles is also concerned with people at the center of the digital transformation process. In contrast to the previous projects, the focus of the competence center is less on research and more on the transfer of digital solutions to textile production. The concept is structured with the five tools of informing, demonstrating, qualifying, designing, and implementing and is available free of charge to small and medium-sized enterprises (Kemper and Merker 2018, p. 65).

Under the main topic "new social infrastructures of work, qualification, and lifelong learning," the ITA presents showcases around the topic of Work 4.0. With the aim of avoiding errors and minimizing downtimes, assistance systems, for example through pick-by-light technologies, but also head-mounted displays (Fig. 6) are used in the textile production environment. As a second topic area,

Fig. 6 Pick-by-light technology in the textile production environment

solutions for individual competence development and digitally supported learning are demonstrated, for example, through the use of extended realtime. Another aspect is workplace individualization and optimization of processes with human involvement. By improving **ergonomics** and **workplace design**, employee satisfaction is increased.

The Mittelstand 4.0 competence center Textil vernetzt is part of Mittelstand-Digital. With Mittelstand-Digital, the German Federal Ministry for Economic Affairs and Energy supports digitization in small and medium-sized enterprises and the skilled trades. More than twenty Mittelstand 4.0 competence centers now assist entrepreneurs throughout Germany with digitization, networking, and the introduction of Industry 4.0 applications.

3.5 Digital Capability Center

Through a cooperation between McKinsey, PTC and ITA, the Digital Capability Center was created in Aachen in collaboration with various specialized service companies for the textile industry. Intensive collaboration with the company Jacob Müller has resulted in a networked ribbon weaving mill as the first learning and demonstration factory for the digital transformation process in the textile industry (Fig. 7). In addition to the networking of the individual process steps fiber preparation, tape weaving, fusing, printing, testing and finishing, assistance systems, ERP and MES systems as well as logistics and production planning tools are also demonstrated in workshops and training courses.

Fig. 7 Digitized narrow weaving process chain at the Digital Capability Center in Aachen, Germany

3.6 Further Research Projects and Clusters

The study “Strick 4.0” (Knitting 4.0) on the future of the knitting cluster in Baden-Württemberg (Artschwager et al. 2017), which was carried out by the German Institute for Textile and Fiber Research (DITF) in Denkendorf, also sees a changing working environment in German knitting mills and potentials through digital technologies such as **virtual reality** in the training and further education of the workforce. The same conclusion is reached by the project Arbeitswelt 4.0 (Working World 4.0), which is funded as part of the futureTEX initiative and is a collaboration between the Saxon Textile Research Institute (STFI) in Chemnitz and the Otto von Güricke University (OVGU) in Magdeburg, which have set this out in a TourAtlas Arbeitswelt 4.0 (Working World 4.0) (Schmicker 2017) with the three stages of finding, retaining and qualifying employees.

4 Outlook and Conclusion

A wide variety of efforts are underway in the textile industry to drive digitization and the transformation process toward a production and working world of the future. In the next few years, research will focus mainly on the utilization and exploitation of the data obtained from networking. Topics such as **artificial intelligence** in the form of **self-learning algorithms** and autonomous machine processes will determine the research landscape in textile technology in the area of the ongoing digitalization of production processes. Interdisciplinary teams of engineers, sociologists, psychologists, doctors and designers will be needed to process this data for the machine operator and to present it in a user-related manner, since systems in the workplace of the future will no longer be understood as purely technical, but as **socio-technical systems**.

References

Altepost A, Löhrer M, Ziesen N, Saggiomo M, Strüver N, Houben D, Gloy YS (2017) Sociotechnical systems in the textile industry. i-com 16(2):153–164

Artschwager A, Tilebein M, Rieder O et al (2017) Studie Strick 4.0 - Textil digital Die Zukunft der textilen Produktion am Beispiel des Strick-Clusters Baden-Württemberg. Deutsche Institute für Textil- und Faserforschung Denkendorf

Auerbach T, Bauhoff F, Beckers M, Behnen D, Brecher C, Brosze T et al (2011) Selbstoptimierende Produktionssysteme. In: Integrative Produktionstechnik für Hochlohnländer. Springer, Berlin, pp 747–1057

Cloppenburg F (2019) Wirtschaftliche und technische Modellierung und Selbstoptimierung von Vliesstoffkrempeln. Shaker, Düren

Cloppenburg F, Kemper M, Rezaey A, Gries T (2018) Industrie 4.0 und Digitalisierung in der Textilbranche, VDI-Z 160, Nr. 4- April, S. 71–74

Fohn A, Altepost A (2018) Der Beitrag technografischer Analysen zur partizipativen sozio-technischen Assistenzsystemgestaltung für eine Weberei 4.0. Technische Unterstützungssysteme, die die Menschen wirklich wollen, p 371

Gronauer B et al (2017) VDI-Statusreport “Digitaler Transformationsprozess in Unternehmen”. Verein Deutscher Ingenieure

Jacobs JC, Kagermann H, Spath D (Hrsg.) (2017) Arbeit in der digitalen Transformation, Jacobs Foundation und acatech - Deutsche Akademie der Technikwissenschaften

Kemper M, Merker A (2018) Textil vernetzt: a window into new digital production, Book of Abstracts / Aachen-Dresden-Denkendorf International Textile Conference, Aachen 29. - 30. November, 2018. - Aachen: DWI - Leibnitz-Institut für interaktive Materialien, 2018, S. 65

Lauenroth K, Schreiber F, Schreiber F (2016) Maschinen- und Anlagenbau im digitalen Zeitalter - Requirements Engineering als systematische Gestaltungskompetenz für die Fertigungsindustrie - Industrie 4.0, Beuth Verlag, (ISBN 978-3-410-25932-9)

Löhrer M, Lemm J, Kerpen D, Saggiomo M, Gloy YS (2018) Soziotechnische Assistenzsysteme für die Produktionsarbeit in der Textilbranche. In: Zukunft der Arbeit–Eine praxisnahe Betrachtung. Springer Vieweg, Berlin, pp 73–85

Marschall P, Jordan J, Peters R (2018) Spatial implications of Industry 4.0: the case of three traditional German textile regions In: Braun, Boris; Dannenberg, Peter; Fuchs, Martina; Revilla Diez, Javier (Eds.): Abstracts/Dynamics in an Unequal World: 5th Global Conference on Economic Geography 2018 (GCEG 2018), 24 – 28 July 2018, Cologne. - Cologne: GCEG 2018, 2018, S. 115-116, URL: https://www.gceg2018.com/fileadmin/GCEG2018/PDF/GCEG_BoA_final.pdf (ISBN 978-3-00-060351-8)

Park H, Schreiber F, Shin K, Min M, Gloy YS (2017) Smarte Strukturen und Systeme: Tagungsband des 4SMARTS-Symposiums, 21.-22. Juni 2017, Braunschweig / Martin Wiedemann, Tobias Melz (Hrsg.), Herzogenrath / Shaker

Reimer V, Alptekin S, Gries T (2016) Investigation of the relations between the parameters in the radial braiding process. In: Recent developments in braiding and narrow weaving. Springer, Cham, pp 111–120

Saggiomo M, Holtermann T, Schneider D, Gloy YS, Gries T (2014) Integratives Messsystem zur Fehlererkennung in der textilen Flächenherstellung (OnLoom Imaging). Frey, Georg; Schumacher, Walter; Verl, Alexander (Hrsg.): SPS IPC Drives: Elektrische Automatisierung; Systeme und Komponenten, (25)

Schmicker S (2017) Touratlas Arbeitswelten 4.0 – Abschlussdokumentation Basisvorhaben Arbeitswelten 4.0. Institut für Arbeitswissenschaft der Otto-von-Güricke-Universität, Magdeburg

Schreiber F, Felk K (2017) Status-Quo und Wertschöpfungsperspektiven digitaler Geschäftsmodelle in der Textilbranche. In: Digitale Transformation von Geschäftsmodellen. Springer Gabler, Wiesbaden, pp 433–450

Simonis K, Lutz V (2017) Schlussbericht Verbundvorhaben SPEEDFACTORY-Autonomik für die Sportartikelindustrie: Titel des Teilvorhabens: Automatisierte Fertigung dreidimensionaler Gestricke sowie die sensorische Überwachung und die automatisierte Handhabung biegeschlaffer Textilien während des Fügeprozesses: Laufzeit des Vorhabens: 01.10. 2013 bis 31.12. 2016. Institut für Textiltechnik der RWTH Aachen University

Veit D (2012) Neural networks and their application to textile technology. In: Simulation in textile technology, pp 9–71

Lightweight Design 4.0: The Fundamentals and Potential of Structural Health Monitoring

Kai-Uwe Schröder and Andreas Janetzko-Preisler

Contents

1 The Evolution of Lightweight Design

Lightweight design refers to a philosophy of component design which objective is to minimize the mass of a structural component while meeting its requirements on load-bearing capacity and stiffness. There are several reasons for lightweight design. First, a minimum weight target is a necessary requirement to fulfill a certain mission. That is particularly in space engineering the outmost reason for lightweight design. That kind of lightweight design is denoted as **function lightweight design**.

However, if the reduction of operational costs is the main driving reason for minimizing weight as in aviation, where to reduce fuel consumption, this is referred to as **economic lightweight design** (Wiedemann 2007).

During the last decades, lightweight design has undergone rapid evolution. The first stage in the evolution of lightweight design was the provision of design and calculation methods of thin-walled structures including non-linear structural behavior. With the help of that theories, lightweight structures became analytically

K.-U. Schröder · A. Janetzko-Preisler (✉)
Institute of Structural Mechanics and Lightweight Design, RWTH Aachen University, Aachen, Germany
e-mail: andreas.preisler@rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_27

calculable. They could thus be designed with regard to their strength and stiffness and optimized with regard to their weight, taking safety factors into account.

The second stage in the evolution of lightweight design is due to the rapid advancements in computer technology. Powerful and affordable computers enabled the widespread use of computer-aided numerical calculation methods such as the finite element method (FEM). Because of the performance of modern computer systems in combination with commercial FEM programs, even complex structures that are not amenable to analytical calculation can nowadays be computed with sufficient accuracy and optimized with regard to their weight.

The third evolutionary stage of lightweight design is closely linked to the emergence of so-called **lightweight material design** (Klein 2008). Here, different materials are used together in a multi-material structure, whereby each material is optimally used according to its characteristics. Furthermore, **function-integrated lightweight design** has increasingly emerged. Here, additional functions are integrated into the load-bearing structure, whereby additional components can be omitted. The primary structure then fulfils additional functions besides load transfer, such as heat conduction or data transfer.

Despite the aforementioned developments, the potential of lightweight design has not yet been exhausted. Even modern computational methods depend on the accuracy of the input data. However, these are subject to **uncertainties**. They already begin with the definition of the loads. Exact operating load cases over the entire service life of a structure are not known before operation (Boller 2013). The loads are therefore estimated taking into account conservative safety factors. Further uncertainties exist in the material properties. They are determined in the laboratory on coupon samples. However, these results are provided with a significant scatter. Further, production-related deviations can occur in real, complex structures. Another source of uncertainty is the fatigue behavior of the structure. Even when using metallic materials, whose behavior is generally well known, the time of crack initiation and the speed of the subsequent crack propagation can only be determined approximately. Lastly, the delaminations that occur in fiber reinforced plastics (FRP) should be mentioned. In this case, the individual layers inside the fiber composite are separated from each other over a considerable area, which means that the shear bond of the layers is not given, and the stiffness is significantly reduced. This has an effect especially in the case of compressive or bending loads. Delamination can be caused by manufacturing defects, fatigue and impact damage. As inner damage, however, they are hardly or not at all visible to the bare eye.

The design against fatigue during operation of a structure can follow different safety concepts. For example, it is possible to aim for zero damage for the entire service life (**safe life design**), but this can only be achieved by using sufficiently conservative safety factors. Therefore, the concept of a fail-safe structure is to be preferred in lightweight applications (**fail-safe design**), for example by ensuring the integrity of the structure in the event of failure through the introduction of structural redundancies. The concept of **damage-tolerant design** can be seen as a further evolution of fail-safe design. Here, regular inspections and repairs prevent the damage from growing to a critical size. This concept has become widely accepted,

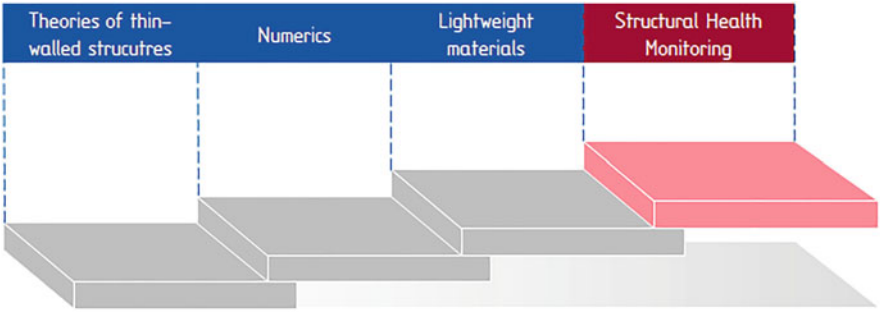


Fig. 1 The four evolutionary steps of lightweight design

particularly in aviation, due to its weight efficiency for metal structures. With the emergence of composite structures, however, the particular problem of **barely visible impact damages (BVID)** arose. As already expressed in the name, this damage is difficult to detect, but can have serious impacts in fiber composite structures in contrast to metallic structures. This consequently means that the presence of undetected damage in the fiber composite structure cannot be excluded and thus the design must be carried out for the damaged structure. Due to the excellent fatigue behavior of fiber composite structures, however, this design can be carried out in such a way that further growth of this damage is prevented (**safe life—flaw tolerant design**). In the design of lightweight structures, weight and inspection effort are thus in direct interaction with each other.

Structural Health Monitoring (SHM) is one way of mitigating these uncertainties and extending inspection intervals. A real-time, reliable SHM system enables a component to provide information about its structural health at any time during operation. And the more accurate and reliable the SHM system is, the more conservatisms in safety factors can be reduced, which directly translates into a lighter structure. Therefore, the SHM, as shown in Fig. 1, is considered the fourth evolutionary step of lightweight design—or Lightweight Design 4.0.

2 Fundamentals of Structural Health Monitorings

Structural Health Monitoring (SHM) denotes the automated monitoring of structural integrity during the operation of a system. The requirements, expectations and boundary conditions of a SHM vary significantly depending on the area of application. Therefore, the systems are typically categorized according to the significance of their results (Rytter 1993):

- Level 1 (detection) provides information on whether damage is present.
- Level 2 (localization) determines the approximate location of the detected damage.

- Level 3 (classification) additionally determines the size and severity of the damage.
- Level 4 (consequence) assesses the impact of the detected damage on the safety of the structure.

As the level increases, so does the complexity of the monitoring system required. A level 1 system can be relatively simple, since it only has to detect the mere presence of damage from the change in structural behavior. Level 4, on the other hand, refers to a very complex system that must be able to make autonomous decisions. It can therefore assess whether the detected damage is critical and whether an immediate shutdown of the machine or the entire system is necessary. In the case of non-critical damage, it can also make predictions about the residual strength of the structure and provide an assessment of the remaining service life. This information is necessary to plan the maintenance of the structure at an early stage (and thus in an economically sensible way). It should be noted that the ability of the SHM system to make decisions goes beyond the initial meaning of the SHM. Therefore, such a system is also referred to as **Structural Health Control** (SHC) (Viechtbauer et al. 2012).

Regular inspections can be regarded as a preliminary stage of SHM, why it is often referred to as scheduled SHM. To detect inner damage, particularly in composite structures, **non-destructive testing (NDT)** methods such as ultrasound or thermography are used (Konstantopoulos et al. 2014). However, many of these methods cannot be applied to the permanent use of a structure in operation due to their complexity or the heavy and expensive measurement and evaluation devices.

In the last decades, a variety of sensor technologies and measurement concepts have been tested in the context of SHM. Due to the multiple use in structural tests, strain sensors (strain gauges, piezoelectric sensors, and fiber optic sensors) are considered reliable and generally accepted. In the context of damage detection, they measure the change in the strain pattern caused by the damage. In the case of structures that are exposed to oscillations and vibrations, damage can be detected with the help of acceleration sensors. This takes advantage of the fact that in the event of damage, the stiffness of the structure and thus the measured vibration response of the system changes. Since the global vibration behavior is only conditionally sensitive to a local change in stiffness (as a result of damage), other approaches consider a significantly higher frequency range. One example of this is the use of so-called directional elastic waves. With the help of piezoelectric actuators, elastic waves of a narrow frequency range are introduced into the structure and then measured. Classically, a distinction is made between two variants: (1) the actuator acts simultaneously as a receiver and measures the echo of the transmitted waves or (2) the transmitted signal is received at one or more other positions. In case of damage, several characteristics of the measured signal change: The amplitude of the main frequency decreases, a shift of the main frequency may be present and other frequencies in the oscillation are measured. However, the direct correlation of these changes with the extent of damage is very complex and not straightforward. Such methods are therefore mostly used for a localization of a damage (i.e., for a level

2 system). For this, the signal is evaluated at several receivers to determine along which paths there is a signal deviation.

Despite the extensive potential of modern sensor technology, most existing SHM systems are categorized in the first two levels (detection and localization) (Boller et al. 2009). For higher levels of monitoring, a correlation between the measured signals and the size and severity of damage is needed. In **data-driven approaches**, correlation is determined empirically based on a large amount of data. Statistical methods and machine learning approaches, such as pattern recognition, are used (Farrar and Worden 2007). **Physics-based approaches** (also called model-based approaches) establish the correlation between the measured signals and the extent of damage directly through physical principles using analytical and numerical models.

The models of the physics-based approaches can be used to optimize the measurement concept and sensor positions already during the development of the SHM system. The simulation of the damaged structure is used to determine the influence of damage on the mechanical behavior of the component under consideration. This enables the identification of so-called **Structural Damage Indicators (SDI)** (Huang et al. 2018). SDIs characterize measured variables that have a directly quantifiable correlation to the damage. They are derived from mechanically quantifiable effects that only occur due to the presence of damage. Thus, the concept of structural damage indicators offers many advantages compared to a pure measurement of parameters. For example, simple damage detection is made possible by determining the occurrence of such a damage effect in the sense of a true/false statement (Preisler et al. 2018). Only when damage is detected in this way is the measured value processed or recorded. In addition, the damage indicators also enable a direct assessment of the damage by correlating their measured level with a threshold value that resulted in the simulations at a critical damage level. An estimation of the residual strength is also possible in this way. Since the concept of structural damage indicators is based on the evaluation of damage impacts determined with the help of mechanical models, an initial measurement to determine a reference threshold is not necessary. Thus, a SHM system based on structural damage indicators can also be used for quality assurance already in production, as deviations from the ideal state specified in the model can be detected at an early phase.

The **absolute SDI**

$$\iota_{abs} = X(x,t) - X_0(x) = \Delta X(x,t) \tag{1}$$

is defined as the deviation in the structural response X at position x at time t from the reference value (expected measured value of the damage indicator) X_0 (Preisler et al. 2018). While Eq. (1) describes the absolute deviation from the expected measured value, a greater significance with regard to the sensitivity can be obtained by considering a **relative SDI**

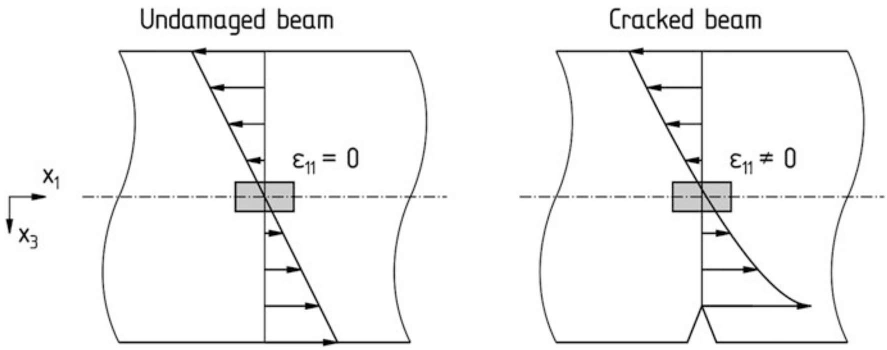


Fig. 2 Stress distribution of an undamaged cross-section of a beam under pure bending (left) and after crack initiation (right) (Preisler et al. 2018)

$$\iota_{rel} = \frac{X(x,t) - X_0(x)}{X_0(x)} = \frac{\Delta X(x,t)}{X_0(x)}. \tag{2}$$

Eq. (2) thus describes the relative deviation from the expected measured value of the damage indicator. It can be seen that the use of a reference value of zero (or close to zero) is specific. In this case, the relative damage indicator grows beyond all limits. In practical application, this means that even a relatively small deviation ΔX results in a significant signal. In the case of a large reference value, on the other hand, a small deviation would be difficult to detect due to the expected measurement noise. Choosing a reference value of zero thus significantly improves the sensitivity of the relative damage indicator. The simplest example to demonstrate this relationship is the observation of a simply supported beam under pure bending load (Preisler et al. 2018).

Figure 2 shows the stress distribution of an undamaged cross-section of a beam (left) and of a crack one (right). In the initial state, the neutral axis, which is defined as the zero stress point, is at the centre of gravity. A strain sensor (e.g., a strain gauge) placed at the centre of gravity thus ideally measures zero in the undamaged beam. After crack initiation, the centre of gravity of the cross-section shifts and stress concentration occur at the crack tip. Due to the changed stress distribution, the strain sensor now measures a deviation from zero. The amount of this deviation depends on the size of the damage.

Figure 3 shows the reference value's influence on the sensitivity of the damage indicator. Here, the relative damage indicator (Eq. (2)) is shown for a cracked beam. All areas in which the relative deviation is less than 100% are marked in grey. Red represents a deviation larger than 300%. Near the crack tip, as Fig. 2 shows, there is a high absolute change in strain, but due to the high reference value, the relative change is comparatively small and decreases rapidly with increasing distance from the cracked cross-section. At the height of the neutral axis with the reference value zero, however, the damage can be detected reliably even at a larger distance from the damaged cross-section.

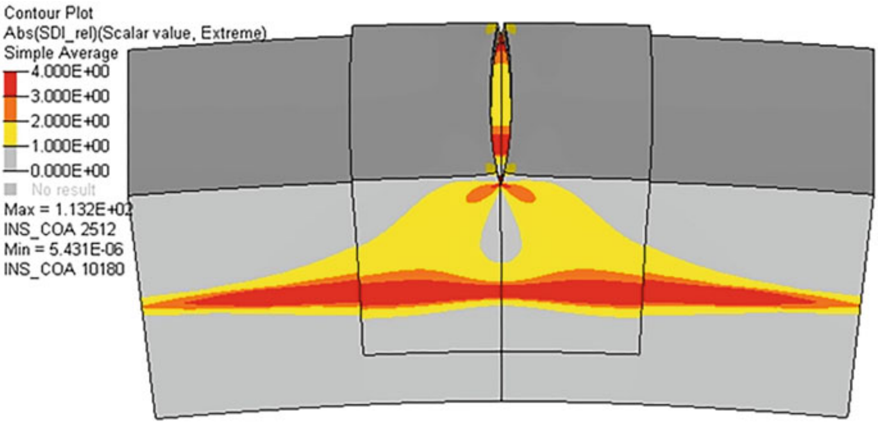


Fig. 3 Relative deviation of longitudinal strain of a cracked beam (cf. Eq. (2)) (Preisler et al. 2018)

3 Demonstrating a SHM System on a Braided Composite Shaft

The implementation of a SHM system offers many options, but its development is also associated with some challenges, mainly related to the integration of all components of the system. To demonstrate this, the development of a SHM system for a braided shaft made of glass-fiber reinforced plastic (GFRP) is discussed below. The textile manufacturing process of the braided shaft offers the option of automated integration of the sensor technology already in the manufacturing process, especially if **fiber optic sensors (FOS)** are used. The braiding of the shaft is done on a soft foam core. In addition, aluminum fittings are woven into both ends to enable a screw joint of the shaft to the test rig.

At this point, the focus of the discussion is on (1) the monitoring concept based on structural damage indicators, (2) the sensor integration in production and (3) the validation of the concept in static tests. Further aspects have been published elsewhere (Preisler 2020).

3.1 Monitoring Concept

For a GFRP shaft, the greatest safety risk comes from a loss of stiffness due to impact damage. The aim of the SHM system is therefore to detect such a loss of stiffness in timely fashion during operation by means of integrated sensor technology. The development of a SHM system starts with the identification of the damage indicator to be monitored. For the shaft, this is tricky because in the case of impact damage, the possible damage location can hardly be localized, since theoretically this can appear on the entire shaft. To identify the damage indicator, the structural behavior of the

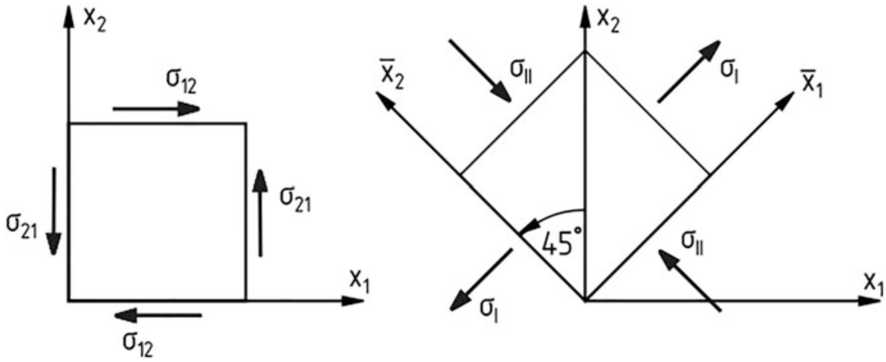
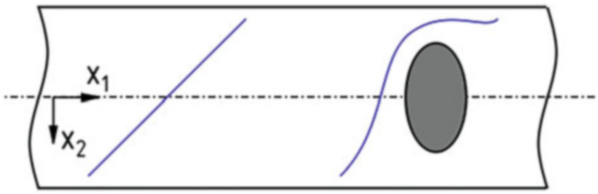


Fig. 4 Plane stresses at pure shear (left) and the corresponding principal stresses with rotated coordinate system (right) (Preisler 2020)

Fig. 5 Direction of principal stresses in an undamaged (left) and damaged (right) shaft (Preisler 2020)



undamaged and the damaged shaft is first examined to be able to describe the load-bearing behavior and the impact of damage. A pure torsional load with clamping on both sides is considered. Due to its geometry, the shaft can be regarded as a thin-walled structure. Therefore, a pure shear load and thus a pure shear deformation is present for the undamaged shaft. Figure 4 shows the plane stress condition prevailing in the shaft in the component coordinate system, in which the x_1 direction corresponds to the component axis. If the reference coordinate system is rotated by 45°, pure shear becomes a pure normal stress state, known as the corresponding principal stress state. This principal stress direction is shown in Fig. 5 as a blue line on the outer contour of the shaft. If there is an impact damage (shown in Fig. 5 (right) as a grey ellipse), the stiffness in this area is reduced and the shear load has to be transferred around the damage. Thus, despite the pure torsion, longitudinal stresses occur in addition to the shear stresses, which changes the corresponding main stress direction at the damage location.

The aforementioned effect can be used as a structural damage indicator in the following way: In the initial state there is pure shear stress. Thus, there are no normal stresses (and thus strains) in the longitudinal direction of the shaft. Only when the shaft is damaged strains appear in the longitudinal direction due to the redistribution of stress. The longitudinal direction of the shaft is therefore referred to as the **zero-strain direction** (Schagerl et al. 2015). Monitoring the strain in this direction thus provides the damage indicator.

Since, as mentioned above, the damage cannot be localized, the shaft must be monitored over its entire area. Fiber optic sensors using **optical frequency domain reflectometry** (Ramakrishnan et al. 2016) are particularly effective for this purpose. It enables a locally resolved strain measurement along the complete FOS. Through the specific placement of the FOS, the entire shaft can be monitored with relatively little effort. An FOS placed in the longitudinal direction of the shaft is thus able to detect damage by measuring a non-zero strain. The x_1 position of the damage can be determined from the location and distribution of the measured zero deviation. A level 2 SHM is thus enabled with an FOS. The amount of zero deviation again depends on the size of the damage and the distance of the damage to the FOS in the circumferential direction of the shaft. If two FOS are used, a level 4 system can be developed from them with the right evaluation algorithm.

3.2 Experimental Setup

To validate the measurement concept, four FOS were integrated as longitudinal fibers in a ±45° braid of glass fibers. The resulting fabric with integrated sensors can be seen in Fig. 6. The layer structure of the shaft consists of three layers of the aforementioned braiding. The sensors are located in the middle layer and thus in the middle braid. ROHACELL 110 IG-F foam was used as the core. Due to its low stiffness, the core can be neglected as a load-bearing element. The shaft is therefore a thin-walled tube with a length of 650 mm, an inner diameter of 65 mm and a wall thickness of 1.5 mm. In addition, aluminum sleeves were inserted at both ends of the shaft. They allow a screw joint to the adapter flanges and thus to the rig.

The finished fabric, including aluminum sleeves and foam core, was then infused with an epoxy matrix and cured in a specially constructed aluminum mold using the

Fig. 6 Glass fiber fabric with integrated FOS as longitudinal fiber (red) (Preisler 2020)

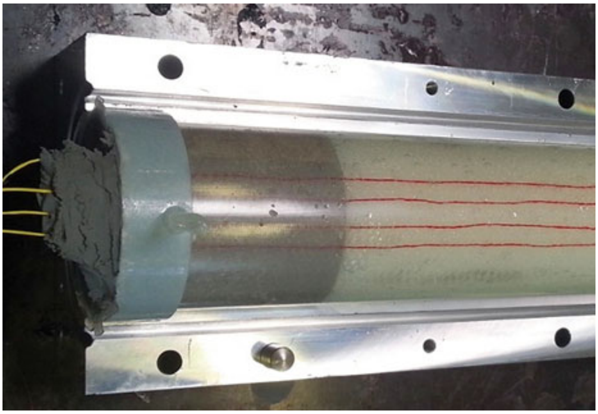


Fig. 7 GFRP shaft after RTM process with aluminum mold (Preisler 2020)

Fig. 8 Inducing damage within a drop tower test (Preisler 2020)

resin transfer molding (RTM) process. By integrating the aluminum sleeves into the RTM process, a strong bonding between the GRP shaft and the aluminum sleeve was achieved. During the RTM process, the FOS were protected with a silicone compound and guided out of the mold (cf. Fig. 7).

Impact damage was then applied to the GFRP shaft in drop tower tests. A square profile with a mass of 3.26 kg was used as the impact device, which impacted the shaft from a height of 0.32 m. The impact energy was thus 10 J. The impact of the square profile was at positions 200 mm and 440 mm in the longitudinal direction. Figure 8 shows the drop tower test with the square profile used.

After being damaged in the drop tower test, the shaft was clamped in a static torsion test. In the torsion test, the load was increased in steps of 50 Nm up to 300 Nm. At each step, the measurement of the FOS was recorded. The complete set-up of the torsion test with the GFRP shaft and the adapter flanges is shown in Fig. 9.

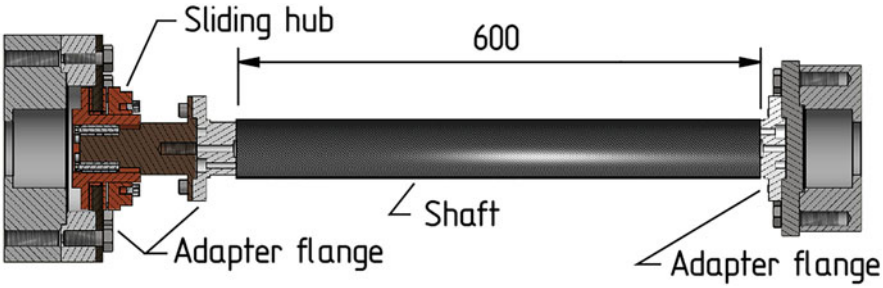


Fig. 9 Schematic illustration of the torsion test (Preisler 2020)

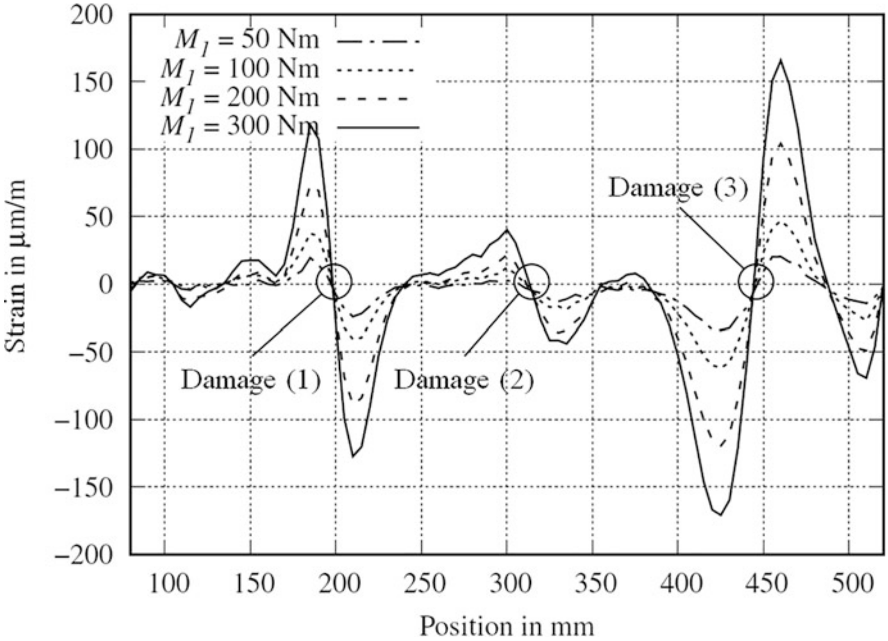


Fig. 10 FOS measured strains of the torsion test (Preisler 2020)

3.3 Validation of the Monitoring Concept

In the static torsion test, the strain along the FOS was recorded for each load step. The measured strain of an FOS is shown for selected load steps in Fig. 10. The range from 80 mm to 520 mm is considered here; thus, the aluminum sleeves fall out of the result shown. Due to the choice of damage indicator, the measured strain in undamaged areas is zero. Smaller deviations are possible due to slight sensor misplacements and measurement noise. Furthermore, a dependence of the measured value on the applied load can be seen in the damaged area. To evaluate the damage, load monitoring is required in addition to the SHM, but this can be much simpler than the SHM system.

In Fig. 10, three characteristic signals can be identified: (1) at a position of approx. 200 mm, (2) at approx. 320 mm and (3) at approx. 440 mm. In all cases there are larger deviations in the measured value from zero before and after the indicated position, while at the indicated position there is a zero pass and thus a change of sign. If the damage is below the FOS, then there is initially a compressive strain that changes to a tensile strain. If the damage is above the FOS, this behavior is flipped.

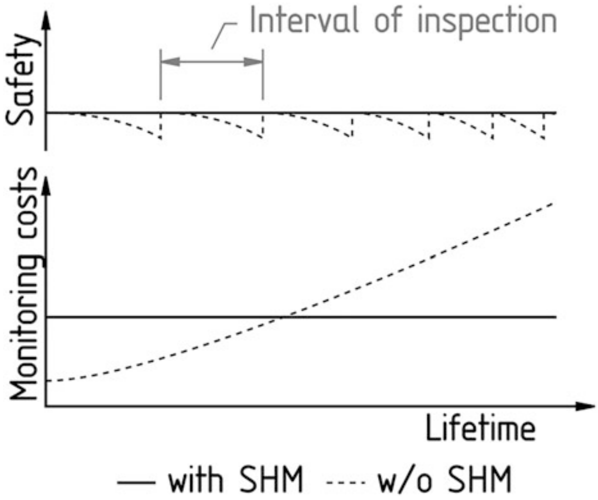
The amount of the measured deviation also depends on the size of the damage and the distance of the damage to the FOS in the circumferential direction. The two damages at 200 mm and 440 mm were produced with the same impact energy and are therefore approximately equivalent. However, damage (1) is positioned more distant from the FOS used in the circumferential direction, therefore the measured deviation decreases compared to damage (3). Damage (2) was not generated in the drop tower test. A conventional strain gauge was attached at this point to validate the FOS measurement. For better adhesion of the strain gauge, the surface was grinded before the sensor application. This is therefore an unintentional surface damage of minor extent. The fact that this surface damage is clearly detectable from the measured values demonstrates the sensitivity of the monitoring concept.

The measurement results show that the use of the zero strain direction as a damage indicator for a braided GFRP shaft under a torque has a very high sensitivity to damage. Both impact damage and also an unintentional surface damage were reliably detected by the FOS. The damage indicator is thus suitable as a basis for the design of a level 4 SHM system that can assess damage in terms of its safety hazard. In combination with the use of FOS with optical frequency domain reflectometry, a very efficient SHM system is also created, since on the one hand a continuous strain measurement along the complete FOS is enabled and on the other hand the integration of the FOS can be carried out directly in the braiding process. This provides the potential of automated sensor placement.

4 Potential of Structural Health Monitoring

The benefits of a SHM system go far beyond the mere detection of damage to a structure in operation. A structure equipped with a SHM system becomes smart. It can detect damage autonomously and, in the case of a level 4 system, assess it directly. The smart structure is thus able to decide whether safe operation is still ensured and can inform the operator in real time about the change in condition. The operator of the structure is enabled to plan any maintenance at an early stage and to carry it out as required. Inspection intervals can be extended and unnecessary downtime avoided. As shown in Fig. 11, the installation of an SHM system maintains the high level of safety of the structure over its entire service life while keeping monitoring costs constant. While the inspection costs of an ageing structure

Fig. 11 Evolution of safety and monitoring costs over a structure's lifetime with and without SHM



that is not monitored by a SHM system increase steadily due to ever shorter maintenance intervals, the costs for implementing the SHM system are only paid once at the beginning. The additional costs of an SHM system are amortized with increasing service life.

An SHM system enables the more accurate assessment of safety factors. If the system is designed accordingly, the condition of the structure in operation is recorded and documented throughout its service life. It is thus documented at which point in time the condition has noticeably deteriorated and what impact other parameters (e.g., load peaks) have on the structure's condition. This creates an extensive database that can be used for the design of future structures. The uncertainties mentioned at the beginning can thus be reduced. Especially with regard to structural bonding, SHM represents a key technology. In many industries, bonding of the primary structure is avoided because there is insufficient confidence in its long-term behavior. This is why the Joint Committee on Adhesive Bonding Technology named SHM as a key element of the roadmap with the goal of "creating confidence" as early as 2015 (Paul 2015).

But SHM does not only provide operational benefits. The use of an SHM system based on damage indicators also enables intrinsic quality assurance during production. With the appropriate choice of damage indicator, the perfect condition, which can be described with a structural-mechanical model, is used as a reference value. The first measurement of the SHM system describes the deviation from this model and can thus be used to evaluate the manufactured component. In the case of an early integration of the sensor system in the manufacturing process, this also enables process monitoring. In the previously described example of a braided GFRP shaft, the FOS could have already been monitored and evaluated during the RTM process.

SHM generates data on the state of the structure during operation or even during production. On the one hand, this data can be used to optimize manufacturing and design, and on the other hand, it represents a key component for the development of

digital twins. To create a digital representation of the real structure, it is not enough to know the external operating parameters of the structure and to map them onto a virtual model. For a comprehensive digital representation, the current condition of the structure must also be known. This is because any damage, but also ageing, have a significant influence on the load-bearing behavior of the structure and must therefore be mapped in the digital twin.

5 Conclusion and Outlook

Depending on the complexity of the SHM system, Structural Health Monitoring enables the detection of damage of a structure up to the automated assessment of the detected damage in real time during operation. SHM ensures a permanently high reliability of the structure at relatively low monitoring costs. In addition, SHM enables the elimination of uncertainties that impede an optimal design process in terms of lightweight construction. In this context, SHM is therefore also referred to as the fourth development stage of lightweight design—or Lightweight Design 4.0.

The presented concept of structural damage indicators results in an efficient and sensitive SHM system that is able to detect damage to the structure with relatively little sensor effort and to quantify and assess this damage directly via the size of the damage indicator. This concept was successfully demonstrated on the example of a braided GFRP shaft by inserting fiber optic sensors in the zero-strain direction of the shaft. The integration of the FOS into the braiding process also enabled automated sensor placement. All damages were reliably detected in the test.

Structural Health Monitoring on real structures outside laboratories and test facilities, however, is difficult to realize. An engineer who wants to implement these basic ideas in reality faces many challenges. Structural behavior must be understood and modelled in both the undamaged and damaged states. To assess the damage, knowledge about the fatigue behavior of the structure is also necessary. In addition to selecting a suitable damage indicator, a sensor that meets the requirements must also be chosen. The measured values of the sensors should ideally be pre-processed at the structure (edge computing) before relevant information is saved and transmitted in packages. Another relevant topic in this context is the supply of energy to the SHM system. Depending on the area of application, energy harvesting approaches play a role here. The development of a SHM system is therefore an interdisciplinary task. Expertise from many disciplines is required. The detached consideration of partial aspects within the SHM, shown in Fig. 12 (left), and the attempt to connect these partial solutions afterwards has not proven to be effective. To successfully implement a SHM system, a holistic approach is required that addresses all sub-aspects from the outset, takes into account their interactions with each other, pays attention to the compatibility of the individual components and combines them to form a working system (cf. Fig. 12 (right)). The development of a SHM system therefore requires a holistic view in the sense of system engineering.

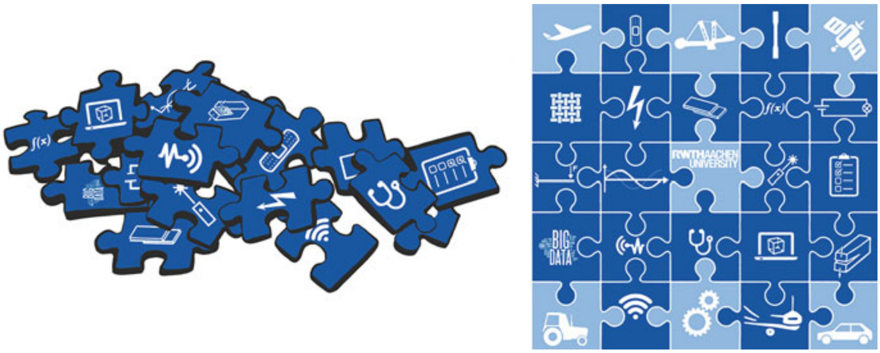


Fig. 12 The holistic approach necessary in the development of a SHM system

Acknowledgement The demonstration of an SHM system on the braided GFRP shaft was carried out in close cooperation with the Institute of Textile Technology, the Institute of Machine Elements and System Development, the Institute of Automotive Engineering and the Institute of Welding and Joining Technology at RWTH Aachen University, the Fraunhofer Institute for Production Technology and Georg Merzenich (enjoy Innovation). The authors would like to take this opportunity to thank all those involved for their fruitful and productive cooperation.

References

Boller C (2013) Structural Health Monitoring – its association and use. In: Ostachowicz W, Güemes A (eds) New trends in Structural Health Monitoring. Springer, Wien Heidelberg New York Dordrecht London

Boller C, Chang FK, Fujino Y (2009) Encyclopedia of Structural Health Monitoring. Wiley, Hoboken

Farrar CR, Worden K (2007) An introduction to Structural Health Monitoring. Philos Trans R Soc A 365:303–315

Huang T, Chaves-Vargas M, Yang J, Schröder K-U (2018) A baseline-free structural damage indicator based on node displacement of structural mode shapes. J Sound Vib 433:366–384

Klein B (2008) Leichtbau-Konstruktion – Berechnungsgrundlagen und Gestaltung. Vieweg +Teubner, Wiesbaden

Konstantopoulos S, Fauster E, Schledjewski R (2014) Monitoring the production of FRP composites: a review of in-line sensing methods. eXPRESS Polym Lett 8:823–840

Paul F (2015) Dem Kleben Vertrauen schenken. Adhäsion Kleben und Dichten 59:20–23

Preisler A (2020) Efficient damage detection and assessment based on structural damage indicators. Dissertation, RWTH Aachen University, Aachen, Germany

Preisler A, Schröder K-U, Schagerl M (2018) Intrinsic damage assessment of beam structures based on structural damage indicators. Am J Eng Res 7:56–70

Ramakrishnan M, Rajan G, Semenova Y, Farrell G (2016) Overview of fiber optic sensor technologies for strain/temperature sensing applications in composite materials. Sensors 16:99

Rytter A (1993) Vibration based inspection of civil engineering structures. Dissertation, Aalborg University, Aalborg, Dänemark

Schagerl M, Viechtbauer C, Schaberger M (2015) Optimal placement of fiber optical sensors along zero-strain trajectories to detect damage in thin-walled structures with highest sensitivity. In: Chang FK (ed) System reliability for verification and implementation. DEStech Publications, Lancaster

Viechtbauer C, Schröder K-U, Schagerl M (2012) Structural Health Control - a comprehensive concept for observation and assessment of damages. In: Scheidel R, Jacoby B (eds) The 13th Mechatronics Forum International Conference Proceedings. Trauner, Linz

Wiedemann J (2007) Leichtbau – Elemente und Konstruktion. Springer, Berlin Heidelberg

Mechanical Engineering and Industry 4.0

Daniel van Geerenstein

Contents

1 Introduction

Mechanical and plant engineering, one of Europe's key industries, has always stood for a constant process of change, adaptation and improvement. Whether with the introduction of the steam engine or the first mechanical loom (First Industrial Revolution), the first electric machines with assembly line production based on the division of labor (Second Industrial Revolution), the use of IT and the connected

D. van Geerenstein (✉)
VDMA, Frankfurt, Germany
e-mail: daniel.vangeerenstein@vdma.org

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_28

automation (Third Industrial Revolution) or now with the intelligent and interconnected digital factory (Fourth Industrial Revolution, Industry 4.0), mechanical engineering has always been an essential driver and addressee of change. By now, mechanical engineering is again at the forefront of a new era in industry. No other sector is currently investing more than mechanical engineering in digital technologies (almost 4% of total turnover).[1]

Accordingly, this results in multi-layered industrial opportunities, such as process optimization in real time, the use of intelligent equipment and the utilization of so-called digital twins.[2] For operators of such plants, the intelligent equipment leads in the best case to production improvement, for example optimization of maintenance and servicing (incl. predictive maintenance). The mechanical engineering industry expects further changes while implementing so-called machine-to-machine communication (M2M communication), in which—in its purest form—products and production machines communicate with each other. Combined with increasing possibilities of data analysis and evaluation (e.g., Big Data analyses), this results in the almost unlimited optimization potential already mentioned: transparent, digital production and logistics processes allow more efficient planning and execution of these processes and will ultimately lead to an intelligent, possibly largely autonomous value chain.

This technological change leads to a high pressure to change in the economy and the employees or trainees employed in this sector: new, innovative approaches emerge, for example in training, where new training content is gradually being included, such as the topics of data security and analysis, information technology based order processing and deadline management, research in clouds and networks, as well as the use of digital learning media and assistance, diagnostic or visualization systems, or new additional qualifications for training courses already in progress.[3] Furthermore, Industry 4.0 also means a new field of training and areas of responsibilities for social science professions such as lawyers, flanked by a continuous duty of further education. Technically coined industrial processes not only pose challenging questions for the law itself, but also for industry and its legal advisors, i.e., lawyers, which will be briefly examined below.

[1] Cf. Ernst & Young, "Industrie 4.0 im deutschen Mittelstand" (2018), p. 9, available at https://www.ey.com/Publication/vwLUAssets/ey-industrie-4-0-im-deutschen-mittelstand-befragungsergebnisse-2018/$FILE/ey-industrie-4-0-im-deutschen-mittelstand-befragungsergebnisse-2018.pdf (as of 26 January 2019).

[2] Cf. joint study by maexpartners and VDMA AG Großanlagenbau, "Potenziale von Industrie 4.0 im Großanlagenbau" (September 2017), p. 4.

[3] Press release of the Federal Ministry for Economic Affairs and Energy (08.06.2018), available at https://www.bmwi.de/Redaktion/DE/Pressemitteilungen/2018/20180608-ausbildung-industrie-4-0-zupacken-statt-zuwarten-in-der-metall-und-elektroindustrie.html (Effective: 04.01.2021).

2 Current Legal Framework and Determining Legal Fields/Issues

The challenges are associated with a certain innovation-hampering legal uncertainty by some companies.[4] The first question to be examined is whether the current legal framework adequately takes the new developments into account. If this question is answered positively, the next question for the legal advisor is: What special features must be taken into account in the context of Industry 4.0 and what urgent legal questions arise for Mechanical Engineering 4.0?

It may seem presumptuous to expect laws, some of which are over a hundred years old, to be able to regulate such modern issues. Technical innovations in particular do indeed pose challenges to the laws, such as the German Civil Code (BGB), which main features have been in force since 1900. Nevertheless, the high level of abstraction often allows principles and concrete regulations of these supposedly outdated bodies of law to be transferred to modern times to answer urgent legal questions satisfactorily. This conclusion—which comes as little surprise to lawyers—was also reached by the first ambitious studies, such as those of the Industry 4.0 platform, although the experts also admitted that in a few cases there is a need for minor adjustments and that in many cases a continuous review of the legal framework appears necessary to reflect current developments.[5]

In addition, finding a common language between technicians and lawyers is another difficulty that has not only been observed since the advent of Industry 4.0, but is becoming increasingly important in this environment. A legal assessment of possible risks and solutions, for example at the level of contract law, necessarily presupposes an unambiguous factual situation; the smallest ambiguities or deviations may lead to a completely different legal assessment. The new armory of a lawyer working in business therefore includes a certain minimum of technical expertise and the will to get involved in new kinds of facts and to include technical, sometimes disruptive concepts in one's own audit procedures from the very beginning. But the technical side in mechanical engineering will also have to develop an appropriate sensitivity for legal risks in the future—only in the interaction of lawyers with technicians and engineers will it be possible to find viable answers to upcoming questions in Mechanical Engineering 4.0.[6]

[4]Cf. Monitoring-Report Wirtschaft Digital, Federal Ministry for Economic Affairs and Energy (October 2017), p. 39, available at https://www.bmwi.de/Redaktion/DE/Downloads/C-D/digitalisierungsprofil-maschinenbau.pdf?__blob=publicationFile&v=4, (Effective: 04.01.2021).

[5]Cf. Plattform Industrie 4.0 result paper "Industrie 4.0 – wie das Recht Schritt hält" (October 2016), available at https://www.bmwi.de/Redaktion/DE/Publikationen/Industrie/industrie-4-0-wie-das-recht-schritt-haelt.html (Effective: 04.01.2021).

[6]Legal magazines, e.g., InTer – Zeitschrift zum Innovations- und Technikrecht account for this circumstance already.

3 Machine Generated Data

Along with the realization that data can have an economic value and is usable as commodity for the economy ("oil or sun of the 21st century"), a quite controversial discussion of science and practice about data sovereignty set in rather early (cf. just Hoeren 2013, pp. 486 et seq.). In this context, some prominent voices have called for the introduction of a virtual and **digital property law** that also includes data.[7] To be able to ensure protection of the value of data, **data ownership** is necessary.[8] This demand finds its background in the supposed legal uncertainties that can arise in relation to data. While data may enjoy occasional legal protection depending on the context (such as copyright, particularly the right of the database producer, protection of business secrets, data protection law in relation to personal data, etc.), a consensus (instructive to this Thalhofer 2017, pp. 225 et seq.) in literature and practice has been reached that the current legal framework does not provide for data ownership in a narrower sense, which gives the person entitled to it an exclusive, *inter omnes right* of control over the corresponding legal position.

Nevertheless, in practice, the creation of such data ownership is rejected by the majority or the necessity of its creation is questioned—both legally and economically, contractual agreements appear to be a more flexible and thus more expedient means of choice, which also counteracts the danger of a possibly innovation-hampering, hasty static allocation of an absolute right (Schlinkert 2017, p. 224). A static allocation by the legislator would also have to be questioned as to the basis on which it was made. Good reasons for the allocation can be found everywhere for every stakeholder.

The machine producer will be able to argue that without his product, not a single datum exists, i.e., he is the creator of the data, while the supplier of parts of the system, such as the sensor manufacturer, will be able to demonstrate just as comprehensibly that without his supplier part, such as the sensor, the datum cannot occur in the machine. However, the customer or operator of a system can also present a thoroughly noteworthy argument for his position as the correct recipient of a property right with the economic purchase or the establishment of ownership of the machine and the operation of the system.

A decision by the legislator in favor of one of these stakeholders with regard to an exclusive, absolute ownership right to the data (and thus the failure of such a right to any other stakeholder) would hardly be justifiable and would be associated with a certain randomness. Therefore, the author strongly advocates the allocation of data under contract law, also in view of the fact that in practice the parties will ensure the

[7] Cf. Speech of the former EU-commissioner for digital economy, Günther Oettinger, https://www.heise.de/newsticker/meldung/Hannover-Messe-Oettinger-fordert-einheitlichen-digitalen-Binnenmarkt-fuer-EU-2602252.html?wt_mc=rss.ho.beitrag.rdf (Effective: 04.01.2021).

[8] Cf. news release of the Bundestag and the chancellor, in which a property of data is demanded https://www.bundesregierung.de/Content/DE/Pressemitteilungen/BPA/2017/03/2017-03-18-podcast.html (Effective: 04.01.2021).

most economically sensible allocation and thus an appropriate balance of interests take place.

3.1 Data Without Personal Reference (Machine Data)

Most economically interesting data in the field of Mechanical Engineering 4.0 will be machine-generated data without any personal reference. While in other sectors the trade and use of data with regard to personalized data sets represent the envisaged or already implemented business models,[9] such data in mechanical engineering is rather unwelcomed or should be regarded as by-catch that does not represent the core element of data collection and the subsequent utilization models.

A typical example taken from practice in mechanical engineering will serve as the basic case for the following legal classification. The complexity of the case constellations hardly allows for the formation of further case groups, especially since each individual case may lead to a different outcome due to the variating interests and business models. Nevertheless, the following case may provide some guidance to enable a more detailed individual case assessment.

Example: Machine manufacturer M produces, among other things, intelligent machines. These are equipped with sensors from supplier Z. M's machine offers the possibility to collect, evaluate and display various data in real time. Operator B expects to optimize the products it produces with the system and to design ordering and logistics processes more efficient by using the data sets, in respect to its own customers or in respect to the material inflow required for the products. In addition, through access to the data, he expects to be able to better integrate other machines into the overall system, which may not originate from M.

M, on the other hand, would like to use the data to improve its machines or parts of them, but also to offer its customers new types of services, such as predictive machine-related maintenance. Finally, Z would like to have access to the data sets to be able to offer customer product-related consulting services in addition to optimizing the sensors. The data sets are stored in a (partly local) cloud of B. The local cloud is operated by B himself, the Internet-based cloud by the cloud service provider C.

This example may show how many different interests and parties exist in a common situation, related to the economic good of data. At the same time, it is suitable for deriving theses from this and, in the best case, verifying them.

While—as shown—a property right involves an absolute legal position *inter omnes*, contracts unfold their legal effect solely between the parties involved in the contracts (*inter partes*). It should be obvious that a contract that includes all parties

[9] Der Wert persönlicher Daten – Ist Datenhandel der bessere Datenschutz (the value of personal data—is data trade the better data protection?), study commissioned by Council of Experts for consumer's questions at the BMJV, June 2017, available at http://www.svr-verbraucherfragen.de/wp-content/uploads/Open_Knowledge_Foundation_Studie.pdf, S. 16ff. (Effective: 04.01.2021).

listed in the above example is neither practical nor realistic. In addition, it should be noted that the data sets are in principle immaterial (except of data storages) and can be used multiple times by several users without the data being consumed, i.e., the economic good is infinite, with the consequence that a limitation of the resource is only possible by technical mechanisms. Only through such measures is it possible to establish de facto/factual exclusivity (cf. also Stender-Vorwachs and Steege 2018, p. 1362).

3.1.1 Data Sovereignty Regulation

The starting point for considerations on contractual agreements on data sovereignty are the conditions for de facto exclusivity just mentioned. In practice, this often means that data generation and storage is factually controlled by one party, for example by setting up access authorizations with regard to encrypted or protected data storage. Due to the merely de facto exclusivity and the infinity of the resource already mentioned, a certain conceptual proximity to licensing law almost suggests itself. There, too, the resource, such as the trademark, the technical solution described and protected by a patent or the copyrighted work, is infinite and only subject to a de facto limitation or monopolization due to the special protection by trademark, patent or copyright law, but a real consumption in the sense of an exhaustion of the resource does not take place even in the case of multiple use (cf. Denga 2018, p. 1373).

The contractual regulation of data sovereignty, taking into account licensing law thought processes and principles, also offers the flexibility of allocation that is necessary in the mechanical engineering industry. Thus, these contractual regulations can provide for license-like agreements on the restriction, exclusivities, sub-licensing possibilities and the like, which enable a precise allocation of the rights to the data in accordance with the interests of the parties.

The regulation of data sovereignty via such agreements requires the inclusion of the usual points of regulation for licensing agreements, such as the object of the license or protection (**definition of data, categorization**), **holders of rights to use** and the scope of the **rights of use** granted.

Transferred to the practical example outlined above, these preliminary considerations result in the following legal approaches for drawing up contracts on data sovereignty. The respective interests in the data should find their way into the underlying contractual relationship and be mapped there along the lines of contracts under licensing principles. A distinction can also be made between data categories, for example. In this way, certain data categories can be defined—usually with the help of the technically experienced actors—which enable the respective data sets to be determined. In practice, a distinction is sometimes made between machine-related and product-related data sets, whereby the first data category refers to data sets that can provide information about the condition of the machine (wear data, etc.), while product-related data sets are also recorded by the machine but refer to the manufactured product (composition, etc.).

After the categorization of the data sets (definition), the beneficiary must be determined. In practice, it has been shown that certain data sets are covered by a special interest of one of the parties, while the other party shows no or only very little interest in them. In the practical example, the machine manufacturer M, for example, may have little to no interest in product-related data sets, as he does not manufacture any products of this kind himself, while the plant operator B has a particularly great interest in it. On the one hand, because he would like to use this data for his product optimization, on the other hand, because this data may be particularly worthy of protection from the operator's point of view and may even represent trade secrets which under no circumstances shall leave his control (for example, recipes, temperatures, compositions, etc.) and particularly must not be known to any competitor.

On the other hand, the plant operator B may only be interested in the machine-related data to a very limited extent, unless he can use it to make independent improvements to the machine, which will not often be the case. The value of the machine-related data often does not result from the concrete value of the data of a single operated machine, but rather from the evaluation of a multitude of data sets of different machines, so that for the machine operator B there may hardly be any added value in the data sets of his own machine, while such data sets can also be of high value for the machine manufacturer M in the interaction with the possibilities of Big Data analysis already mentioned at the beginning, to enable a further development and improvement of his machines.

Ultimately, the scope of right of use must be determined. Are the rights exclusive, is sub-licensing possible (in the example, for example, to supplier Z), which types of use by the parties should be prohibited? This opens up a wide range of possibilities. In this context, it also seems important to determine the technical protection of the data records in order not to compromise the de facto control over the data and thus to ensure the contractually agreed possibilities of use.

3.1.2 Data Access

The aforementioned discussion about the supposed necessity of data ownership led in practice and probably also in the majority of academia to the conclusion that the creation of such a property right was not necessary to ensure the protection of the data and the parties involved (cf. Hornung and Hofmann 2018, p. 27). However, this was increasingly followed by a discussion on the extent to which a **data access right** should be introduced to address (de facto) monopolies that may hinder competition. The basic assumption of these discussions is that the sometimes huge amounts of data of a fully networked industry can have a particular competitive relevance and that the exclusion of these data or the denial of data access could possibly mean a considerable competitive disadvantage.[10]

[10] About the whole and cf. result paper Industrie 4.0 – Kartellrechtliche Betrachtungen, Plattform Industrie 4.0 (Industry 4.0—cartel law considerations, Plattform Industrie 4.0) (April 2018),

For the mechanical engineering lawyer, this raises the question of whether, for example, the refusal of access to certain data sets is legally permissible or—seen from the other side—whether access to important machine data can be legally demanded. What if the contractual partner demands access to data, but the own company does not want to grant such access for various reasons? Or: The own company needs it for future analyses, but the contractual partner refuses this request and a compromise cannot be found?

The answers to this question can be found in national and European antitrust law: antitrust authorities recognize the relevance of data access as a power factor in the sense of market power under antitrust law,[11] and the German legislator introduced "access to competition-relevant data" as a criterion for assessing the market position of companies in the 9th amendment to the Act against Restraints of Competition (GWB), Section 18 (3a) No. 4 GWB. However, it should be noted that this does not imply that (exclusive) data sovereignty per se leads to a market-powerful or market-dominant company in the sense of antitrust law, but that this data sovereignty is only one (important) part of the assessment. Furthermore, it must be pointed out that market power, market dominance and even real monopolies are not necessarily prohibited under antitrust law. Only the abuse of a dominant position or—in relation to Germany—the abuse of relative market power is prohibited under antitrust law, Sections 19 and 20 GWB and Article 102 of the Treaty on the Functioning of the European Union (TFEU).

Thus, one will come to the conclusion with the prevailing opinion in the literature that antitrust law in its current form only provides for the obligation to grant data access or a basis for a claim for data access in a few exceptional situations. The main reason for this is that it will be difficult to prove a dominant or powerful market position and even more difficult to prove abuse of such a position (cf. Stender-Vorwachs and Steege 2018, pp. 1366 f.). The same will generally be true in practice. In practice, the same can probably also be assumed for any exploitative or discriminatory structuring of data-related contracts.[12] Moreover, in addition to the question of trade secrets possibly contained in the data, the question must also be asked whether, in the case of granting access to data for the benefit of a competitor, antitrust considerations do not speak precisely against this. If the data contain information relevant to competition or if such information can be analyzed from these data, it would have to be examined in the individual case whether there is no

pp. 17 et seq., available at https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/hm-2018-kartellrecht-ag4.html (Effective: 04.01.2021).

[11] About antitrust relevance of data sovereignty cf. just "Competition Law and Data" of the french Autorité de la concurrence and the German Federal Cartel Office (May 2016) as well as BKArtA "Big Data und Wettbewerb" (2017).

[12] Result paper Industrie 4.0 – Kartellrechtliche Betrachtungen, Plattform Industrie 4.0 (April 2018), (Industry 4.0—cartel law considerations, Plattform Industrie 4.0) (April 2018), p. 21, available at https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/hm-2018-kartellrecht-ag4.html (Effective: 04.01.2021).

exchange of information that is inadmissible under antitrust law (instructive Zimmer 2014, § 1 GWB paras 265 et. seq.).

For the mechanical engineering lawyer entrusted with the examination of possible data access issues, the above remarks mean that, according to the current state of German and European law, a claim to data access can only be assumed in specific individual cases. Nevertheless, it will be necessary to keep an eye on further legal developments on this issue, especially at the European level (e.g., free flow of data activities of the European Commission[13]).

3.2 *Data With Personal Reference (Data Protection)*

Machine-generated data can—intentionally or unintentionally—also contain data records that allow or make possible the identification of natural persons. Article 4 No. 1 of the General Data Protection Regulation (GDPR) shows a broad understanding of personal data (cf. Ernst 2018, Art. 4 para 3), which leads to complex questions and tasks in the practice of mechanical and plant engineering. For example, data that initially appears to be completely irrelevant in terms of data protection law can be relevant to data protection if the information—possibly taking into account additional information, which under certain circumstances does not even have to be specifically available in the respective company—can basically be assigned to a person (about the status of the dispute about whether a relative personal reference, which must also consider knowledge and resource from third parties, is sufficient to identify persons: Klar and Kühling 2018, Art. 4 para 25 et. seq.). Especially against the background of increasingly detailed information and the advancing possibilities of Big Data analysis, a far-reaching examination of data protection compliance will be necessary (cf. Mantz and Spittka 2017, pp. 148 f.). The most urgent task for lawyers will therefore be to examine the data records, if necessary, with the help of technically experienced experts, to determine whether they are related to a person or can be related to a person and how this should be dealt with under data protection law in the event of an affirmative answer or if there are doubts.

3.2.1 Anonymous Data/Anonymization

Since in the vast majority of cases in mechanical and plant engineering, machine-generated, personal data is not part of the business model of the companies or, in some cases, is necessary but can be dispensed with afterwards, companies in the

[13] Extensive remarks (e.g. about the access right for responsible authorities for regulatory control purposes) available at http://europa.eu/rapid/press-release_IP-18-4227_de.htm (Effective: 04.01.2021).

sector sometimes rely on the anonymization of this data to already not fall within the scope of application of the data protection regulations and to have to implement the—in the course of the chapter outlined—far-reaching requirements. According to recital 26 p. 5 of the GDPR, the principles of data protection are not to be applied to anonymous information or to personal data that has been anonymized in such a way that the data subject cannot be identified or can no longer be identified.

Whereas the old Federal Data Protection Act (BDSG-alt) defined anonymization in § 3 (6) BDSG-alt as "the alteration of personal data in such a way that the individual data on personal or factual circumstances can no longer be attributed to a specific or identifiable natural person or can only be attributed to a specific or identifiable natural person with a disproportionate expenditure of time, costs and labor," the new Federal Data Protection Act (BDSG-neu) and the GDPR do not provide for a definition of anonymization. However, the aforementioned recital 26 of the GDPR in p. 4 speaks of the fact that "all objective factors, such as the cost of identification and the time required for it, [should] be taken into account, taking into account the technology and technological developments available at the time of the processing" to determine whether a natural person is (no longer) identifiable and a data is thus anonymous or anonymized.

Insofar as personal data cannot be prevented in advance,[14] recognized methods for anonymization in mechanical engineering practice are, above all, the deletion of identifying characteristics and the aggregation of data.[15] However, it should be noted here that initially, personal data is collected, which is only anonymized in a second step, i.e., only from this point on is it no longer subject to the principles of data protection law. However, the initial collection still requires justification (see also below). Only the subsequent use of this anonymous data (e.g., Big Data analysis) is possible without taking data protection into account.

3.2.2 Data Protection Principles

The central principle in data protection law is the prohibition with reservation of permission, whereby the prerequisites for the lawfulness of the processing result from Art. 6 (1) GDPR. This means that any processing of data is prohibited unless it is permitted by law or the consent of the data subject. In practice, data processing for the fulfilment of contractual obligations (Art. 6 (1) lit. b) GDPR) and processing within the framework of the balancing of interests (Art. 6 (1) lit. f) GDPR) come into consideration as grounds for permission in mechanical and plant engineering. In addition, the possibility of the data subject's consent to the data processing (which may also be implied) must always be considered, Art. 6 (1) lit. a) GDPR. When

[14] Principles of data minimization must be considered, Art. 5 (1) lit. c) GDPR respectively the Privacy by Design and Privacy by Default, Art. 25 GDPR will have to be considered.

[15] Cf. also: Stellungnahme 5/2014 (WP216) zu Anonymisierungstechniken der Artikel-29-Datenschutzgruppe (10.04.2014), including guideline (appendix, pp. 32 et seq.).

assessing the possibilities of justification, attention should be paid to the efficient use of the legal grounds and consent should only be used if the legal grounds for justification do not apply (Rücker 2015, p. 37).

Other principles to be particularly observed are the data protection requirements of "**Privacy by Design"** and **"Privacy by Default**," Art. 25 GDPR. Through this data protection by technology design or data protection-friendly default settings, machine and plant manufacturers will already be required to include appropriate data protection precautions in the planning when developing new products and business models. This applies both in the case where the personal data is (also) processed by the manufacturer itself, as well as in the case where, for example, with the sale of the machine (only) the operator of the system carries out the data processing. While in the latter case there is no position as a data controller in the sense of data protection law, it must be possible for the products to be used by the data controller—i.e., the operator—in compliance with data protection law. If this is not the case, the manufacturer cannot be accused of a data protection violation, but the operator of the system is in breach of applicable data protection law, which may result in warranty claims against the manufacturer (Schuster and Hunzinger 2017, pp. 146 f.). Therefore, manufacturers will only be able to market products that comply with data protection requirements (Mantz and Spittka 2017, p. 155, para 41).

The mechanical and plant engineering sector will also have to pay special attention to cross-border data transfers in the context of Industry 4.0. The transfer of personal data will in many cases have a cross-border connection simply due to technical circumstances (e.g., cloud services that operate servers abroad for these services). If a transfer only takes place within the EU, only the regulations of the GDPR (and, if applicable, national regulations of the Member States due to so-called opening clauses within the GDPR) must be observed.

In the case of a transfer to third countries (states outside the EU or the EEA), Art. 44 et seq. of the GDPR impose far-reaching verification obligations on the controller with regard to the permissibility of the data transfer. If there is no so-called adequacy decision of the European Commission on the level of protection of the third country (Art. 45 GDP[16]), suitable guarantees such as the EU standard data protection clauses of the EU Commission or binding internal data protection regulations (Art. 47 GDPR) can also justify the permissibility of a data transfer to a third country. Finally, it should be noted that even if the transfer of data is permissible in principle, the transfer itself requires a suitable legal basis, such as the permissibility of data processing in the context of processing for the purpose of fulfilling a contract, etc.

[16] Cf. https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/adequacy-protection-personal-data-non-eu-countries_de. an adequacy decision exists for Andorra, Argentina, Canada, Faroe-Islands, Guernsey, Israel, Isle of Man, Jersey, New Zealand, Switzerland, Uruguay and the USA (limited to the regulations of the Privacy Shield) (Effective: 04.01.2021).

4 Legal Liability: Product and Producer Liability

Like any revolutionary or evolutionary economic and technical development, Industry 4.0 also raises various liability issues in mechanical engineering—while in practice the contractual liability agreements are often decisive (see also II. D. b) on the problems arising from this), where there are no direct contractual relationships between the actors, statutory liability must be considered. Accordingly, product and producer liability will be examined in more detail below, especially taking into account the increasing networking and an Industry 4.0-like organized production process.

4.1 Producer-Liability of the Manufacturer

The **producer-liability** can be traced back to the principle that when a product is placed on the market (cf. Higher Regional Court Hamm NJW-RR 2011, 893), a duty of care arises which the producer violates with a defective product. The producer must ensure, within the bounds of what is technically possible and economically reasonable, that customers, users of the product and other third parties are not harmed (Förster 2018, § 823 para 673). The producer's obligations arise specifically with regard to design, manufacturing, instruction and product monitoring obligations. Although the term producer is not defined by law, it can be defined as the person who significantly directs and controls the development and production process, whereby this includes the producer of a partial product as well as the (end)producer of the final product (cf. Wende 2017, p. 71 para 8).

In case law, **construction obligations** are understood to mean the observance of the required safety level already at the time of design or the planning and development phase (Wagner 2017, § 823 para 818). The product might fall short of the required safety standard in this phase, which results from the current state of science and technology (cf. just Federal Supreme Court from 16.06.2009. AZ VI ZR 107/08, NJW 2009, 2952 – Airbag). In the context of Industry 4.0, special features or problems not yet known in industry will probably only rarely become apparent here (also Wende 2017, p. 71 para 15)—especially since the practical relevance in case law appears to be rather subordinate anyway (cf. to verdict Wagner 2017, § 823 para 818).

On the other hand, the producer's compliance with its **manufacturing obligations** can be considered more relevant—in practical terms also due to the new Industry 4.0-like production processes. A breach of the manufacturing obligations occurs if the product, which is in itself planned with the required degree of safety, is implemented incorrectly in the manufacturing process. In addition to the aspects of the autonomation of industrial processes, which may make it important to take necessary and reasonable measures to ensure the exclusion of human error with regard to the new planning of manufacturing processes (Wende 2017, p. 73 para 17), "smart factories" will present machine manufacturers with factual hurdles. On manufacturing obligations, it is recognized that the risks arising from partial products supplied by third parties for the overall product are borne by the manufacturer of the overall/final product (cf. Wagner 2017, § 823 para 708). In a production that (partially) autonomously requests supplier products and processes them to manufacture the entire product, the demands on the organization of the production process (e.g., on sampling, cf. Higher Regional Court Oldenburg NJW-RR 2005, 1337 - Fahrradpedale) will increase.

The **instruction-duties** will also be of increased importance in the Industry 4.0 environment. Accordingly, the manufacturer is obliged to warn the user of a product of those (residual) hazards which are imminent in the case of intended use or obvious misuse, and which are not part of the general hazard knowledge of users (BGHZ 181, 253, NJW 2009, 2952 para 23 with further references). It is important to note that—far-reaching—instructions do not exempt from the obligation to comply with the (preceding) design obligation and refer to the residual dangers—despite the implementation of the design obligation. In view of the increasingly complex manufacturing processes of networked products, producers will have to pay attention to the documentation accompanying the product, such as the operating instructions. It is true that the duty to instruct cannot be interpreted without limits, especially since there are likely to be an almost infinite number of sources of danger. Within the bounds of what is reasonable, however, the manufacturer is obliged to make a case-by-case assessment of the severity of the impending damage and the probability of its occurrence if no warning is given (cf. Wagner 2017, § 823 para 830), and—depending on the sphere concerned—to consider adapted, comprehensible instructions for use or very clear (warning) notices.

Ultimately, the manufacturer's duty also includes **product-monitoring obligations**, i.e., the duty to assume responsibility for the product even after it has been placed on the market, beyond the scope of responsibility at the time of placing the product on the market (and thus the state of knowledge and possibilities at that time) and, in the event of danger to the user or third parties, to take measures, if necessary, to avert this danger. The intensity of the duty to observe the product is determined by the dependency between the extent of the possible damage, the type and intensity of the possible hazard, the probability of occurrence or the severity of the hazard for the legal asset at risk and what is possible and economically reasonable.

A distinction is made between a passive and an active product monitoring obligation. While the passive obligation is limited to receiving, collecting and systematically evaluating cases of damage and indications of safety deficits via a complaint management from customers and third parties (BGHZ 99, 167, 170f. – Lenkerverkleidung), the active obligation is seen as independently generating information about possible risks of damage, for example via experience with the same or similar products, the evaluation of scientific-technical literature, etc. The active obligation is seen as a duty to monitor the product (Wagner 2017, § 823 para 838). Especially in connection with Industry 4.0-typical, complex new developments such as software-controlled devices, the intensity of the monitoring obligations will increase and gain in complexity. If a manufacturer identifies hazards, he is obliged to react, such as changing production, warning users of product hazards or, ultimately, recalling products that have already been marketed.

Producer liability—unlike the liability under the Product Liability Act—ultimately presupposes fault on the part of the producer, i.e., a culpable violation of their **duty to implement safety precautions**. The aggrieved party bears the burden of proof with regard to the product defect and its causality for the legal injury that occurred (BGHZ 51, 91, 102). Since it is often impossible for the aggrieved party to gain detailed insight into the production processes, case law has developed certain forms of facilitating proof, so that the manufacturer must prove that there has been no culpable breach of his duty of care with regard to the alleged breach of duty to implement safety precautions (Wagner 2017, § 823 para 838). Especially on the (partially) autonomous or self-learning systems, the question arises whether fault-based producer liability has reached its limits. Apart from the discussion (cf. Palmerini and Bertolini 2015, p. 244 with further references) of the extent to which today's systems actually contain or will ever achieve a certain autonomy in their actions or always remain in principle a machine that decides (and possibly also makes new decisions) purely based on the given parameters, there will always be a duty to take into account that the producers must make preparations on the design side that take into account and avoid system-immanent safety risks in accordance with the principles of producer liability (Wende 2017, p. 82 para 70). However, this requires foreseeability, which may not necessarily be given in the case of—almost—independently developing systems, e.g., because atypical user behavior led to a further development of the system causing the damage. In this case, the developer of the system can only be held culpable to a very limited extent (Bräutigam and Klindt 2015, p. 1137; Horner and Kaulartz 2016, p. 24), e.g., for allowing the system to develop further in the first place (which will often be the goal of such Industry 4.0 systems). In this respect, the traceability of product development will become increasingly important to be able to clearly identify sources of error and those responsible for them (Wendt and Oberländer 2018, p. 64).

4.2 Product Liability

In contrast to fault-based producer liability, the manufacturer's product liability is linked solely to the placing of a defective product on the market, without it being relevant whether the manufacturer did so culpably (strict liability, § 1 para. 1 German Product Liability Act). The manufacturer within the meaning of the Product Liability Act is the person who manufactured the final product, a basic material or a partial product, § 4 para. 1 Product Liability Act. Liability is also extended to the so-called quasi-manufacturer who claims to be the manufacturer by affixing his name, trademark or distinctive mark, § 4 para. 1 sentence 2 Product Liability Act. Under § 1 para. 1 Product Liability Act, the producer is liable for damage caused by a defective product with regard to the physical integrity and predominantly privately used property of the consumer.

The Product Liability Act considers a product to be defective if it does not offer the safety that can reasonably be expected taking into account all circumstances, § 3 para. 1 Product Liability Act. Circumstances to be particularly considered are the presentation, the use that can reasonably be expected, and the time when it was put on the market. In determining whether a product defect exists, the case law is based on the defect categories already explained above in the context of producer liability in tort. Exceptions to the producers' liability result from § 1 para. 2 Product Liability Act. Thus, the liability to pay compensation is excluded if, for example, the manufacturer has not (yet) placed the product on the market (§ 1 para. 2 no. 1 Product Liability Act), the defect was not yet present at the time of placing the product on the market (no. 2) or could not be detected based on the state of the art in science and technology (no. 5).

Special features on liability concepts arise only to a limited extent in the context of Industry 4.0, so that in practice and literature, despite the new developments in Mechanical Engineering 4.0, no fundamental need for reform is recognized (Spindler 2015, pp. 766 et seq.). Existing ambiguities of responsibilities are to the detriment of the manufacturer of the end product (Hornung and Hofmann 2018, p. 42), so that especially machine builders—as already explained in the context of producer liability—must in future pay more attention to clearly comprehensible product development and production processes to document responsibilities at least in the internal relationship.

It should also be noted that at the time of writing this book, the Product Liability Directive (Directive 85/374/EEC) is being evaluated at European level.[17] Whether an adaptation is necessary against the background of Industry 4.0 is widely doubted by the business community.[18]

4.3 Special Case IT Security

Machine builders will increasingly have to face the question of liability in terms of the extent to which software they have developed themselves, e.g., for the operation and/or networking of machines, are to be evaluated from the point of view of **IT security**. With the increasing networking of machines, cyberattacks of various contents are becoming more and more likely. The economic significance of such attacks is increasing every year, regardless of whether they are business interruptions, blackmail attempts or industrial espionage (more on liability relations for cyberattacks: Mehrbrey and Schreibauer 2016, pp. 75 et seq.). Accordingly, there

[17] https://ec.europa.eu/growth/single-market/goods/free-movement-sectors/liability-defective-products_en.

[18] Cf. statement ORGALIM https://www.orgalim.eu/position-papers/orgalime-comments-evaluation-product-liability-directive.

is also an increasing risk that such cyberattacks will cause damage that may have to be considered in the context of producer and product liability.

Irrespective of the disputed question of whether software even constitutes a product within the meaning of the Product Liability Act, the protective purpose of the Product Liability Act (consumer protection, privately used objects, personal injury) means that a basis for a claim relevant to industrial practice cannot yet be assumed. However, since a broader scope of liability may exist in the context of tortious producer liability on, e.g., further damage (interest in integrity), the discussion must be held on the extent to which inadequate IT security in the use of software constitutes a product defect. Under Section 2 (2) German BSI Act, security is defined as the compliance with certain security standards concerning the availability, integrity or confidentiality of information and is ensured by security precautions in and during the use of IT systems, components, or processes.

In principle, the responsibility of the software manufacturer ends where a third party intentionally and abusively interferes with the event or the product and damage occur as a result (Rockstroh and Kunkel 2017, p. 81). However, this may be different if—as is likely to be increasingly the case—the user could, for example, expect software that is protected against attacks and the manufacturer should have foreseen such security problems in the software according to the state of the art in science and technology (Wende 2017, p. 76 para 42). In such cases, the manufacturer's claim will have to be investigated, which also seems appropriate if a latent danger caused by the software is realized (cf. Spindler 2004, p. 3146). The decisive factor will therefore be whether the producer of the software should have been aware of the security vulnerability at the time it was placed on the market (Spindler 2007, para 123). If the manufacturer does not take all possible and reasonable measures when he is aware of it, he is exposed to a corresponding liability risk (Rockstroh and Kunkel 2017, p. 82). Mechanical Engineering 4.0 will therefore have to devote itself intensively to the analysis of cyberattacks and their gateways (Wende 2017, p. 77 para 42), also to be able to fulfil its obligation to monitor products after they have been placed on the market. Finally, companies should also consider cyber insurance to cover existential liability risks in addition to the indispensable prevention.

5 Software in Mechanical Engineering 4.0

While German machine manufacturers dominate their core business, machine, and plant construction, on an international scale and offer leading products worldwide, the last decade has seen a development that is now (also) reflected in the context of Industry 4.0. The use, but also the own development, of increasingly complex software poses challenges for the machine builder 4.0, mostly also of a legal nature. These will be briefly outlined below, also and especially to achieve a certain sensitivity among practitioners for this unfortunately often neglected and therefore sometimes quite risky legal topic.

5.1 Software License Agreements

Under Section 69a (1) German Copyright Act, computer programs are programs in any form, including design material, which are protected under Section 69a (2) sentence 1 and (3) sentence 1 Copyright Act in any form of expression, insofar as they constitute individual works in the sense that they are the result of their author's own intellectual creation. In this respect, copyright protection is governed by the provisions applicable to linguistic works, Section 69a (4) Copyright Act. A legal definition for computer software, however, is not found in the Copyright Act. However, the Federal Court of Justice and the predominant legal literature define computer programs as "the means of expression, chosen in any form, language and notation or in any code, of a sequence of instructions intended to cause a computer to perform a particular task or function" (Marly 2018, para 74 with further references). It can therefore be stated that the vast majority of software developed and used by mechanical engineers will constitute a copyrightable work within the meaning of the Copyright Act.

Based on the principle that copyright cannot be transferred by legal transaction (unlike, for example, ownership of movable property, §§ 929ff. German Civil Code), there is a dispute in the legal literature about the typification of the transfer of software. On the one hand, due to the fact that the transfer of copyright is excluded by copyright law, the transfer is considered to be characterized by licensing law, whereby licensing law represents the sum of all those rules that result from the general provisions of the law of obligations (Hilty 2003, p. 15). On the other hand, it is argued that the term "license" is unclear and does not indicate which legal consequences are associated with it (Schug and Rockstroh 2018, para 1 with further references), so that due to the contractual typology of the German Civil Code and the lack of a "license law" (Marly 2018, paras 736 et seq.), a distinction must be made as to whether a transfer is made for a limited period of time (in which case it is more likely to be classified as a lease) or for an unlimited period of time (in which case it is classified as a purchase). However, it should be noted—without having to decide the dispute at this point—that a legal contract between the manufacturer of the software and the user is to be assumed with the transfer, even if no explicit arrangements on this have been reached between the parties.

5.2 Source Code of the Software

The problem of a lack of a formulated contractual basis for the transfer of software always becomes apparent when the parties are in dispute about which rights to the software were actually granted. In addition to questions of warranty law (which are based on the above-mentioned typification and the resulting legal consequences of these types of contracts and cannot be dealt with here due to the limited scope of this chapter), the questions of decompilation possibilities and access to the source code have proven to be particularly relevant in practice in mechanical engineering. While the decompilation of a computer program is (compulsory) governed by the extensive provisions of Section 69e Copyright Act and is generally only possible without the consent of the right holder if the reproduction of the code or the translation of the code form is indispensable to obtain the necessary information for establishing interoperability with other programs (cf. about the whole Marly 2018, paras 255 et seq.) and this represents the last (appropriate) means of establishing interoperability (Grützmacher 2014, § 69e para 13), the question of the purchaser's claim to surrender of the source code is an ongoing issue that occupies Mechanical Engineering 4.0.

Primarily, it should be noted that a clear answer regarding the obligation to hand over the source code is difficult if no explicit contractual agreements have been made in this respect (in such a case, it is usually the case that the manufacturer states that the software is only handed over in the object code and that a handover of the source code is not owed). In the absence of such an explicit provision, according to the opinion of the highest court and the literature, an interpretation of the contract and the elaboration of the purpose of the contract must be used to determine an answer (Federal Supreme Court, 16.12.2003, AZ X ZR 129/01). According to the fundamental decision of the Federal Court of Justice, three criteria in particular are decisive. The decisive factor is who is responsible for error correction, maintenance, and modification work on the software. At least if the buyer has to take over the maintenance of the program himself, he will be entitled to demand the source code for this purpose. Furthermore, it must be asked whether the software was recognizably intended from the beginning (also) for further distribution to third parties. The buyer could not make the usual additions and changes to the program in this context without knowledge of the source code. Ultimately, the agreed remuneration can provide conclusions as to whether, in addition to the mere creation of the software program, the handing over of the source code is also appropriate as a further obligation.

It seems important in this context that according to the so-called purpose transfer doctrine of copyright under Section 31 (5) Copyright Act (in the absence of specific provisions, only those rights are granted that are necessary to achieve the purpose of the contract), the rights remain as far as possible with the author in case of doubt (Marly 2018, para 688). This is to be particularly considered with regard to the right of the customer to change the software under certain circumstances and thus to be able to demand the surrender of the source code if necessary.

5.3 Use of Open-Source Software

Companies in mechanical and plant engineering are increasingly relying on **open-source software (OSS)** in the context of Industry 4.0, for example in the area of embedded systems. OSS promises a multitude of advantages in Mechanical Engineering 4.0, from the collaborative approach to creation to cost benefits. However, it will be the mandatory task of the lawyer to balance the undoubtedly existing operational advantages of OSS with the legal risks of its use and to rely on close interaction with the software developers. It will obviously be difficult for the lawyer to recognize, for example, in the aftermath of system developments, whether OSS has been used and what consequences this may have for the company's proprietary software.

The main risk here is the so-called "viral effect." In the case of the use of OSS, which provides for strict "copyleft" licensing, proprietary software is threatened with the fate of falling under far-reaching OSS obligations.[19] This means that when using such OSS, the user is obliged to make the modified program available free of charge and to allow other users to further develop it (Marly 2018, para 927; Marschollek and Wirwas 2017, p. 430, para 39), i.e., particularly to disclose the source code. It is therefore advisable to oblige software developers to document the OSS used or to draw up framework conditions for the use of OSS in advance of development, which, for example, exclude the use of OSS under certain license models for sub-areas.[20]

In practice, the above remarks mean that Mechanical Engineering 4.0 will have to pay much more attention to software law issues and particularly ensure that (contractual) regulations on software licensing are in place. The use of OSS will offer far-reaching possibilities, especially for SMEs, whereby it is imperative to observe the licensing conditions to avoid undesirable legal consequences.

6 Contract Design/Terms and Conditions

The comments in this chapter, for example on data sovereignty and software licensing, show that Mechanical Engineering 4.0 often finds its place in the existing legal framework. However, it has also become clear that innovative business processes and business models are absolutely dependent on robust contractual regulations. Standardized contracts prepared for a large number of cases will play an important role here. In practice, hardly any contract will be individually negotiated, i.e., the core content of the individual clauses, which is not in line with the law, will be seriously questioned and the negotiating partner will be given freedom of design to protect their own interests, but instead prefabricated standard contracts will be used (Basedow 2017, § 305 para 35). This leads to the difficulty that under German law the contracts are subject to the regulations on general terms and conditions (GTC) under § 305 Civil Code, and particularly to the GTC-content control under § 307 et seq. Civil Code.

The general transfer of the case law in the area of consumers to purely business-to-business transactions (B2B) therefore presents lawyers with serious problems

[19] About copyleft as a whole: https://www.gnu.org/licenses/copyleft.de.html.

[20] A collection of OSS license models is available at https://opensource.org/licenses/alphabetical.

when attempting to draw up robust contractual agreements for the envisaged business models. While protection against overreaching by the entrepreneur is certainly justified with regard to consumers, the tendency of the Supreme Court to successively extend consumer protection to B2B matters is a disadvantage for Germany as an industrial location that can hardly be tolerated. A reform of the law on general terms and conditions on B2B matters is therefore called for.[21] The Frankfurter Initiative zur Fortentwicklung des AGB-Rechts (the Frankfurt Initiative for the developing of the GTC-law) demands a reform of the GTC-law in respect of B2B-issues.[22]

As long as this obvious locational disadvantage exists, the lawyer will often have to consider choosing a foreign, more flexible legal system, in which the effectiveness of agreed performance commitments, risk distributions, liability definitions and allocations can be agreed in a legally secure manner, especially in Industry 4.0 processes, which often do not end at national borders (Schlinkert 2017, p. 222).

References

Basedow J (2017) In: Säcker, Rixecker, Oetker, Limperg (eds) Münchener Kommentar zum Bürgerlichen Gesetzbuch, 7th edn

Bräutigam P, Klindt T (2015) Industrie 4.0, das Internet der Dinge und das Recht, NJW 2015, 1137

Denga M (2018) Gemengelage privaten Datenrechts, NJW 2018, 1371

Ernst S (2018) In: Paal, Pauly (eds) Datenschutz-Grundverordnung Bundesdatenschutzgesetz, 2nd edn

Förster C (2018) In: Bamberger, Roth (eds) Beck'scher Online-Kommentar BGB, 48th edn

Grützmacher M (2014) In: Wandtke, Bullinger (eds) Praxiskommentar zum Urheberrecht, 4th edn

Hilty R (2003) Der Softwarevertrag – ein Blick in die Zukunft, Konsequenzen der trägerlosen Nutzung des patentrechtlichen Schutzes von Software, MMR 2003, 3

Hoeren T (2013) Dateneigentum - Versuch einer Anwendung von § 303a StGB im Zivilrecht, MMR 2013, 486

Horner S, Kaulartz M (2016) Haftung 4.0, InTeR 2016, 22

Hornung G, Hofmann K (2018) Industrie 4.0 und das Recht: Drei zentrale Herausforderungen, In: Hornung (ed) Rechtsfragen der Industrie 4.0

Klar M, Kühling J (2018) In: Kühling, Buchner (eds) DS-GVO BDSG, 2nd edn

Mantz R, Spittka J (2017) Datenschutz und IT-Sicherheit. In: Sassenberg, Faber (eds) Rechtshandbuch Industrie 4.0 und Internet of Things

Marly J (2018) Praxishandbuch Softwarerecht, 7th edn

Marschollek JP, Wirwas M (2017) Digitalisierung der Elektroindustrie. In: Sassenberg, Faber (eds) Rechtshandbuch Industrie 4.0 und Internet of Things

Mehrbrey KL, Schreibauer M (2016) Haftungsverhältnisse bei Cyber-Angriffen, MMR 2016, 75

[21] Cf. only association statement "AGB-Recht für Unternehmen modernisieren – Wirtschaftsstandort Deutschland stärken" (Modernise GTC law for businesses—strengthen the industrial base Germany), available at https://www.vdma.org/v2viewer/-/v2article/render/26919255 (Effective: 04.01.2021).

[22] https://www.bmjv.de/SharedDocs/Artikel/DE/2015/02092015_AGB_Recht.html.

Palmerini E, Bertolini A (2015) Liability and risk management in robotics. In: Schulze, Staudenmeyer (eds) Digital revolution: challenges for contract law in practice

Rockstroh S, Kunkel H (2017) IT-Sicherheit in Produktionsumgebungen, MMR 2017, 77

Rücker D (2015) In: Bräutigam, Klindt (eds) BDI-Gutachten Digitalisierte Wirtschaft / Industrie 4.0

Schlinkert H-J (2017) Industrie 4.0 – wie das Recht Schritt hält, ZRP 2017, 222

Schug C, Rockstroh S (2018) In: Weise, Krauß (eds) Beck'sche Online-Formulare Vertrag, 46th edn

Schuster F, Hunzinger S (2017) Pflichten zur Datenschutzeignung von Software, CR 2017, 141, 146f

Spindler G (2004) IT-Sicherheit und Produkthaftung – Sicherheitslücken, Pflichten der Hersteller und der Softwarenutzer, NJW 2004, 3145

Spindler G (2007) BSI-Gutachten Verantwortlichkeiten von IT-Herstellern, Nutzern und Intermediären

Spindler G (2015) Roboter, Automation, künstliche Intelligenz, selbst-steuernde Kfz – Braucht das Recht neue Haftungskategorien?, CR 2015, 766

Stender-Vorwachs J, Steege H (2018) Wem gehören unsere Daten?, NJOZ 2018, 1361

Thalhofer T (2017) Recht an Daten in der Smart Factory, GRUR-Prax 2017, 225

Wagner G (2017) In: Säcker, Rixecker, Oetker, Limperg (eds) Münchener Kommentar zum Bürgerlichen Gesetzbuch, 7th edn

Wende S (2017) Haftungsfragen bei vernetzten und autonomen Systemen. In: Sassenberg, Faber (eds) Rechtshandbuch Industrie 4.0 und Internet of Things

Wendt J, Oberländer M (2018) Produkt- und Produzentenhaftung bei selbständig veränderlichen Systemen, InTeR 2016, 58

Zimmer D (2014) In: Immenga, Mestmäcker (eds) Wettbewerbsrecht, 5th edn

Part III
Transport, Logistics and Construction

Electromobility: Trends and Challenges of Future Mass Production

Achim Kampker, Kim Kreisköther, Patrick Treichel, Tom Möller, Yannick Boelsen, and Daniel Neb

Contents

1 Initial Situation e-Mobility

Contrary to the general perception that "electromobility" is an invention of the twenty-first century, its beginnings date back to the nineteenth century, when the American Thomas Davenport developed the first battery-powered electric vehicle.

However, the first electric vehicles were still designed with non-rechargeable batteries. Only the invention of lead accumulators made it possible for Gustavo Trouve to design an electrically powered vehicle in such a way that the battery did not require replacement (Becker 2010).

A. Kampker (✉) · K. Kreisköther · P. Treichel · T. Möller · Y. Boelsen · D. Neb
Chair of Production Engineering of E-Mobility Components, RWTH Aachen University, Aachen, Germany
e-mail: A.Kampker@pem.rwth-aachen.de; kim.kreiskoether@pem.rwth-aachen.de; P.Treichel@pem.rwth-aachen.de; t.moeller@pem.rwth-aachen.de; y.boelsen@pem.rwth-aachen.de; d.neb@pem.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_29

Initially, the electric drive train was even able to prevail over the combustion engine. Due to the resulting advantages such as reliability, safety and low noise emissions, the electric motor was used in 38% of all cars produced in the USA at the end of the nineteenth century (Kampker 2014). The electric car industry reached its peak in 1912 when 34,000 cars with an electrified drive train were produced worldwide (Möser 2002). With the development of the electric starter at the same time, the internal combustion engine began to gain acceptance, mainly due to its greater driving range and greater comfort. For this reason the electrified powertrain was only used in niches at the end of the Second World War (Kampker 2014).

Since the 1970s, there has been a growing environmental awareness among the population worldwide. In addition to reducing the consumption of resources, this also refers to the reduction of global pollutant emissions. The ambitious goal of international climate policy of keeping global warming below two degrees Celsius will only be possible if CO_2 emissions are reduced worldwide. To achieve this goal, not only industry and households, but particularly the transport sector must be considered.

A rethinking from conventional vehicles with combustion engines to electrified vehicles powered by renewable energies is an essential approach to achieve the climate targets set. The need for electromobility was also recognized by the German government, which adopted the "National Electromobility Development Plan" in 2009, which aimed to have one million electric vehicles on German roads by 2020 (Nationale Plattform Elektromobilität 2011). Along with this plan, politicians intend to invest in research and development of electric energy storage systems, vehicle technology and system and network integration.

Since then, electromobility has also been gaining in importance in Germany. The increasing market success of new manufacturers of electrified vehicles such as Tesla or BYD and the customers' mistrust in diesel combustion engines since the so-called "emissionsgate" in 2015 have led to all renowned car manufacturers increasingly following the trend of electric mobility with their own hybrid or electric vehicles (Amelang and Wehrmann 2020).

Numerous factors play a role in the market upturn. On the one hand, companies orient their research and development to market opportunities. Only when it can be assumed that the market is ready to accept disruptive technologies, companies will be compelled to react to changing market conditions. On the other hand, customers also expect improvements in technology.

Current major obstacles to the purchase of an electric vehicle are still the lack of a charging infrastructure, short ranges and high acquisition costs. While the charging infrastructure has already been addressed by government subsidy programs, it is necessary to overcome the low ranges and high initial costs through product and production technology innovations.

The need for this is particularly evident in the close coupling of the perceived low range of electric vehicles and the high acquisition costs, as the range is highly dependent on the capacity and number of battery cells used. To achieve ranges similar to those of the combustion engine while maintaining cost parity, the battery cells used in the vehicle must fall significantly below the limit of 100 €/kWh.

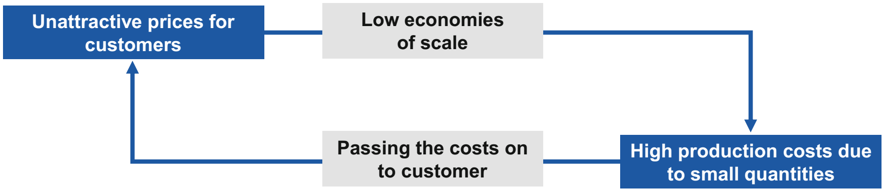


Fig. 1 Vicious circle of small quantities for electrified vehicles

One reason for the high initial costs or the economic challenges of electrified vehicles with regard to the battery lies in the small quantities in which the derivatives are currently produced. This prevents the occurrence of economies of scale and thus increases the corresponding costs, which are ultimately borne by the customer (Kampker 2014). An initially low demand on the side of the customer causes high production costs, since no economies of scale can be achieved. The increased production costs must be passed on to the customer due to market mechanisms, which leads to a price-related unattractiveness and to a stagnating small demand.

On the one hand, government measures such as promotional programs can directly subsidize the purchase of electric vehicles, thus enabling a break out of the vicious circle. On the other hand, the production capacities of automobile manufacturers (OEM) can be expanded and further developed at increased entrepreneurial risk to achieve the necessary economies of scale. Currently, both mechanisms are evident in the market (Fig. 1).

Not only car manufacturers and end users are affected by the electrification of the vehicle fleet, but also the entire automotive supplier industry, which is usually less known by end users.

Due to the substantial change in the powertrain, these particular supplier companies are confronted with a changed and strongly fluctuating demand for products from the "classic product portfolio" as well as with new types of production processes (Burggräf et al. 2011).

Previous production technologies for the conventional powertrain based on internal combustion engines, for example, are mainly focused on machining production technologies. The previously usual high machining times are now reduced by up to 74% in the electric drive train. In the course of electrification, the focus is now on forming, stamping and joining technologies (Eberhard Abele 2009).

The current challenge for the entire automotive industry is to uncover new technology potential, to flexibly shape and thus minimize the costs of the manufacturing process.

2 Introduction to the Fields of Action

Based on the changing components in the electric vehicle, the following sections discuss the trends and challenges of these fields of action from a production technology point of view with regard to the upcoming high-volume product:

1. Complete vehicle
2. Battery systems
3. Electric drive
4. Vehicle body
5. Recycling

In these fields of action, on the one hand, the components and underlying production technologies are being changed, and on the other hand, components from the conventional powertrain are being substituted by new components with new production technologies. However, new fields of action are also emerging, such as the recycling industry, which have an impact on the entire value chain over the entire product life cycle. These too must be considered and designed before the ramp-up of large-scale production.

2.1 Field of Action 1: Complete Vehicle

The change from the conventional vehicle powered by an internal combustion engine to an electric vehicle leads to a drastic change in components and vehicle structure. Thus, the combustion engine as well as the exhaust and tank system are no longer required. Due to the electrification of the powertrain, the transmission, wheel suspension, air conditioning, braking and steering system (especially auxiliary units) as well as the heating and thermal insulation have to be redesigned and redeveloped or adapted to the new requirements. The vehicle is completed by an electric motor, high-voltage battery including battery management system (BMS), power electronics and charging connection (Kampker 2014).

The change and pervasion of electric vehicles will take place gradually and in addition to the purely electrically powered vehicles, there will also be a large proportion of hybrid vehicles. The design of hybrid electric vehicles can generally be divided into three different categories that have an impact on the overall vehicle design (Fig. 2).

The **serial hybrid** consists of an internal combustion engine coupled to a generator and one or more electric motors responsible for the drive. There is no mechanical connection between the combustion engine and the drive axle. The required energy is produced by the internal combustion engine, converted by a generator and used to charge the battery or directly to operate the electric motors depending on the requirements (Hilgers 2016).

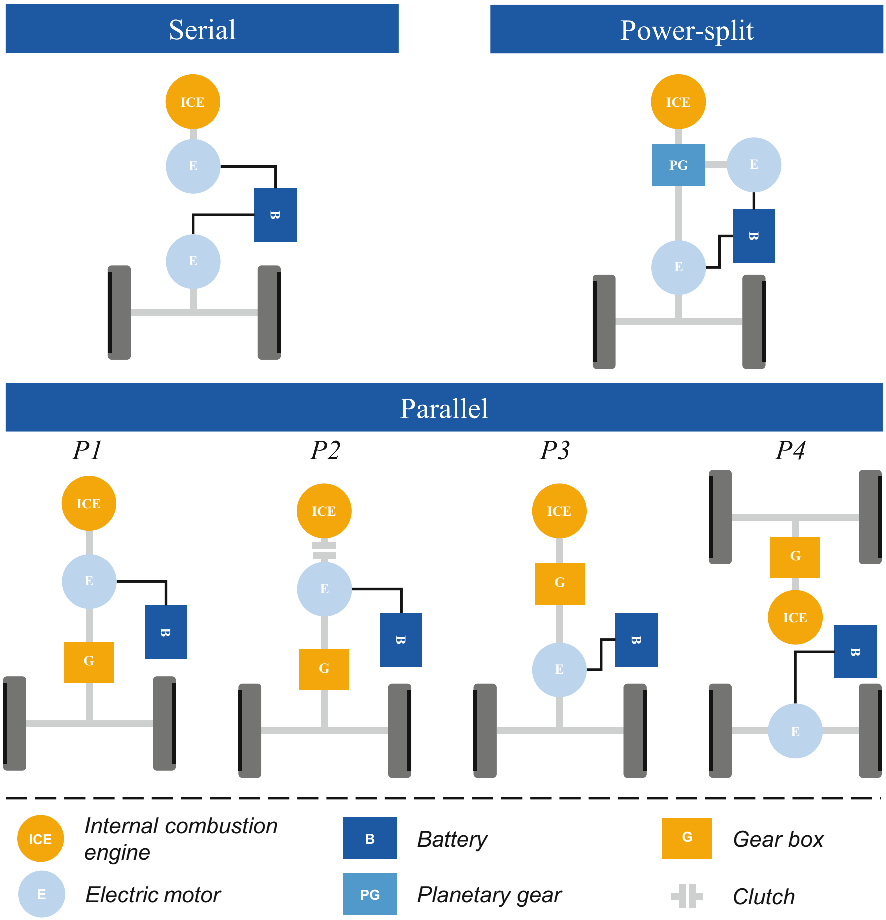


Fig. 2 Overview of powertrain topologies

In the so-called **parallel hybrid,** the combustion engine is located parallel to the electric motor so that the power of both systems can be used separately or superimposed. The electric motor of a parallel hybrid can be positioned at different locations within the powertrain. For the designation of the different positionings, the designation P1, P2, P3 and P4 of the Daimler AG has become generally accepted (Hilgers 2016).

In a **power-split hybrid**, the power of the combustion engines is divided into a mechanical and an electrical path. Hereby a part of the energy is directly transmitted to the drive axle. The other part drives a generator, which either feeds the battery or an additional electric drive motor (Görke 2016).

The variety of variants resulting from hybridization and the significant change in the product imply a series of challenges and trends in the production of road vehicles (Fig. 3).

	Product	Process
Trends	• Autonomous driving • High amount of variants with small quantities • Fully connected vehicles	• Intelligent resequencing in assembly • Early integration of powertrain components into the vehicle • Realization of the fixed-point-lose/ mobile assembly • Self-propelled chassis in production • Use of Augmented Reality for Agile Low-Cost Assembly • Intelligent production halls
Challenges	• Increased complexity during commissioning • Short development times	• Scalable production systems • High-voltage safety in assembly • Short ramp-up times • Resource flexibilization • New manufacturing processes unknown to OEMs

Fig. 3 Field of action: complete vehicle

The trend and the spread of autonomously driving, fully connected vehicles causes many changes both on the process and product side. In production and assembly, the early integration of the powertrain into the vehicle allows a fixed-point free assembly. This allows existing inflexible manufacturing structures to be replaced. In addition, the use of self-propelled chassis within the vehicle production can partially eliminate the need for cost-intensive investments such as the conveyor system. The automated and intelligent navigation of the vehicles within the production environment can be modeled by the vehicle's internal sensor technology, which means that precise sensor commissioning must be ensured in early process phases. Instead of using the internal sensor technology, there are approaches to equip the production hall with sensor technology and optical measuring equipment. In this way, navigation can be organized in a centralized manner, so that even variants without the corresponding sensor technology can be maneuvered autonomously through the production hall. An intelligent production control system can for example by resequencing the assembly processes, reduce model and variant mix losses and thus level out station utilization.

New vehicle structures and materials also bring fundamental process changes both in assembly and in the preceding processes. Novel material mixes imply different handling and joining processes that are partly unknown to the OEM. The increase of the system voltage from 12 to up to 800 volts requires new concepts in the area of safety and employee training, both in battery production and in the final assembly of the entire vehicle.

To support the employees efficiently and purposefully in new processes, augmented reality is used among other technologies. The implementation of such new methodology offers added value especially considering the high number of variants with a low amount of repetitions. Augmented reality can be used especially in the shorter and more frequent product launches. In addition to qualified employees, scalable production systems are also required that can be operated cost-efficiently in a variety of different operating modes. A particular challenge is currently the flexibilization of the production resources.

2.2 Field of Action 2: Battery Systems

For electric vehicles, high-voltage **battery systems** are essential, which production is currently highly complex and cost-intensive. The battery systems are basically made up of battery cells, which are connected to battery modules, a defined number of which are installed into a battery pack. The battery pack is then installed in the vehicle. The striving for cost reduction with the aim of achieving cost parity with vehicles with conventional drive trains represents one of the major challenges for the success of electrified vehicles in the future. In today's systems, the battery system is the biggest cost driver of the electric drive train. Against this background, battery cells particularly play an important role, since they account for about 70% of the total costs of the battery system. As a positive market development, it can be noted that the manufacturing costs for lithium-ion battery packs have fallen constantly in recent years. A further cost reduction of 62% compared to 2014 is expected by 2025 (Pillot 2017). It can be assumed that in 2025 the battery pack of an electric car will be about as expensive as the engine and emission reduction technology of a comparable fuel-powered vehicle combined (Seiwert et al. 2015). Until then, however, further development potentials must be consistently identified, their development accelerated, and their implementation realized to achieve the desired target state of cost parity while at the same time achieving high quality and customer benefit.

Today's storage systems are largely based on lithium-ion battery cells, which, compared to alternative cell chemistries, are mainly used in electric vehicles due to their high volumetric and gravimetric energy density (Fig. 4). In this sense, the lithium-ion cell represents the smallest unit of a battery system (Kampker 2014). Three main battery cell formats have become established in the production of energy storage devices for automotive applications: the pouch/flat cell, the cylindrical/round cell and the prismatic cell. The cells do not differ in their mode of operation but in their geometric shape and the production steps (Kampker 2014). The geometrical shape has a particular influence on the volumetric energy density. The energy density is decisive for the range of the electric vehicles and thus the capacity of the lithium-ion battery cell can be increased by an improved volume utilization. In general cylindrical cells currently have the highest energy density per unit volume compared to prismatic or pouch cells (Michaelis et al. 2018). The main components are the positive electrode (cathode), the negative electrode (anode), two current collectors, separator material between the electrodes, liquid electrolyte and the housing. The task of the electrolyte is to enable the transport of ions between cathode and anode. The ion permeable separator ensures that the electrodes are isolated from each other to prevent an internal short circuit (Kampker 2014). Individual lithium-ion battery cells are then electrically interconnected via a cell contacting system and assembled to form a battery module. In addition to electrical contacting, the battery

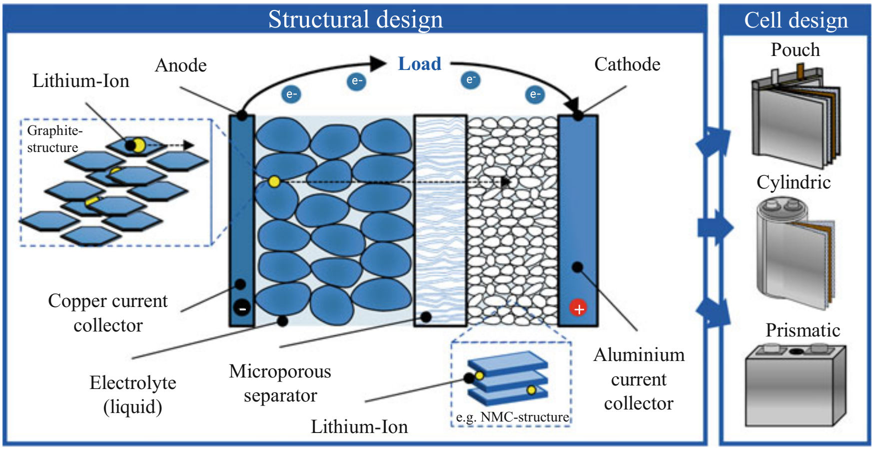


Fig. 4 Structure of a lithium-ion battery cell

module's universal tasks include tensioning of the cells, ensuring electrical insulation from the outside and dissipating any heat generated by the battery cells. Furthermore, it must be ensured that the frame concept of the module prevents any force from being applied to the cell contacting system, that any cell tolerances that occur are compensated and that mechanical integration in the battery storage tray can be guaranteed. The design of the battery modules depends strongly on the cell design used. In the case of prismatic cells, a module concept consisting of two pressure or end plates and two metallic tie rods on each side of the battery module is predominantly used in the automotive context. Finally, several battery modules are combined to form a battery pack, which is installed with peripheral system components. The following derivation of development trends and challenges will be considered first from a product and then from a process perspective (Fig. 5).

On the process side, there are efforts to increase the coating and drying speed in the context of battery cell production. Furthermore, the reduction of energy consumption in terms of environmental sustainability is playing an increasingly important role in battery production. Individual process steps, such as the drying process of the active material, still pose a challenge with regard to the energy efficiency of the production equipment. Innovative system technologies, such as the use of laser technology to dry the active material, can make a decisive contribution to reducing the ecological footprint of the battery cell manufacturing process.

2.3 Field of Action 3: Electric Drive

The **electric motor**, an electromechanical converter that converts electrical energy into mechanical energy, is not a new or unknown product, but the basic technology

	Product	Process
Trends	• Standardisation of the cells • Large format cells • High system voltage • Standardization of module formats • Increase of the fast charging capability	• Minimization of quality inspection • Creation of stable processes • Use of thin, innovative carrier foils/materials • Increase cell coating speeds • Dry coating in the cell production
Challenges	• Availability of materials, especially of cobalt • Market dominance in Asia • Increase of gravimetric energy density • Temperature control and adjustment of the operating area • Compensate aging processes with a suitable charging strategy	• Design for Remanufacturing • Solvent reduction in the coating process • Building a fair supply chain • Acceleration of formation and aging • Reduction of energy consumption within the production chain

Fig. 5 Field of action: battery

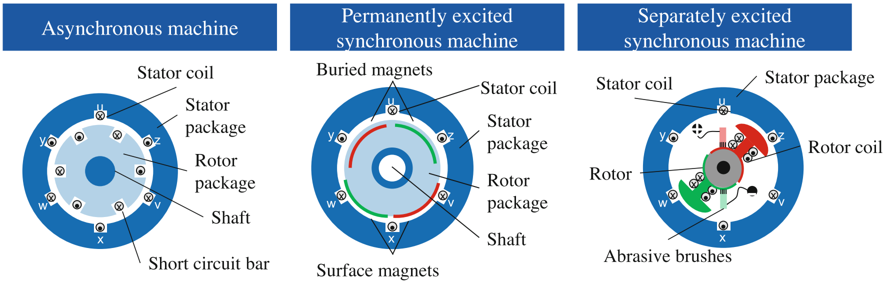


Fig. 6 Electric motor technologies

has been around since the late nineteenth century. In essence, a distinction can be made between direct current and alternating current or three-phase machines, whereby the alternating current machines have become generally accepted in the traction range due to their better efficiency, higher power density, lower costs and increased reliability (Hayes and Goodarzi 2018).

In practice, the main type of machines that are used for traction applications in relevant quantities are permanently excited **synchronous machines** (PSM), **separately excited synchronous machines** (FSM) and **asynchronous machines** (ASM) (Füßel 2017) (Fig. 6). In all three designs, the stator winding is supplied with three-phase current to generate the rotating magnetic field.

2.3.1 Asynchronous Machine

In the asynchronous machine, a change in the stator field induces current in the rotor's short-circuit bars, thus generating a counter-field. This generates torque between the stator and rotor. To maintain the self-induction, the rotor is always slower than the stator. The greater the required torque is, the greater is also the slip. Due to their relatively simple design, asynchronous machines are particularly cheap, but have the disadvantage that power density and efficiency are lower compared to the permanently excited synchronous machine, which makes increased cooling capacity necessary (Füßel 2017).

2.3.2 Permanently Excited Synchronous Machine

In the permanently excited synchronous machine, the rotor follows the magnetic field of the stator without slippage due to its magnetic field of constant polarity generated by permanent magnets: Accordingly, the speed of the rotor is proportional to the speed of the magnetic field and is therefore called synchronous speed. Due to its high power density and high efficiency, the synchronous machine is widely used despite the high cost of the required magnetic materials (Füßel 2017).

2.3.3 Separately Excited Synchronous Machine

In this design, the rotor of a separately excited synchronous machine contains a coil system which is supplied with the direct current required for excitation by slip rings. An electromagnetic field is generated in the rotor, which synchronously follows the rotating field formed by the stator. The simple and inexpensive design is counterbalanced by the wear and tear of the slip brushes (Bolte 2018).

The ongoing electrification of the global vehicle fleet is leading to an increasing importance of electric motors in the automotive industry. Many automakers have therefore integrated the production of these into their value chain to maintain the established level of value creation, differentiation, and quality leadership.

From the point of view of process and production technology, the requirements for electric traction motors have changed significantly, especially with regard to quality-relevant aspects.

Compared to the first-generation electric motors, the electric motors of the future will be produced on highly automated and strongly interlinked production lines. The production facilities are fully digitally connected and an exact link between process and product can be established. This allows stochastic processes and quality fluctuations such as winding to be optimized and better controlled (Fig. 7).

One technological innovation in the field of winding is the use of hairpin technology, in which the round wire is wound into the stator by means of shaped coils which are inserted into the stator by forming and assembly processes and then interconnected, for example by laser welding. Due to the highly interdependent production chain, the repeatable and dimensionally accurate production o**f hairpins** and the subsequent interconnection of these still poses challenges for system manufacturers and OEMs, particularly with regard to variant flexibility. The use of hairpin technology can also increase efficiency on the material side by reducing

	Product	Process
Trends	• Downsizing of engines and reduction of the power-to-weight ratio • Increase in material efficiency • Integration of subcomponents • High rotation speed concepts • High voltage motors (800V) • Unconventional e-motor topologies	• Highly automated and interlinked systems • Use of hairpin technology • Networked and digital production systems • Innovative groove insulation processes • Reduction of the cycle time • Stronger focus on joining technology
Challenges	• Renunciation of rare earths • Heat dissipation and cooling concepts • Efficiency increase on system and component level • Improvement of the NVH behavior • Resistance to ageing of impregnation and insulation • Reduction of the winding head	• Stochastic winding process • Variant flexibility, especially in stator production • Process stability in hairpin production and hairpin welding

Fig. 7 Field of action: electric drive

the winding head and at the same time increasing the groove filling factor, thus enabling the production of motors with a higher power-to-weight ratio. Due to the more compact design, but above all due to the integration of sub-components such as power electronics and reduction stage, the system is subject to an increased thermal load, which cannot be compensated for by increasing the individual efficiencies alone. Accordingly, new concepts are required for adequate heat dissipation, for example by the integration of cooled hollow rotor shafts or innovative thin-wall insulation to minimize the thermal bridge between the copper wire and the lamination stack.

The development of hairpinstator technology is currently characterized by cost-intensive equipment investments and high personnel expenditure already in the early product development phase. The Chair of Production Engineering of E-Mobility Components (PEM) of the RWTH Aachen University uses the so-called return-on-engineering approach for cost- and time-optimized development of products under consideration of a scalable production. In the production of hairpinstators, this can be particularly used to optimize machine investments to identify the optimal production line concepts and limiting factors as well as to uncover cross-process interdependencies.

In addition to the outlined field of action in stator production, there are efforts to reduce system costs sustainably, for example by reducing the magnetic material and alternative motor topologies such as axial flux designs and switched reluctance machines.

2.4 Field of Action 4: Vehicle Body

From an automotive perspective, the introduction of electromobility is characterized above all by the substitution of certain components in the drive train and the energy storage system. With regard to the vehicle architecture, this opens up completely new possibilities for the arrangement of these elements. This also has an influence on vehicle components that are designed based on this arrangement. The best example for this is the body of the vehicle.

The **body** of a vehicle forms the structural and design basis of a vehicle. In its classical passenger car structure, it is the basis for the cohesion of all vehicle components as well as for the protection of these and the vehicle occupants from environmental influences and in the event of a crash. In addition to protection, crash test safety and aerodynamics, the design basis for the body is also the geometric configuration of the interior vehicle components (Grote and Feldhusen 2014).

In addition to some special designs, these are usually designed in a self-supporting structure of deep-drawn sheet metal shells in the so-called shell construction method or extruded profiles, cast nodes and sheet metal forms in the space-frame construction method for vehicles manufactured in large series (Friedrich 2013). Self-supporting means that the chassis is integrated into the body and thus the body functions as a stand-alone element for stiffness and as the most important aggregate carrier (Grote and Feldhusen 2014). The body absorbs impact energies in the best possible way and diverts them for the protection of the passengers (Grote and Feldhusen 2014).

Within automotive production, the production of the body-in-white is classified, depending on the definition, as the first or second stage before the surface and either after or together with the press shop. The press shop is where the basic components are manufactured, which are mounted into the body-in-white using the three basic operations of handling, fixing and joining. Hereby the raw body follows a certain three-part architecture. The underbody production reproduces the chassis and is thus responsible for the most important stiffness provider in the body. The upper body or "hat" includes the side walls and the roof. In the third and last step, doors, hatches and fenders are mounted into the body (Haunstetter 2010). The body shop is generally highly automated with automation levels of over 90% and follows fixed production lines based on the herringbone principle. Subassemblies are manufactured in robot areas and joined to the overall body in the main line. Despite the resulting high investment requirements in the three-digit million range, car body construction is regarded by carmakers as a core competence with an in-house production depth of 60–100%.

With regard to **crash safety**, passenger protection and battery protection must be combined when introducing electromobility. The design focus in conventional vehicles is generally on crumple zones with the highest possible ductility for absorbing impact energy to keep the impact on the driver as low as possible (Pischinger and Seiffert 2016). If this approach is maintained in the case of electric vehicles, battery damage would represent another massive risk factor. This must be

	Product	Process
Trends	• Use of new materials • Lightweight construction using a multi-material mix • More degrees of freedom through modified product architecture • Integration of elements for tolerance compensation	• Multi-material design suitable joining process • Flexcells and island production • Joining with plug connections • Component-integrated fixtures • Use of thermoforming
Challenges	• Availability of materials • Crash safety of batteries • NVH behavior	• Tolerance management in fixtureless body construction • Complexity of the joining processes • Additional costs due to alternative joining methods • High variance in highly automated production

Fig. 8 Field of action: vehicle body

counteracted by the stiffest possible construction in the battery housing. While the battery can thus basically serve conceptually as the central component of the overall vehicle stiffness, this aspect is in competition with the typical arrangement of the battery and the passenger cell to protect both elements (Marx 2014).

On customer demand for long **range** e-vehicles with heavy batteries, the issue of **lightweight design** in body construction is increasingly becoming relevant (Fig. 8). Consequently, the conventional limits of steel-intensive shell construction and space-frame construction, which is characterized by the use of aluminum, increasingly merge into each other, and more and more bodies are being produced in the **multi-material design** (MMD) structure (Braess and Seiffert 2012). This means that, depending on the function of the car body at a particular location, different appropriate materials are used to intelligently reconcile crash safety and lightweight construction. While in the underbody aluminum cast nodes are supplemented by hot-formed ultra-high-strength steels, in the front of the vehicle high energy absorption in a frontal crash is achieved through high aluminum usage. In the body, stiffness is then generated by a shell construction with hot-formed steels, while components with only limited crash relevance, such as the vehicle hatches, are manufactured with carbon fiber composites (Schindler and Sievers 2008). At the same time, however, this poses major challenges for body shop production. The most common joining process in conventional steel body construction is resistance spot welding (RSW). This process is currently hardly applicable for multi-material joints. In general, thermal joining processes pose a problem due to widely differences in the melting points of different metals. Resistance spot welding is therefore increasingly being replaced by mechanical joining processes such as riveting, clinching, friction welding and flow-drill screws for steel-aluminum joints or gluing and screwing when joining polymers (Hage and Schulz 2009). This requires new corresponding process understanding and complexity management in the joining technology of automotive manufacturers.

While the component topology hardly varies at all in conventional drive trains, the e-drive components offer significantly greater degrees of freedom in the arrangement. This leads to a multitude of possibilities to design future vehicles more individually with regard to their application function. During the integration phase of the new electric vehicle technologies into the existing automotive market, a distinction is made between two different design principles. The simple replacement of conventional elements of the drive train with e-drives in an existing body structure is called conversion design. If a vehicle body is designed with regard to the e-drives and thus product-specific electromobility cost innovations are exploited, this is called purpose design. While **conversion design** vehicles have hardly any new requirements in terms of body construction compared to conventional vehicles and can therefore be integrated into existing body production systems, new body production facilities must be set up for **purpose design** vehicles (Kampker 2014). This additional variance is in line with a number of market trends, such as shorter innovation cycles, increased customer demand for individualization, global market uncertainties and the resulting fluctuations in unit numbers, which further intensify the challenges in body construction (Wemhöner 2006). Since automated production processes are usually difficult to change, tolerances in body construction must be ensured by variant-specific equipment such as picking tools and welding devices. With increasing variance, the cost and space requirements for these variant-specific resources also increase. To counteract this, the inflexibility of the production processes must be broken down, the variant-specificity of the operating resources must be eliminated and the operating resource expenditure must be reduced. A first applied approach in the automotive industry is the development of the production line structure towards a network of modular flexible production islands which are approached by different car body variants in different order depending on the joining sequence. Instead of a rigid conveyor system, driverless transport systems can be used for car body transport through flexible production equipment.

To reduce the fixture effort, several solutions are available in the outlook. If it is assumed that fixtures in car body construction have mainly the functions "shaping," "fixing" and "tolerance setting," then the functions "shaping" and "fixing" can be transferred to the component by intelligent implementation of connector elements and the function "tolerance setting" can be taken over by adaptive process technologies. Newly introduced joining processes, such as the one-sided remote laser welding technology, facilitate accessibility to joining points and can thus reduce a factor of complexity of fixtures. To reduce the design and production time of equipment, 3D printing of some fixture, tool elements or forming tools, for example, can be used, which can be quickly adapted to the respective variant in an automated intelligent design process.

2.5 Field of Action 5: Recycling

The environmental motivation behind electromobility and the resulting willingness and speed to innovate makes electromobility a potential pioneer of a resource-sustainable recycling economy in the automotive industry. Recycling management is defined as the decoupling of economic growth from the environmentally harmful extraction of resources required for the manufacture of products (Jung 2008). This is achieved by **recycling** of already used raw materials at the end of the product life cycle of the respective product, as well as the general extension of product life cycles through a basic structure that can be further developed and evaluated through **remanufacturing**.

Looking at the environmental impact of battery electric vehicles (BEVs) compared to internal combustion engine driven vehicles (ICEVs), it can be calculated that the emissions of BEVs are lower over the whole product life cycle of the vehicle. The prerequisite for this observation is the current average European energy mix, which is covered by approximately 30% from renewable energies (Electric vehicles from life cycle and circular economy perspectives 2018). At the same time, however, the influence of raw material and production loads is increasing with the progressive introduction of electric mobility (Kampker et al. 2016). The production of e-vehicles also requires certain raw materials, such as copper, lithium and cobalt, which are also needed in the manufacture of conventional cars, but are nevertheless available in limited quantities.

One approach to improving the ecological fingerprint is circular economy, the aim of which is the most complete possible energy and material recovery at the end of a conventional product life. This is in contrast to the linear economic production of previous vehicles. Based on the so-called utilization pyramid, there are numerous different levels of recycling and further reuse (Fig. 9). The basic idea is to extend the

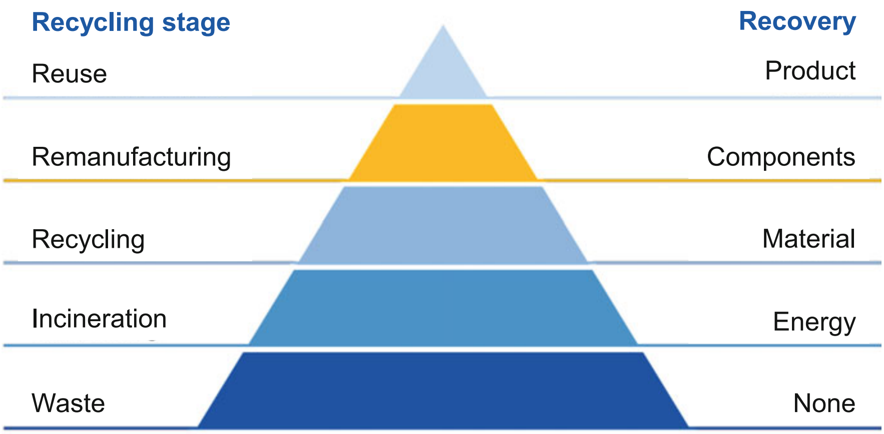


Fig. 9 Recycling pyramid (Kampker et al. 2016)

	Product	Process
Trends	• Early inclusion of the remanufacturing idea in the development • Creation of open interfaces for updates • Use of raw materials	• Efficient utilization of remanufacturing potentials • Implementation of remanufacturing and recycling in production systems • Development of new business models
Challenges	• Increased acqusition and development costs • Material mix trends and use of adhesives	• High stochastic wear variance • Lifecycle data collection to evaluate the end-of-life processes • Recyclability of new materials

Fig. 10 Field of action: recycling

product life and thus also has economic advantages, since the **total cost of ownership** (TCO) is reduced and the factor of high acquisition costs is put into perspective. This could also help the electromobility to achieve a higher market penetration (Kampker et al. 2016).

The potentially applicable methods can be summarized under the terms repair, reconditioning, refurbishing and remanufacturing. While repair is only used to restore functionality after a failure, reconditioning includes the proactive replacement of components that are close to failure. Both methods do not pursue the goal of achieving a new condition of the products. The refurbishing process addresses this at least visually, but does not serve to create a new state of functionality. The remanufacturing process has the goal of bringing a product into the new state both optically and functionally, including the associated warranty and product life extension (Fig. 10).

Great potential in electromobility arises from the smaller number of components in an e-vehicle and its flexibility in the arrangement in the vehicle due to lower mechanical complexity. In addition, there is the corresponding possibility of redesigning the vehicle. This comes into effect as soon as components have to be replaced or renewed in the context of the recycling cycle and thus easy accessibility is required (Kamper et al. 2018). A modularized construction is the best solution here.

The battery represents the largest cost factor among the components of an e-vehicle and at the same time one of the most resource-intensive. As an example, modular standardized battery concepts with generally less raw material input or even substitution technologies prove to be particularly suitable for the application of a circular economy. The first concepts in this direction were implemented by Nissan as early as 2016, through the sale of e-vehicles without batteries, which are then offered through a leasing contract. This gives the vehicle manufacturer the flexibility for simple to apply updates through new battery innovations without changing the basic vehicle. A prerequisite for this concept is the open design of interfaces, e.g., to the battery management system. Within the battery modules, there are also efforts towards the possibility of replacing individual battery cells that have failed without having to replacing the entire module (Kampker et al. 2016).

The idea of **remanufacturing** not only makes it easy to extend the product life of an electric vehicle, but also allows for continuous upgrading with newer technologies and designs. As already described, the general prerequisite here is the creation of open interfaces. For example, it would be conceivable to separate the powertrain from the passenger cell so that the passenger cell can be easily replaced during redesign without modifying the basic body, chassis or powertrain. Simplified assembly of external components, which allow easy disassembly and reassembly, not only offers the potential for simple repairs in the event of an accident, but also for upgrading the design while retaining the basic structure.

The use of alternative materials such as composite or hybrid materials, e.g., in the car body construction, has made recycling more difficult from a technological and cost perspective. On the other hand, these materials have a significantly longer service life and therefore form a promising basis for a corresponding basic framework, whose implemented components can be treated in terms of recycling management.

In addition to the basic functionality of the vehicle, the speed of innovation in information and communication technology also influences on the outcome of e-mobility. The upgrading of the vehicle platform, for example with the latest driver assistance systems, offers the opportunity to create new business models within the automotive industry, added to the ecological factors of the circular economy. In addition to this platform strategy, the concept of the environmental service branch also facilitates concepts such as shared mobility or commercial niche use as a fleet concept for e-vehicles offered directly by the automobile manufacturer. In addition, the constant flow of vehicles back to the manufacturer provides the latter with a data basis on, for example, wear and tear along the life cycle of a product, which can be directly incorporated into the product design (Kampker et al. 2016).

3 Summary and Outlook for Large-Scale Production of Electric Mobility

Electric mobility is becoming increasingly important for German end users. The automotive OEMs and the supplier industry behind them must now shape the change to the electric drive train to maintain their innovation leadership and defend their market share.

Since the first generation of electric vehicles from German OEMs is still represented in the market in small quantities, the knowledge gained from the development and production of these vehicles must be used to efficiently establish the second and third generation under large-scale production conditions in Germany as a high-wage location.

Based on the changing components from the conventional to the electric powertrain, the product and process fields of action are the complete vehicle, battery (systems), electric motors, vehicle body and recycling.

In the overall vehicle, the electric drive train in combination with the future implementation of sensors for autonomous driving will enable cost-efficiently mobile or freely linked final assembly in defined scenarios. By equipping the vehicles with the electric drive train and sensors at an early stage, the vehicles can drive autonomously through the assembly process. This reduces on the one hand the system costs, and on the other hand offers many degrees of freedom in resequencing, for example for capacity alignment.

The battery will remain the main cost driver of an electric vehicle. The high cost of the battery cell can be compensated by improved active materials and consequently higher energy densities at lower raw material costs. At the same time, standardization measures at cell, module and pack level can also reduce costs. Innovations in production technologies, for example to improve coating quality and reduce drying and aging times, also offer a high potential for quality optimization and cost reduction.

In the field of electric motors, there is a technological change from classic winding technology to so-called hairpin coils, especially in the stator production. These technological advances have the potential to offer efficiency benefits on the product side due to a higher copper factor, but also to further reduce production costs through highly automated production with automotive process stability. The new technologies present the OEM and supplier industry with the challenge of qualified manufacturing processes for large-scale production (Kampker et al. 2018).

The new arrangement of the electric-vehicle components also creates significantly more freedom in the overall design of the vehicle body. While this leads to new design possibilities on the product side, the resulting variance, in addition to the new material mixes used for lightweight construction reasons, presents the usually highly automated body shop with a multitude of complexity challenges. Flexible component and equipment design, intelligent automation and more flexible joining technologies and production structures must replace today’s inflexible manufacturing processes and equipment in car body manufacturing.

The ecological motivation and the significantly less complex vehicle topology in electromobility provide opportunities for a circular economy in the automotive industry. The use of resources increases with the introduction of e-vehicles through components such as the battery, the electric motor and the multi-material mix in the bodywork. However, while the battery is easily recyclable, the electric motor and the possibility of redesigning car bodies offer great potential for increasing the lifetime of vehicles. The corresponding consideration in vehicle development could in future turn electric vehicles into a life-long platform with the possibility of technology and design updates. From an economic point of view, this would not only improve the total cost of ownership and thus make electromobility more attractive on the consumer market but would also open up opportunities for completely new business models in the automotive sector.

References

Amelang S, Wehrmann B (2020) "Dieselgate" - a timeline of the car emissions fraud scandal in Germany. Clean Energy Wire - Journalism for the energy transition, Berlin

Becker J (2010) Geschichte der Elektroautos. Hg. v. Süddeutsche Zeitung

Bolte E (2018) Elektrische Maschinen. Springer, Berlin, Heidelberg

Braess H-H, Seiffert U (2012) Vieweg Handbuch Kraftfahrzeugtechnik. Vieweg+Teubner Verlag, Wiesbaden

Burggräf P, Kampker A, Nee C (2011) Elektromobilität wälzt Produktionstechnik um. In: *Mobilität Morgen* 1

Eberhard Abele (2009) Wandel im PKW-Antriebsstrang: Auswirkungen auf Produktionskonzepte. Unter Mitarbeit von Julien Hohenstein, Patrick Pfeiffer und Eric von Wihl. In: Maschinen- und Metallbearbeitung

Electric vehicles from life cycle and circular economy perspectives. TERM 2018: Transport and Environment Reporting Mechanism (TERM) report (2018). Luxembourg: Publications Office of the European Union (EEA report, No 13/2018)

Friedrich HE (2013) Leichtbau in der Fahrzeugtechnik. Springer Fachmedien Wiesbaden, Wiesbaden

Füßel A (2017) Technische Potenzialanalyse der Elektromobilität. Springer Fachmedien Wiesbaden, Wiesbaden

Görke D (2016) Untersuchungen zur kraftstoffoptimalen Betriebsweise von Parallelhybridfahrzeugen und darauf basierende Auslegung regelbasierter Betriebsstrategien. Springer Fachmedien Wiesbaden, Wiesbaden

Grote K-H, Feldhusen J (Hg.) (2014) Dubbel. Taschenbuch für den Maschinenbau : mit mehr als 3000 Abbildungen und Tabellen. 24. aktualisierte und erweiterte Auflage. Berlin: Springer Berlin

Hage M, Schulz R (2009) Verbindungstechnik im Multi-Material-Design künftiger Leichtbau-Karosserien. Lightweight Des 2(2):42–45. https://doi.org/10.1007/BF03223568

Haunstetter T (2010) Methoden des Data Mining in Anwendung an die Flexibilitätsattribute des automobilen Karosseriebaus zur strategischen Investitionsplanung. Hamburg, Helmut-Schmidt-Univ., Diss., 2010. Helmut-Schmidt-Universität, Bibliothek, Hamburg. Online verfügbar unter http://opus.unibw-hamburg.de/opus/volltexte/2010/2546

Hayes JG, Goodarzi GA (2018) Electric powertrain. Energy systems, power electronics et drives for hybrid, electric et fuel cell vehicles. John Wiley et Sons, Hoboken, NJ

Hilgers M (2016) Alternative Antriebe und Ergänzungen zum konventionellen Antrieb. Springer Fachmedien Wiesbaden, Wiesbaden

Jung G (2008) Kreislaufwirtschaftsland Rheinland-Pfalz. Hg. v. Gottfried Jung. Mainz: Ministerium für Umwelt, Forsten und Verbraucherschutz; Ministerium für Wirtschaft, Verkehr, Landwirtschaft und Weinbau

Kampker A (2014) Elektromobilproduktion. Hg. v. Christoph Nowacki. Springer Vieweg, Berlin

Kampker A, Kreisköther K, Hollah A, Lienemann C (2016) CoFAT 2016 - Electromobile Remanufacturing - Nutzenpotenziale für batterieelektrische Fahrzeuge

Kampker A, Triebs J, Hollah A, Lienemann C (2018) Remanufacturing of electric vehicles. Challenges in production planning and control. Hg. v. Global Conference on Sustainable Manufacturing

Marx B (2014) Crash-Sicherheit von Elektrofahrzeugen: Herausforderungen und Lösungsansätze. In: Gerhard Tecklenburg (Hg.): Karosseriebautage Hamburg. Wiesbaden: Springer Fachmedien Wiesbaden (Proceedings), S. 43–52

Michaelis S, Kampker A, Heimes H (2018) Roadmap Batterie-Produktionsmittel 2030. Hg. v. VDMA Batterieproduktion

Möser K (2002) Geschichte des Autos. Campus, Frankfurt am Main, New York

Nationalen Plattform Elektromobilität (2011) Zweiter Bericht der Nationalen Plattform Elektromobilität. Hg. v. Gemeinsame Geschäftsstelle Elektromobilität der Bundesregierung. Berlin

Pillot C (2017) The Rechargeable Battery Market. Main trends 2016-2025. Hg. v. Avicenne Energy

Pischinger S, Seiffert U (2016) Vieweg Handbuch Kraftfahrzeugtechnik. Springer Fachmedien Wiesbaden, Wiesbaden

Schindler V, Sievers I (2008) Forschung für das Auto von morgen. Aus Tradition entsteht Zukunft. Springer, Berlin, New York

Seiwert M, Hönscheid AT, Stölzel T, Menn A (2015) Autozulieferer - Angriff auf die Jobmaschine. Hg. v. Wirtschaftswoche

Wemhöner N (2006) Flexibilitätsoptimierung zur Auslastungssteigerung im Automobilrohbau. Zugl.: Aachen, Techn. Hochsch., Diss., 2005. Shaker, Aachen. Online verfügbar unter http://darwin.bth.rwth-aachen.de/opus/volltexte/2006/1496/pdf/Wemhoener_Nils.pdf

Role and Effects of Industry 4.0 on the Design of Autonomous Mobility

Ingrid Isenhardt, Thomas Otte, Anas Abdelrazeq, Alexia Fenollar Solvay, Christoph Henke, and Max Haberstroh

Contents

1 Introduction

The ongoing digital transformation affects all areas of our society and encompasses technical, organizational as well as social changes. This transformation not only results in an increasing data collection and data availability but also in an increasing

I. Isenhardt (✉) · T. Otte · A. Abdelrazeq · A. F. Solvay · C. Henke · M. Haberstroh
Institute for Information Management in Mechanical Engineering, RWTH Aachen University, Aachen, Germany
e-mail: isenhardt.office@ima.rwth-aachen.de; thomas.otte@ima.rwth-aachen.de; anas.abdelrazeq@ima.rwth-aachen.de; alexia.solvay@umlaut.com; christoph.henke@ifu.rwth-aachen.de; max.haberstroh@umlaut.com

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_30

utilization of this data—for example, when developing new digital business models and services. One area that was affected early on by the ongoing transformation is the mobility of people and freight (cf. logistics). In addition to the digital mapping of almost all traffic routes and the high availability of digital traffic information, the use of assistance systems in vehicles is driving this development as well as the emergence of e-commerce and new mobility services. In road traffic, this development is culminating in the implementation of the long-standing vision of self-driving vehicles, which is now imminent or has already become reality in tests and specific applications such as at airports or in intralogistics.

Although the term autonomous is strongly associated with self-driving cars in the context of mobility (cf. Süddeutsche Zeitung 2019), it generally describes a characteristic of independence. In principle, the term autonomous driving describes any driverless means of transportation in different areas of application. Correspondingly, the topic concerns not only road transport but also, for example, the rail-bound transport of passengers (cf. Flughafen Düsseldorf GmbH 2021), as well as the transport of goods within a production facility (cf. Ullrich 2011).

The aim of this article is not to provide a complete overview of the broad field of autonomous driving, but rather to outline the development of autonomous driving and the relationship to Industry 4.0 using examples. First, in Sect. 2, the background of autonomous driving and autonomous mobility is highlighted and placed in the context of Industry 4.0. Section 3 discusses the effects of Industry 4.0 on mobility and logistics. The transferability to other fields of application is then presented in Sect. 4 using three industries as examples—these are intralogistics, the construction industry and an air transport system for individual passenger transport. The article concludes with a summary and outlook.

2 From Industry 4.0 to Mobility 4.0

During the First Industrial Revolution and as early as the end of the eighteenth century, the transportation of people and goods by motor vehicles experienced its beginnings. This development of the automobile along with the industrial revolutions is summarized in Fig. 1. Starting from the original, functional benefit of locomotion, the automobile experienced continuous development and an increase in safety, comfort, and efficiency.

In the twentieth century, in addition to continuous improvements regarding the basic functions of the automobile (e.g., the powertrain), there were further developments particularly in the areas of information and communication technologies (cf. ICT). Consequently, various driver assistance systems were introduced to support drivers in performing their driving tasks at different levels.

Along with the Fourth Industrial Revolution and an expansion of the collection, availability, and processing of data, as well as the overarching networking of various technical and non-technical systems, there has been an ever-increasing expansion of the connectivity of vehicles and the sensors required for the sensing and recording of

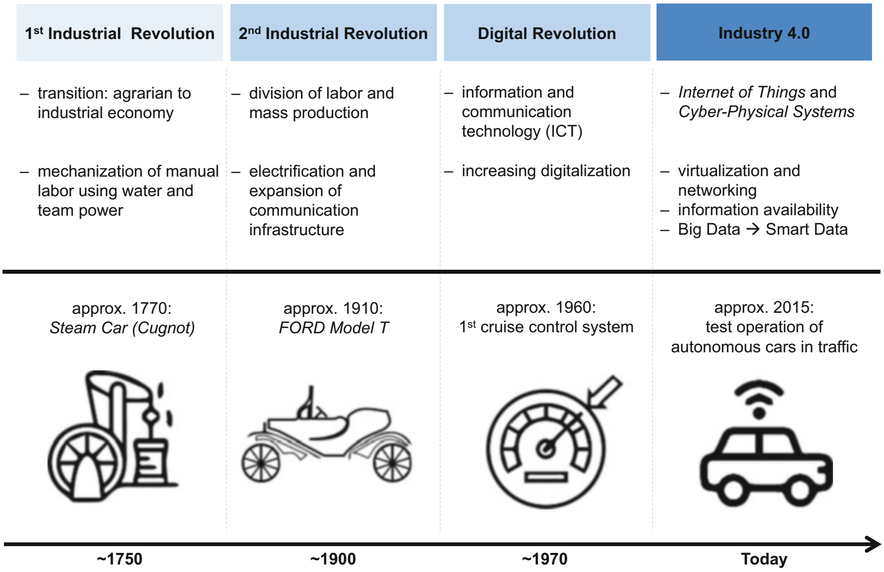


Fig. 1 Development of the automobile along with the industrial revolutions (own illustration based on Jeschke et al. 2017)

system states and environmental conditions in recent years. In combination with the (further) development of algorithms for processing the resulting data, this development has led to advanced driver assistance systems that can support the driver in almost all aspects of the driving task or completely relieve him of it.

The technical feasibility of (partially) autonomous mobility has been investigated and demonstrated in several research and development projects in recent years. Between 1996 and 2003, for example, the Promote Chauffeur I and II projects dealt with the electronic follow-up driving of trucks. In 2009, the KONVOI project, which was funded by the German Federal Ministry of Economics and Technology (BMWi—now the German Federal Ministry for Economic Affairs and Energy), realized first test drives with electronically coupled trucks in real traffic. Following this first successful demonstration of the feasibility of electronically coupled convoys journeys on German freeways, the topic was taken up in the European context in the SARTRE project and funded by the European Commission (cf. European Commission 2021).

In the passenger car sector, the Grand Challenge organized by the Defense Advanced Research Projects Agency (DARPA) of the U.S. Department of Defense between 2004 and 2007 contributed to the further technical development of autonomous vehicles (cf. DARPA 2019). After the last DARPA Grand Challenge in 2007, the development of autonomous vehicles was pursued not only by established automobile manufacturers but also by companies that had previously played no direct role in the automotive industry. For the development of its self-driving car, for example, Google hired the scientist Sebastian Thrun, who had previously led the

	SAE LEVEL 0	SAE LEVEL 1	SAE LEVEL 2	SAE LEVEL 3	SAE LEVEL 4	SAE LEVEL 5
What does the human in the driver's seat have to do?	You are driving whenever these driver support features are engaged - even if your feet are off the pedals and you are not steering.			You are not driving when these automated driving features are engaged - even if you are seated in 'the driver's seat'.		
	You must constantly supervise these support features; you must steer, brake or accelerate as needed to maintain safety.			When the feature requests, you must drive.	These automated driving features will not require you to take over driving.	
	These are driver support features:			These are automated driving features:		
What do these features do?	These features are limited to providing warnings and momentary assistance.	These features provide steering **OR** brake / acceleration support to the driver.	These features provide steering **AND** brake / acceleration support to the driver.	These features can drive the vehicle under limited conditions and will not operate unless all required conditions are met.		This feature can drive the vehicle under all conditions.
Example Features	automatic emergency braking blind spot warning lane departure warning	lane centering **OR** adaptive cruise control	lane centering **AND** adaptive cruise control at the same time	traffic jam chauffeur	local driverless taxi pedals / steering wheel may or may not be installed	Same as level 4, but feature can drive everywhere in all conditions.

Fig. 2 Levels of driving automation following SAE J3016 (cf. SAE 2018)

Stanford Racing Team that won the DARPA Grand Challenge in 2005 and came in second in 2007. In another example, Tesla Motors attracted particular attention with its announcement that it would offer its Autopilot as a downloadable add-on function for its customers, thus launching a beta test phase for autonomous driving, which was unusual for the automotive industry at the time.

To be able to classify different types of assistance systems that can take over parts of the driving task, the SAE (abbreviation for "Society of Automotive Engineers") has developed definitions for different levels of driving automation. As a distinct delimitation of the individual (partially) autonomous driving states can have decisive implications for various areas of society—including jurisdiction—institutions such as the German Federal Highway Research Institute (BASt) have joined professional associations such as the SAE in a dedicated project group for a legally secure definition of autonomy levels in autonomous driving (cf. Gasser et al. 2012). The resulting gradations start at level 0 (SAE) or level 1 (BASt) with no automation, continuing up to complete automation of all tasks in every possible driving state (SAE level 5) or the takeover of the lateral and longitudinal guidance by the system in defined use cases (BASt level 5) accordingly. For illustration purposes, the roles of the driver as well as the tasks that are taken over by the technical system along with the assigned automation levels are presented with examples in Fig. 2.

3 Effect of Industry 4.0 on Mobility and Logistics

The developments described above are going to influence mobility and logistics far beyond the technical design of future vehicles. Even if this change is not sudden, as the term revolution suggests, but rather gradual, the transformation is still profound, sustainable, and comprehensive. The sum of incremental changes is leading to an expansion of the tasks that vehicles can perform without the intervention of a driver. Moreover, in addition to technical components, mobility and user behavior are also changing (e.g., through the introduction of new mobility services and offers) and thus the entire mobility sector, the traffic infrastructure (e.g., through traffic lights that communicate with approaching vehicles), and the business as well as legal framework conditions (e.g., through new types of business models or emerging liability issues).

These developments are also corresponding to the Fourth Industrial Revolution, whose core is the intelligent networking of production resources, planning systems and manufacturing automation systems to make production processes more efficient, orienting them more closely to current customer requirements and being able to react adequately and efficiently in the event of problems. Ultimately, developments can be summarized under the label 4.0 if they include two core aspects: (1) the digitization of all processes and (2) a complete networking of actors, infrastructure, and technical resources. The term *Mobility 4.0* can accordingly be used to describe the comprehensive digitization and inherent networking of technical and natural systems to create value concerning the transport of people or goods based on the intelligent merging of information. Increasing information transparency and the availability of information in real-time can make processes more flexible and responsive. Finally, comprehensive digitization and networking allow the usage of autonomous systems as they are not only equipped with the necessary technology themselves but also their networking enables the exchange of information and interaction with other systems necessary for safe and efficient autonomous operation.

In the following, the importance of the digital transformation in general and the development of autonomous driving in particular will be discussed for the context of passenger and freight transport.

3.1 Effect on Passenger Transport: Mobility 4.0

A change that goes hand in hand with the development and introduction of autonomous vehicles can be observed in the mobility behavior. Overall, mobility, supported by new mobility services such as car sharing (e.g., Car2Go, DriveNow, Cambio), ride-sharing (e.g., CleverShuttle) or ride-hailing (e.g., Uber), is increasingly perceived as a service (cf. Mobility-as-a-Service—MaaS). The introduction of autonomous vehicles is likely to reinforce this trend in two respects. On the one hand, the cost-effectiveness of MaaS services will increase due to savings in personnel costs and higher availability, fewer accidents, the elimination of social periods and rest periods, and more efficient trip planning thanks to intelligent algorithms. On the other hand, MaaS offers to eliminate additional direct costs for the user, e.g., for purchase, repair, insurance, or parking space (even if these costs are at least partially passed on to the user). In addition, actual door-to-door mobility is made possible without, for example, the need to search for a parking space in private transport or the waiting times and transfers in public transport.

This change in mobility behavior in turn has a direct impact on the established providers inside the market. On the one hand, automobile manufacturers can expect further declining sales figures for individual vehicles (cf. VDA 2019) and a changed customer profile. In the future, MaaS providers will play a more important role as customers of car manufacturers, which means a new type of requirement situation for manufacturers in the next step. For example, Uber already approached the Daimler AG in 2016 with a request to buy 100,000 Mercedes Benz S-Classes with the condition that they drive autonomously (cf. Freitag 2016). On the other hand, traditional mobility services such as public transport and taxis are now competing not only with private transport but also with new market participants and business models, as it can be observed in the cases of Uber and CleverShuttle, for example.

3.2 Effect on Freight Transport: Logistics 4.0

The transfer of Industry 4.0 and its underlying principles poses a major challenge to all players in the logistics industry, as it can fundamentally change traditional logistics processes, services, and processes (cf. Lierow et al. 2014; Vogel-Heuser et al. 2017; Otte et al. 2020a, 2020c, 2021, 2022). Following the definition of Industry 4.0 (cf. intelligent networking of machines and processes for industry with the help of information and communication technology—BMWi 2021), Logistics 4.0 can be defined as follows: stakeholders (cf. people) such as shippers, transportation companies, logistics and mobility service providers, or customers attempt to transport goods (cf. objects) from an origin to a destination using various vehicles and handling equipment (cf. systems), forming a transportation network. To meet the full definition of Industry 4.0, the actors and processes of the transportation network must be dynamically connected, real-time optimized, and self-organized.

The transformation towards a Logistics 4.0 framework follows two major trends. On the one hand, the establishment of communication and network standards (cf. Doll et al. 2014) as the basis for large-scale information availability and data exchange for the automation of processes. On the other hand, the creation of service-oriented collaboration platforms and thus transparency to support integrated procurement, traffic planning and the resulting traffic flows (cf. Ten Hompel et al. 2014; Lumatrak 2019). Aspects influencing this transformation are presented in Fig. 3.

The digitization and automation of the supply chain are leading to a greater convergence of all sectors of a supply network and are having a particular impact on players such as manufacturers and freight companies. The ability to automatically

Fig. 3 Key terms in Logistics 4.0

derive the optimal value creation flow at any point in time from the data is strengthened by the further development of machine communication and self-learning systems and offers modularization and flexibility of the processes (cf. Jeschke et al. 2017). This availability of information at the right place and at the right time is the key to an efficient use of all resources and is therefore also of particularly relevant for the logistics domain.

Regionalization and decentralization of production processes create new demands on logistic systems, as the boundaries between production, transport and logistics are being reduced. In this context, decentralization describes systems that cooperate autonomously with each other and can thus independently develop solutions to problems (cf. Jeschke et al. 2017). Examples of this decentralization already exist today through the establishment of artificial intelligence, machine-learning systems, or multi-agent systems (e.g., for traffic planning and route optimization).

4 Transferability to Industry-Specific Fields of Application

The digital transformation is not limited to individual sub-areas of our society, which means that progress in one area influences developments in other areas. Like Industry 4.0, the development of autonomous vehicles is not only changing the existing traffic situation and mobility behavior but is also acting as an innovation driver in other areas. Often, challenges arise in these other areas of application that have not yet been considered or have been considered too little. Sometimes in this course, however, solutions emerge that in turn drive progress in the adjacent domains.

For the relationship between Industry 4.0 and the development of autonomous vehicles, a variety of other fields of application are of importance, out of which a selection is briefly presented in the further course of this section. Correspondingly, the aim of the following is not to provide a complete representation, but rather to demonstrate how the digital transformation is progressing through an overarching and reciprocal interaction of a wide variety of areas.

4.1 Intralogistics

The use of the so-called automated guided vehicles (AGVs) has been already a common practice in intralogistics for several years. Up to now, however, they have mostly been used in well-defined environments separated by protective fences, for example. For instance, track-bound systems (e.g., via induction loops or guiding lines) ensure the material flow in certain areas of intralogistics. These systems are currently being gradually replaced by trackless systems, which perform the material flow without infrastructural interventions such as the insertion of markers or lines, mainly based on optical sensors in the vehicle (e.g., laser scanners or cameras). With the significantly higher flexibility achieved in this way, the complexity of the task also increases.

The RoboCup Logistics League, for example, serves as a test environment for the development and research of autonomous systems cooperating in teams within dynamic environments. In the intralogistics scenario, two teams, each with three autonomous mobile robot systems, compete against each other in an Industry 4.0 testing setup (cf. Niemueller et al. 2013). The goal is to map the material flow between distributed modular production systems (MPS) using mobile robots, whereby these automatically plan and implement the material flow. Furthermore, the systems must detect environmental conditions and derive actions based on heterogeneous sensor technologies (cf. Niemueller et al. 2017).

In addition, current challenges in intralogistics lie in the introduction of the systems in situations with mixed traffic, in which AGVs share the available space, for example, with people or manually controlled forklifts. The navigation systems of the AGVs must be able to efficiently avoid dynamic obstacles within the specified driving areas and return to the original route. To ensure safe operation in passenger transport, the use of safety laser scanners, among other things, has been indispensable up to now. However, off-the-shelf systems on the market are currently only able to record the surroundings in two dimensions, so that overhangs cannot be detected reliably by the system. It is usually up to the operator to ensure that the surroundings are clear. The current harmonized Type C machine safety standard (cf. DIN 1997) regulates the requirements on the systems for safe operation but will be replaced by the Type C standard (cf. DIN 2018), which is currently in draft form, following harmonization.

A particular challenge in intralogistics is posed by mobile manipulators for mobile handling, which is not sufficiently considered in the current standards. The problem is that these systems are intended to perform handling tasks (e.g., for order picking or machine loading) outside structurally separated areas. This means that the systems are used in the close proximity of people and must therefore demonstrate functional safety. Today, there are design guidelines for manipulators for collaborative operation near people and, as already mentioned, for the use of AGVs. However, their combination has not yet been sufficiently recorded and supported by design guidelines to enable direct use of these systems without major restrictions and losses in efficiency. On the one hand, the movement of the manipulator while the AGV is moving and the gripping of objects with different geometric characteristics pose a particular problem.

4.1.1 Use of Autonomous Systems in e-Vehicle Assembly

Particular potential in the realization of material flow with AGVs arises in electric vehicle production. Since the vehicle is already functional at an early stage during final assembly, the vehicle itself can already be integrated into the material flow as a freely navigating AGV, e.g., through additional sensors attached to the vehicle, guided by external sensors (e.g., cameras in the environment) or using the driver assistance systems available in the vehicle. In this way, a more flexible assembly process can be implemented.

4.1.2 Use of Autonomous Systems in Order Picking

In the field of order picking (e.g., in department stores or in work preparation), two approaches to the use of mobile and stationary robot systems are currently being pursued—robot-to-goods and goods-to-robot. The first approach focuses on the use of AGVs, which provide an order picker, e.g., with a goods carrier (cf. Amazon—Kiva Systems, cf. Guizzo 2008) or GreyOrange (cf. Crowe 2018) and the order picker accordingly picks a single positions to put together an overall order. The goods-to-robot approach focuses on the provision of goods to a stationary robotic system that performs the pick. For this purpose, the goods must either be available in sequence or be actively separated or, alternatively, the robot must be able to recognize objects, derive gripping spots and perform a pick using additional sensors. The approaches mentioned here are currently mainly carried out outside of passenger transport and are usually physically separated.

4.1.3 Outlook: Mobile Manipulation for Stationary Handling

The current introduction of mobile, freely navigating robot systems in production and intralogistics has brought about an initial break with rigid process steps.

However, these systems have so far only carried out simple material transport steps to fixed workstations and robot cells. To continue this trend and to increase flexibility concerning the purpose of mobile robot systems, mobile manipulators are increasingly being considered. These are mobile robot systems equipped with industrial robots as handling units. The first manipulators are already being used in the field of order picking (e.g., *Magazino* (cf. Ackermann 2016), Fetch Robotics (cf. Ackermann 2015), *Kuka KMR iiwa* (cf. Dömel et al. 2017)). The advantages of these systems lie on the one hand in their ability to handle, assemble, and process workpieces from any location, and on the other hand in the inclusion of travel time in the value creation process through the ability to carry out assembly processes or quality assurance checks during the transportation task.

4.2 Construction Industry

Compared to stationary production processes and other fields of application of autonomous vehicles, the construction industry must cope with special challenges. Construction projects are, among others, characterized by:

(1) a high diversity of variants
(2) complex and highly interdependent work processes
(3) the use and integration of a wide variety of machine types from different manufacturers and product generations
(4) strong interaction with people
(5) a continuous change in site topology and a highly polluted environment

The digitization and networking of construction sites through cyber-physical systems consequently requires the further development and testing of existing technologies with special consideration of these challenges (cf. Otte et al. 2020b; Zhou et al. 2020). More specifically, current research and development projects are looking, for example, at site-specific approaches for machine control as well as assistance and automation systems, interface and interoperability management, and the integration of autonomous construction machines into the overall context of the construction site. The development of suitable operating strategies, the adaptation of task and process planning as well as the development of approaches for digital mapping and navigation under the mentioned conditions are important research questions.

When implementing assistance and automation systems to the machine control, the safety and efficiency of the construction process are in the foreground. To a strong extent, particularly the efficiency has depended on the expertise of the machine operators up to now. For example, the driver's influence on fuel consumption during handling work with a wheel loader can be amounted up to 50% (cf. Jacobs 2013). For successful process optimization through semi-autonomous construction machinery, this influence must either be considered or minimized through the use of assistance or automation systems.

The introduction of cross-manufacturer interface and communication standards as a fundament for successful interface and interoperability management often fails in construction processes due to the high heterogeneity of machines, manufacturers, users, and clients, which continues to further increase due to globalization (cf. Kramer et al. 2019). In addition, the often-lengthy development process of standards, especially in the international context, makes it almost impossible to keep pace with the generally observed shortening of innovation cycles. To nevertheless allow the necessary flexibility in the realization of construction projects, modular approaches and intermediary systems with clear interface definitions are therefore necessary.

4.3 Passenger Air Transport

Ground-based transportation infrastructures (e.g., roads or rail systems) have systemic capacity limits. One approach to overcome the existing limits is to develop the third dimension for individual passenger transport. Until now, the possibilities for an extensive expansion of the existing transport system with individual and automated passenger air transport services has been limited by several central challenges. In the context of the Fourth Industrial Revolution, it now appears that some of these key challenges may be solved. For the introduction of an autonomous passenger air transport service (e.g., in Germany), the following three fields of action, among others, are of central importance:

4.3.1 Technological Development of Autonomous Aircraft

This field of activity relates to the development of individual technical components as well as the entire autonomous aircraft. The research and development departments of various international companies are currently working on this topic, including Volocopter, Lilium, Airbus and Ehang, as well as Aston Martin and Audi. The existing debate in industry and academia is contributing to the further development of the individual sub-areas of autonomous passenger aviation, thereby laying the foundation for a future, comprehensive introduction of autonomous aircraft into the national or supranational transport network.

4.3.2 Interactions Between People and Technology

Another field of action arises from the fact that autonomous aircraft will operate the flight without pilots and (possibly) without additional personnel to accompany the flight, while the passengers typically do not have the necessary knowledge to intervene in emergencies. This leads to special requirements regarding the interaction between humans and machines within the aircraft and raises questions about, for example, which inputs can be made by a passenger, which information is provided by the aircraft during the flight, how potential emergencies should be handled or how the technical systems should be designed to maximize the well-being of the human passengers. These aspects have been conceptually investigated in the course of the IndiLuV research project (cf. BMBF 2018).

In addition to this individual component, an expansion of existing mobility services to include autonomous aircraft must clarify whether, to what extent, and under what conditions urban areas can be overflown by them. From a societal perspective, the question of how autonomous aircraft can be regulated and operated in urban airspace must therefore be clarified.

4.3.3 Integration Into the Existing Transport Infrastructure

In addition, the integration into the existing transport network will be a central field of action in course of the introduction of this type of mobility service. The possibility of creating new direct connections between urban agglomerations can result in gains for passengers in terms of travel time and travel comfort. To leverage these potentials of the overall system on multimodal door-to-door connections, transition points (so-called mobility hubs) between the different modes of transport as well as take-off and landing areas in urban areas must be developed. Promoting these approaches to expanding the existing transport infrastructure, the European Commission published its Flightpath 2050 development goals for aviation already back in 2011 (cf. European Commission 2011). According to this, by 2050, within Europe . . .

– . . . 90% of all travelers should be able to travel door-to-door within a maximum of 4 h.
– . . . a uniform ground infrastructure to handle the traffic volume should exist.
– . . . an air traffic management system can (additionally) handle autonomous air traffic.

In summary, it can be concluded that the sustainable transfer of such an innovative mobility offer into people's everyday lives requires more than a purely technical development (cf. Otte et al. 2018). It requires knowledge regarding the integrability into the existing transport network as well as regarding the interaction of people, organization, and technology. In addition to the direct interaction between passengers and aircraft, this particularly includes the interaction between the aircraft and society, which will also be looking for a way to deal with the new technology.

5 Summary and Outlook

Mobility and traffic, as well as the vehicles required for it, have always been closely linked to the social developments from which they emerge and which, for their part, play a decisive role in shaping them. This is particularly true for mechanically driven vehicles, which were and are not only the symbol but also the engine of the industrial revolutions in the literal sense of the word. In the development of Mobility 4.0, the core principles and key technologies, but also the challenges of Industry 4.0, can be found almost archetypically. This can be seen, for example, in the digitization of the entire traffic area, the ubiquitous availability of traffic information, or the use of autonomous vehicles in various areas of application. In particular, it is evident that the core principles of Industry 4.0, greater information transparency and networking, are leading to greater interlinking of different areas of application.

Accordingly, the development of autonomous vehicles is by no means limited to road traffic, but also extends to all other modes of transport and increasingly to other areas such as intralogistics or the construction industry. The example of autonomous passenger air transport clearly shows how far-reaching the current changes are. Similarly, it is not only technical progress that is driving the transformation but also changes in social requirements and the associated willingness to accept or at least discuss new solution approaches.

Since the fourth industrial revolution is by no means completed, but in full swing, it is almost impossible to predict which developments will ultimately be successful and how the sum of these developments will change mobility. However, there is no question today about the technical feasibility of autonomous vehicles on all modes of transport and their inherent potential to contribute to solving some of the central challenges of today's mobility sector. Thus, it can be assumed that autonomous vehicles will not only be more efficient and resource-saving in the medium term, but also safer than current solutions. In addition, there are economic factors such as potential savings in personnel costs and greater overall availability.

Largely unresolved are socially relevant issues such as the legal framework or the possibility of participation by all levels of society. In particular, there is still a lack of a conscious social opinion-forming process that does justice to the high pace of innovation and the complexity of current developments. Similarly, the core principles of Industry 4.0—information transparency and networking—could be applied.

6 Additional Notes

This book chapter embodies an English version of the following German article previously published by Springer:

Isenhardt I, Fenollar Solvay A, Otte T, Henke C, Haberstroh M (2020) Rolle und Einfluss der Industrie 4.0 auf die Gestaltung autonomer Mobilität. In: Frenz W (ed) Handbuch Industrie 4.0: Recht, Technik, Gesellschaft. https://doi.org/10.1007/978-3-662-58474-3_35

References

Ackermann E (2015) Fetch robotics introduces fetch and freight: your warehouse is now automated. IEEE Spectrum. Retrieved from https://spectrum.ieee.org/automaton/robotics/industrial-robots/fetch-robotics-introduces-fetch-and-freight-your-warehouse-is-now-automated

Ackermann E (2016) German warehouse robots tackle picking tasks. IEEE Spectrum. Retrieved from https://spectrum.ieee.org/automaton/robotics/industrial-robots/german-warehouse-robots-tackle-picking-tasks

BMBF (2018) Projekt IndiLuV. Retrieved from https://www.technik-zum-menschen-bringen.de/projekte/indiluv

BMWi (2021, May 4) What is Industrie 4.0? Retrieved from https://www.plattform-i40.de/PI40/Navigation/EN/Industrie40/WhatIsIndustrie40/what-is-industrie40.html

Crowe S (2018) GreyOrange raises $140M for warehouse robots. Retrieved from https://www.therobotreport.com/greyorange-raises-140m-series-c-warehouse-robots/

DARPA (2019) The grand challenge. Retrieved from https://www.darpa.mil/about-us/timeline/-grand-challenge-for-autonomous-vehicles

DIN (1997) Sicherheit von Flurförderzeugen: Fahrerlose Flurförderzeuge und ihre Systeme (No. DIN EN1525:1997-12)

DIN (2018) Entwurf Flurförderzeuge - Sicherheitstechnische Anforderungen und Verifizierung: Teil 4: Fahrerlose Flurförderzeuge und ihre Systeme (No. DIN EN ISO 3691-4:2018-05)

Doll A, Friebel D, Rückriegel M, Schwarzmüller C (2014) Global logistics markets. Retrieved from https://www.rolandberger.com/de/Publications/pub_global_logistics_markets.html

Dömel A, Kriegel S, Kaßecker M, Brucker M, Bodenmüller T, Suppa M (2017) Toward fully autonomous mobile manipulation for industrial environments. Int J Adv Robot Syst 14(4). https://doi.org/10.1177/1729881417718588

European Commission (2011) Flightpath 2050: Europe's vision for aviation: maintaining global leadership and serving society's needs. Retrieved from https://op.europa.eu/de/publication-detail/-/publication/296a9bd7-fef9-4ae8-82c4-a21ff48be673

European Commission (2021) Safe road trains for the environment. SARTRE. Retrieved from https://cordis.europa.eu/project/rcn/92577/factsheet/de

Flughafen Düsseldorf GmbH (2021) SkyTrain. Retrieved from https://www.dus.com/de-de/anreisen/bus-und-bahn

Freitag M (2016, March 18) Uber an autonom fahrenden S-Klassen interessiert: Daimler offenbar vor Megaauftrag von Uber. Manager Magazin. Retrieved from https://www.manager-magazin.de/unternehmen/autoindustrie/daimler-uber-soll-autonom-steuernde-s-klassen-bestellt-haben-a-1082886.html

Gasser TM, Arzt C, Ayoubi M, Bartels A, Bürkle L, Eier J et al (2012) Rechtsfolgen zunehmender Fahrzeugautomatisierung. Retrieved from https://www.bast.de/BASt_2017/DE/Publikationen/Berichte/unterreihe-f/2013-2012/f83.html

Guizzo E (2008) A look under the hood of Kiva Systems warehouse robots. Retrieved from https://spectrum.ieee.org/automaton/robotics/robotics-software/a_look_under_the_hood_of_kiva_systems_warehouse_robots

Jacobs G (2013) Energieeffiziente Antriebe für Bau- Und Landmaschinen. ATZoffhighway 6(1): 50–58. https://doi.org/10.1365/s35746-013-0068-0

Jeschke S, Hoffmann M, Fenollar Solvay A, Schouba P, Stahl M (2017) Studie zum Einzug der 4.0-Entwicklungen in die Transportlogistik-Branche mit aktuellen Megatrends und Handlungsfeldern

Kramer A, Jacobs G, Fenollar Solvay A, Lipp J, Gentz C (2019) FutureSite – test center for innovative construction site technologies and automation. Retrieved from https://www.vdi-wissensforum.de/en/forms/whitepaper/va/1274/

Lierow M, Janssen S, D'Incà J (2014) Disruptive logistics: the new frontier for E-Commerce. Retrieved from https://www.oliverwyman.com/our-expertise/insights/2014/sep/disruptive-logistics-the-new-frontier-for-ecommerce.html

Lumatrak (2019) How End-to-End Supply Chain Visibility (E2ESCV) adds tangible value. Retrieved from https://www.lumatrak.com/e2escv-faqs-the-basics-of-end-to-end-supply-chain-visibility-e2escv/
Niemueller T, Ewert D, Reuter S, Ferrein A, Jeschke S, Lakemeyer G (2013) RoboCup Logistics League sponsored by Festo: a competitive factory automation testbed, 8371:336–347. https://doi.org/10.1007/978-3-662-44468-9_30
Niemueller T, Neumann T, Henke C, Schönitz S, Reuter S, Ferrein A et al (2017) Improvements for a robust production in the RoboCup Logistics League 2016, 9776:589–600. https://doi.org/10.1007/978-3-319-68792-6_49
Otte T, Metzner N, Lipp J, Schwienhorst MS, Solvay AF, Meisen T (2018) User-centered integration of automated air mobility into urban transportation networks. In: 2018 IEEE AIAA 37th Digital Avion-Ics Systems Conference (DASC) Proceedings, 1–10. https://doi.org/10.1109/DASC.2018.8569820
Otte T, Fenollar Solvay A, Meisen T (2020a) The future of urban freight transport: shifting the cities role from observation to operative steering. (RWTH-2020-02971), 1–10. https://doi.org/10.18154/RWTH-2020-02971
Otte T, Zhou HA, Gannouni A, Odenthal J, Abdelrazeq A, Hees F (2020b) A manufacturer-independent digital process platform for future construction sites, 96–101. https://doi.org/10.1109/IIT50501.2020.9298971
Otte T, Gannouni A, Meisen T (2020c) The future of urban freight transport: a data-driven consolidation approach to accompany cities on their path to zero redundancy in freight movement. In: 27th ITS World Congress via 1st Virtual ITS European Congress (COVID-19), pp 1–10. https://doi.org/10.18154/RWTH-2020-06989
Otte T, Meisen T (2021) A reference framework for the performance-based decision support of city authorities in urban freight transport. In: IEEE 2021 8th International Conference on ICT for Smart Society (ICISS), pp 1–7. https://doi.org/10.1109/ICISS53185.2021.9533210
Otte T, Posada-Moreno AF, Hübenthal F, Haßler M, Bartels H, Abdelrazeq A, Hees F (2022) Condition monitoring of rail infrastructure and rolling stock using acceleration sensor data of on-rail freight wagons. In: Proceedings of the 11th international conference on pattern recognition applications and methods, pp 432–439. ISBN 978-989-758-549-4
SAE (2018) Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. Retrieved from https://www.sae.org/standards/content/j3016_201806/
Süddeutsche Zeitung (2019) Themen – Autonomes Fahren. Süddeutsche Zeitung. Retrieved from https://www.sueddeutsche.de/thema/Autonomes_Fahren
Ten Hompel M, Rehof J, Heistermann F (2014) Logistik und IT als Innovationstreiber für den Wirtschaftsstandort Deutschland: Die neue Führungsrolle der Logistik in der Informationstechnologie. Retrieved from BVL website: https://www.bvl.de/positionspapier-it
Ullrich G (2011) Geschichte der Fahrerlosen Transportsysteme, 1–12. https://doi.org/10.1007/978-3-8348-9890-6_1
VDA V. d. A. (2019) Zahlen und Daten. Retrieved from https://www.vda.de/de/services/zahlen-und-daten.html
Vogel-Heuser B, Bauernhansl T, ten Hompel M (2017) Handbuch Industrie 4.0 Bd.3 (2., erweiterte und bearbeitete Auflage). Springer Reference Technik. Berlin, Germany: Springer Vieweg. https://doi.org/10.1007/978-3-662-53251-5
Zhou HA, Gannouni A, Otte T, Odenthal J, Abdelrazeq A, Hees F (2020) Towards a digital process platform for future construction sites, 1–7. https://doi.org/10.23919/ITUK50268.2020.9303198

Ethics of Digitalization in the Automotive Industry Using the Example of Self-Driving Cars

Arne Manzeschke and Alexander Brink

Contents

At the TEDx SiliconValley Conference of 2011, Damon Horowitz, former Director of Engineering and Google's in-house philosopher, impressively demonstrated to his audience that many people are more familiar with the technical features of their mobile devices than with the ethical foundations of their own actions (Horowitz 2011).

In the interaction between man and car, it is possible to study vividly over time how the **man-machine relationship** is changing: from more instrumental use to a complex digital environment, which enables a multi-layered concept of mobility, logistics and life. Autonomous mobility and autonomous driving are now an integral part of modern society—even if they are still in their infancy. What was long considered fiction is increasingly becoming reality. With the help of cameras and sensors, autonomously driving vehicles will participate in road traffic. According to the German Association of the Automotive Industry, six **levels of automated driving** are currently defined: Level 0 (driver), Level 1 (assisted), Level 2 (semi-automated), Level 3 (highly automated), Level 4 (fully automated), and Level 5 (driverless) (VDA 2015, p. 15).

Global corporations such as Daimler see the future of **mobility** in intuitive mobility supported by technology, which is characterized by the term CASE (connected, autonomous, shared & services, and electric) (Daimler 2018). Vehicles that drive autonomously avoid human error, reduce the number of accidents on the

A. Manzeschke (✉)
Lutheran University of Applied Sciences - Evangelische Hochschule Nürnberg, Nürnberg, Germany
e-mail: arne.manzeschke@evhn.de

A. Brink
Bayreuth University, Bayreuth, Germany
e-mail: alexander.brink@uni-bayreuth.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_31

road, and save energy by optimizing traffic flow. In addition to preventing the negative effects of **mobility**, the various positive effects of autonomous mobility concepts for society are discussed in the context of so-called **shared value approaches**, which go beyond a mere increase in comfort for the customer. These include, for example, the integration of elderly or disabled people into road traffic, or the collection of data to optimize traffic flows. The mobility and logistics concept associated with this are understood by Winner (see also Manzeschke and Brink in this volume "Ethics of Digitalization in Industry") as a constitutional structure that decisively determines how we will live (well).

In the future, people will leave far-reaching decisions to self-driving cars. In certain situations these will be life-saving, such as automatic braking systems or distance controllers. Here, there are hardly any ethical differences. The situation is different in collisions, where different damage scenarios conflict with one another. In this case, it would have to be defined ex ante for an autonomously driving car how it should react—and this would have to be based on an ethical decision-making process. In fact, autonomous **mobility** is associated with the moral requirement to avoid accidents, injuries, and deaths. Without going into the technical, economic, or legal aspects (for a good overview, see Maurer et al. 2015), we want to concentrate on some specific ethical problems (BMWi 2017).

As soon as vehicles can be fully automated or even driverless, they have to replicate human decisions—including moral ones (Lin 2015, p. 69). One could argue that hard **ethical decisions** are not made, because the fully automated vehicle simply brakes in critical cases, or the driver takes the wheel (as would be possible at least for stages 1 to 3) and makes a decision based on their own ethical reflection. The more complex the situation and the more urgent the need for a decision, the less possibilities there are to refer to these two options. The Ethics Commission of the Federal Ministry of Transport, headed by Udo di Fabio, published a statement on automated and networked driving (BMWi 2017; Hevelke and Nida-Rümelin 2015). The Institute of Electrical and Electronics Engineers (IEEE) formulates a clear moral claim in its **Code of Ethics**: "to treat fairly all persons and to not engage in acts of discrimination based on race, religion, gender, disability, age, national origin, sexual orientation, gender identity, or gender expression" (IEEE 2018).

The fundamental ethical problem of autonomous **mobility**, which falls into the first category of ethical problems (see also Manzeschke and Brink, Ethics of Digitalization and Healthcare), can be illustrated particularly well by using the **trolley dilemma**, which is why we list it here (Cathcart 2013; Lin 2015). In trolley dilemmas, moral situations are described which cannot be solved satisfactorily, because conflicting moral claims cannot be convincingly weighed against each other (Ryle 1953). A solution is always at the expense of one of the two "cruxes" of the dilemma, i.e., a moral good, a moral norm or attitude that one wants to see maintained. **Trolley dilemmas** under the conditions of autonomous **mobility** are also of this kind. For example, in the case of an unavoidable collision for the autonomous vehicle, it is up to its artificial intelligence to "decide" who should be harmed and who should not: e.g., three young people (with more life expectancy and potential), two older people (less life expectancy, but more life experience), or the one occupant of the autonomous vehicle (self-protection, or number of accident victims).

It is worthy of concern with such a situation that we would not expect a human driver to resolve this situation morally convincingly. Rather, we would concede that they could not, in such a case, act morally at all. Provided they had not violated any duty of care, they would in such a scenario be at least morally, perhaps even legally, blameless. However, we would expect more from artificial intelligence under the same conditions. Since the course of the vehicle has to be determined, decision parameters must be applied. If a decision is to be made, it must be made in a morally convincing manner, or in such a way that our previous moral sensibilities are not too affected. The persuasive power of AI's "decisions" would have to be based on rules or arguments that we humans, as moral actors, would have to be able to understand and accept. Thus, one could program an AI strictly calculating from a consequentialist ethics point of view, which determines a certain good (life expectancy, life experience, merits towards society, or number of injured people) as a target value for the weighing process. In principle, the self-protection of vehicle's occupants could be given priority over the protection of other road users, or even the opposite (BMWi 2017). If we now expect AI to decipher a problem that is morally unsolvable for us humans, the solution in this case would be based on a bare subsumption, such as the rule "protection of others before self-protection" (or vice versa). But this rule does not offer any room for moral considerations, from which the responsibility of the employee results. In another case, the solution could be based on coincidence. Since there is obviously no truly moral solution, one could consistently demand that such a solution be created by a random generator – analogous to accidents caused by humans, in which chance also seems to prevail. The decision that AI makes in the respective situation, deciding which party should be harmed or protected from harm, is then accepted as the "right" solution. In this context, the question of mandatory or personal ethics is widely discussed in literature (e.g., Müller and Gogoll 2016). Theoretically, it would be conceivable that we humans would accept the results of AI as a moral rule—as a higher insight, so to speak. However, this would then have nothing to do with ethics and morals, but would be the acceptance of instructions, without further insight into the reasons.

Such a construction presumably arouses astonishment and emotional unease. It shows, first of all, that such a firmly programmed logic has little or nothing to do with ethics, which in such difficult situations offers freedom, allows the moral actor to fail, or dispenses them from moral demands. For artificial intelligence, by contrast, such a "solution space" seems unconvincing. It would have to find better solutions based on all the available data and the high speed of calculation and reaction. What can be programmed is a procedural rule which—often enriched by a few exceptional variants—must be executed.

Ex-ante programming, in which an individualized procedural rule is determined for a driver based on a **utilitarian**, **deontological**, or **virtuous-ethical argument**, seems unrealistic and impractical. This applies all the more so if several occupants with different moral positions are sitting in the vehicle.

Artificial intelligence could be allowed to “learn” from accident constellations about which decisions are ethically accepted. Instead of an explicit rule-based approach (e.g., self-protection takes precedence over external protection), implicit patterns are extracted from evaluated decisions from the training agenda. Apart from the technical and social problems that come with the collection of such training data, further ethical questions arise: it could be that algorithms do not always learn the best morally, and do not come to good decisions. In the case of Microsoft’s Chatbot Tay, which was supposed to “learn” communication using Twitter data, it has been shown how it developed into a racist bot in a very short time based on this data (Vincent 2016).

From an ethical perspective, the variants shown are unsatisfactory and hardly communicable. In the long run, AI-based “solutions” to such moral dilemmas could even damage our moral sense, and our sense of responsibility. The constellation is reminiscent of Paul Watzlawick’s “**ultra solution**,” in German (“Patendlösung”) it is a play on words combining patent and final solution, a solution that does away with the problem and just about everything else (Watzlawick 1980). In other words: The attempt to delegate moral problems to an artificial intelligence makes our own moral judgement diminish, and under certain circumstances, also eliminates morality.

This shows once again that none of the solutions outlined are really pacifying or morally convincing—even if majorities could actually be won for one option. From an ethical perspective, it takes the human being as moral authority to recognize and shape such developments.

A further aspect discussed among technicians, lawyers, and sociologists is the question of the extent to which autonomously decisive actors can be held liable for erroneous behavior. This is therefore not only a moral but also a legal **responsibility**. Does this mean that cars are autonomous actors? Certainly not as things stand today. The decision which is made by an autonomous car is ultimately programmed by man, and takes into account technical, legal, economic, and ethical aspects. The autonomous car is programmed to react in a way that also leads to an effect (e.g., in a positive case, human lives are saved), but the actual action is not attributed to the car, but a person. At the same time, this is only the current status, and upcoming changes will force us to think further. This is also shown by the initiative of the European Parliament, to regulate “ever more sophisticated robots, bots, androids and other manifestations of artificial intelligence” (EU 2015). Before the detailed questions arise from autonomous driving, the basic questions outlined here must be discussed: how we want to understand these artificial intelligences, and what status we will assign to them in **socio-technical arrangements**. Afterwards, the ethical questions are to be negotiated, which directly refer to the normative area: how do we deal with the loss of hundreds of thousands of jobs among professional drivers (trucks, taxis, rescue services, etc.)? Which invasions of privacy are associated with the permanent monitoring of vehicles and their occupants? What influence on social and individual life do organizations that shape our **mobility** in complex digital worlds through their digital infrastructure have? The quasi-monopolistic position of some **platform operators** is already proving to be problematic, and without appropriate regulation in the area of mobility, it would be expected that possible advantages (damage reduction, optimized traffic flow, lower land consumption) would be counteracted by considerable losses (in the areas of privacy, justice, but above all, moral judgement and moral "leeway"). Further ethical questions that can be expected from the combination of different data from different spheres (such as mobility, communication, and health) cannot be precisely formulated at present, but this does not mean they are not relevant.

References

BMWi (2017) Ethik-Kommission Automatisiertes und vernetztes Fahren. https://www.bmvi.de/SharedDocs/DE/Publikationen/DG/bericht-der-ethik-kommission.pdf?__blob=publicationFile. Accessed 5 Jan 2022

Cathcart T (2013) The trolley problem, or would you throw the fat guy off the bridge? Workman Publishing Company, New York

Daimler (2018). https://www.daimler.com/innovation/case-2.html. Zugegriffen am 5 Jan 2022

EU (2015) European Parliament on rules for robotics, 2015/2103(INL) – 16/02/2017 Text ad- opted by Parliament, single reading. http://www.europarl.europa.eu/oeil/popups/summary.do?id=1477231&t=d&l=en. Accessed 5 Jan 2022

Hevelke J, Nida-Rümelin J (2015) Selbstfahrende Autos und Trolley-Probleme: Zum Aufrechnen von Menschenleben im Falle unausweichlicher Unfälle. In: Sturma D, Heinrichs, B, Honnefelder L (Hrsg.) Jahrbuch für Wissenschaft und Ethik, Bd 19, 1. Aufl. de Gruyter, Berlin, S 5–24

Horowitz D (2011) We need a "moral operating system". https://www.ted.com/talks/damon_horowitz. Accessed 5 Jan 2022

IEEE | Institute of Electrical and Electronics Engineers (2018) IEEE code of ethics. http://www.ieee.org/about/corporate/governance/p7-8.html (2014). Accessed 5 Jan 2022

Lin P (2015) Why ethics matters for autonomous cars. In: Maurer M, Gerdes JC, Lenz B, Winner H (Hrsg) Autonomes Fahren – Technische, rechtliche und gesellschaftliche Aspekte. Springer, Berlin/Heidelberg, S 69–85

Maurer M, Gerdes JC, Lenz B, Winner H (eds) (2015) Autonomes Fahren – Technische, rechtli- che und gesellschaftliche Aspekte. Springer, Berlin/Heidelberg

Müller J, Gogoll J (2016) Autonomous cars: in favor of a mandatory ethics setting. Sci Eng Ethics 23(3):681–700

Ryle G (1953) Dilemmas. The Tarner Lectures 1953. Cambridge University Press, Cambridge

VDA | Verband der Automobilindustrie (Hrsg) (2015) Automatisierung. Von Fahrerassistenzsystemen zum automatisierten Fahren, Berlin

Vincent J (2016) Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist. Accessed 5 Jan 2022

Watzlawick P (1980) Ultra Solutions. How to fail most successfully. W. W. Norton Company, New York/London

Rail Transport Technology 4.0

Christian Schindler

Contents

The term **Industry 4.0** first appeared in an article about the Hannover trade fair in 2011 in VDI-Nachrichten[1] and referred to the so-called "fourth industrial revolution in production" (Kagermann et al. 2011). Since then, the suffix "4.0" has spread into many fields of technology and its original meaning has been partially changed or completed. According to Jaspernite et al., Industry 4.0 is the networking of the involved technical systems with each other and with people with the aim to better support them through digitization by information, enabling the technical systems to

[1] Weekly newspaper of the Verein Deutscher Ingenieure (Association of Germen Engineers).

C. Schindler (✉)
IFS, RWTH Aachen University, Aachen, Germany
e-mail: schindler@ifs.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_32

make independent decisions and to perform their tasks as autonomously as possible (Jaspernite and Niggemann 2012).

The preliminary version of the "Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0"[2] from the working group of the same name commissioned by former Federal Government dates from 2012. Among others, the term **Smart Mobility** describes a field of application from Industry 4.0. Here, mobility is comprehensively understood as an integrated consideration of passenger and freight transport and therefore also covers the area of logistics. A distinction between "Smart Mobility" and "Smart Logistics" is not made here and the terms are only explained to the extent that they deal with networked transport (Kagermann et al. 2012).

Eight months before the published report of the Promoter Group Mobility, neither the term "Industry 4.0" nor "Smart Mobility" appeared. However, in the sense of Industry 4.0, the **autonomization** and **networking** of transport systems and their means of transport aside from electrification is identified as a future challenge. Nevertheless, the focus of the study is on automotive vehicles, while water, air, and rail vehicles are only mentioned as a footnote (Kreimeyer et al. 2012).

In the brochure on the high-tech strategy 2014, in which the term "Industry 4.0" appeared for the first time in connection with the technology goals of the German Federal Government, the word "rail" appeared only once in connection with new vehicle technologies, while eventually a small sub-chapter was devoted to the modes of transport air and water (BMBF 2014).

Finally, in the federal government's **5-point strategy** the term "Smart Mobility" appeared in connection with the term service in the rail transport sector in a publication on the future of rail as a mode of transport. There was no mention of the term "Industry 4.0." However, the published speech was about **Mobility 4.0** and the characteristics of Industry 4.0 "automated" and "networked" were mentioned as essential goals within the 5-point strategy (BMVI 2016).

The term **Rail 4.0** was mentioned for the first time in the title of the innovation program, which several associations in the railway sector published as a recommendation for the previous federal government. There, Industry 4.0-related goals such as "digitizing infrastructure," "**digitalization** and **big data** in maintenance," and "automation" were named, some of which had already been requested 6 months earlier in the rail freight-transport master plan without referring to the term "X 4.0." (Möbius et al. 2017; BMVI 2017). The implementation of the master plan had been incorporated in the coalition agreement of the federal government of 2018 and was confirmed by the actual government.

Deutsche Bahn AG (DB) has previously published a digitalization strategy DB 4.0 in 2015 (DB 2015). In the current technology strategy, the core objectives of "increasing passenger comfort," "improving system performance," and "securing profitability" are mentioned. On this subject, one can refer to modern technologies arising from digitalization, such as "Internet of Things," "Robots and Drones,"

[2]Engl. "Implementation recommendations for the future project Industry 4.0."

Grade of Automation	Driving Operation	Starting	Braking	Door Closure	Operation in Disruptions
GoA 0	On-sight Driving	Driver	Driver	Driver	Driver
GoA 1	Manual Drive with Train Control (ATP)	Driver	Driver	Driver	Driver
GoA 2	Partly Automatic Train OperationwithDriver (STO)	Automatic	Automatic	Driver	Driver
GoA 3	Attended, Driverless Train Operation (DTO)	Automatic	Automatic	Train Attendant	Train Attendant
GoA 4	Fully Automatic, Driverless Train Operation (UTO)	Automatic	Automatic	Automatic	Automatic

Fig. 1 UITP definition of the grades of automation (GoA) [© IFS, RWTH Aachen according to definition UITP 2011]

"5G," "Block Chain," and "Additive Manufacturing" (Härdi 2018). Other European railway companies, such as the French SNCF with its #DigitalSNCF program and the Swiss Federal Railways (SBB) with Smart Rail 4.0, also have similar strategies that are based on Industry 4.0 (Masse 2018; Rytz and Mandour 2018).

The following pages analyze what seems to be reasonably feasible from the original idea of Industry 4.0 for the rail transport system, what, if any, has already been implemented or is currently being tested, and what could still be realized in the future. The actions of rail transport are reflected in the features of Industry 4.0

- networked
- informative
- self-organizing
- autonomous or automatic

1 Autonomous Driving

It is the original task of rail transport to transfer people or goods. Rail vehicles are track-guided vehicles, i.e., the automatic keeping and changing of direction—generated by steering in automotive vehicles—is inherent to the system. It is therefore sensible to automate the basic task of rail transport, driving, which is one of the abovementioned properties of Industry 4.0.

In 2011, the International Association of Public Transport (Union Internationale des Transports Publics UITP) published a press release on the state of the art of automatic metro systems. The state of the (partial) automation of metros by five **Grades of Automation (GoA)** is therein described (Fig. 1). GoA 0 denotes on-sight driving. In GoA 1, driving is protected by signals and train control systems (**Automated Train Protection** ATP). However, the driver must start and stop the train as well as control the doors and react to malfunctions. In GoA 2, an automatic system assists the driver in starting and braking (Semi-automated Train Operation STO). In GoA 3 the operation is quasi fully automatic, only that instead of the driver, a train attendant takes over the door control and incident management (Driverless Train Operation DTO). In GoA 4, the vehicle drives completely automatically and unattended (**Unattended Train Operation** UTO). Since then, the UITP definition has been adopted by many railway operators (UITP 2011).

1.1 State of the Art

Rail transport is divided into different types. For instance, a distinction must be made between passenger and freight transport. **Passenger transport** is further divided into long-distance and local transport. **Local transport** can be further subdivided into urban and regional transport. In Germany, urban transport is largely operated according to the construction and operating regulations for trams[3] (BOStrab 1987), which also apply to light-rail and metro systems, while regional and commuter train transport, similar to freight and long-distance passenger transport, are regulated according to the Railway Construction and Operating Regulations[4] (EBO 1967).

A train run cannot be carried out uncoordinated and spontaneously, but must be planned so that the route to be traveled is free at the given period, i.e., is not occupied by another vehicle and the switches are adjusted so that the train is directed to its destination (route). The only exception is the tram, which is operated broadly on sight, similar to automotive vehicles and either the driver (according to BOStrab requirements for drivers) adjusts the switches manually or via wireless communication or the vehicle. However, the tram driver must also adhere to a timetable.

The train driver must follow certain guideline parameters on the railway. According to Pachl, they are the approval of a train and the permissible driving speed depending on the train position, as well as stopping at the scheduled stations (train stations and stops) (Pachl 2017). In addition, at the station platform the driver must pay attention to whether there are still passengers in the boarding area and must react to short- or long-term disruptions and monitor the route (including the opposite track) with regard to hazards during the ride. He receives the necessary information from the rail traffic controller at the railway traffic control center, the available timetable, as well as fixed or displayed signals and speed charts in the driver's cabin.

There are several reasons for this type of operation. First, a track-guided vehicle cannot pass another vehicle at an arbitrary location. For this purpose, a passing loop must be available, the usage of which must also be planned. Second, the train

[3] Bau- und Betriebsordnung für Straßenbahnen.

[4] Eisenbahnbau- und -betriebsordnung.

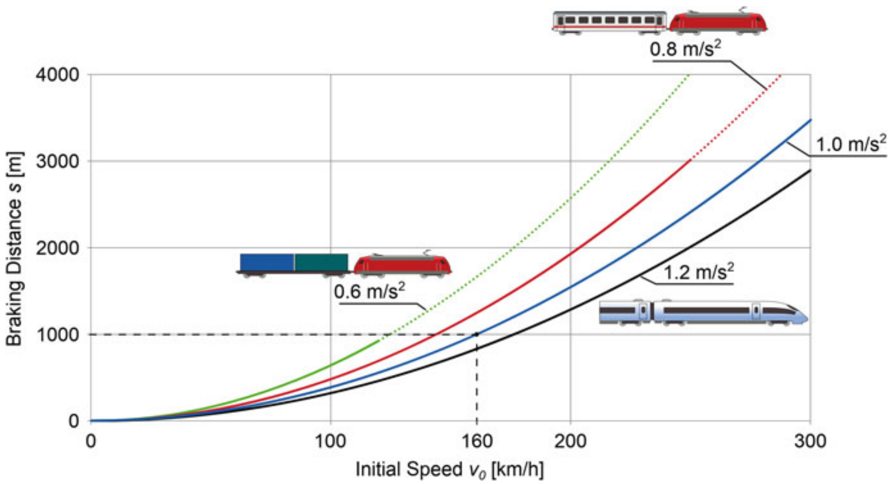


Fig. 2 Braking distances of mainline railway vehicles [© IFS, RWTH Aachen by Haigermoser 2002]

schedule must be followed. Train traffic jams and other unforeseen waiting times must be avoided. Third, the safety of driving operations must be guaranteed. That is, trains must not experience head-on or flank collisions (front and flank protection, the latter can occur in switches and crossings), and there must not occur rear-end collisions (protection in case of one train following a train running ahead).

Since rail vehicles have very long **braking distances** due to the low adhesion coefficient between wheel and rail compared to the road vehicles (tire/road surface), in most cases they cannot be brought to a standstill within the range of vision of the train driver while braking (Fig. 2). Their maximum braking deceleration is between 0.6 m/s^2 and 1.2 m/s^2, which leads to braking distances of up to 1000 m for freight trains from an initial speed of 100 km/h and for passenger trains from 160 km/h to standstill. A **high-speed train** at 300 km/h speed needs up to 3000 m braking distance to stop safely, despite of a higher braking capacity than other train types.

The train driver must know whether and when to brake on time. This is made sure by signaling and protection systems, which, if disregarded, affect the safety of the train travel to the safe side. In the case of **intermittent automatic train control** (in Germany PZB[5]), i.e., stationary signal transmission, the position of the main signal at maximum braking distance is shown to the train driver by a distant signal. If the train overruns a distant signal that indicates a stop, the train is emergency braked so that it comes to a standstill before the main signal.

With modern PZB systems, the speed curve between the distant and the main signal and the highest speed, maximum 160 km/h, are monitored by lineside sensors, so-called track magnets. Newer train control systems use transponders that can send

[5] PZB—Punktförmige Zugbeeinflussung.

data telegrams in case of an overrun, which contain more information than merely a stopping or driving speed specification (Pachl 2004).

Since passenger trains that are faster than 160 km/h require a braking distance longer than 1000 m to come to a standstill, they cannot be operated with the existing stationary signals and PZB systems. **Continuous automatic train control** (LZB[6]) was introduced at the German Federal Railway (Deutsche Bundesbahn[7]) in 1975, which made it possible initially for intercity long-distance trains to drive at speeds of up to 200 km/h and later from 1991 for the ICE trains up to 300 km/h. It involves continuous monitoring of the train speed via conductor cables installed in the track and a display in the vehicle indicating permanently the momentary maximum speed and that after the next velocity change, as well as the distance to that point. Since the adopted line cables provide the location as well and LZB-guided vehicles are mostly equipped with an **automatic driving and braking control** (AFB[8]), they can operate without driver action. The train driver only must be always ready to intervene in station areas, at the transition from an LZB to a PZB line, and in the event of disruptions. This level of automation could therefore be denoted as GoA 1. Incidentally, LZB systems are also utilized with subway and light rail vehicle systems, with the latter, of course, only on lines with an independent track, that is provided in tunnels and elevated sections as well as with tracks that are separated by fences (Fig. 3).

In Europe alone, there are over 20 different train control systems (Poré 2007). It was therefore logical in the course of the attempts of the European Union for a consistent and safe rail transport in Europe (**interoperability**) to replace the slowly aging national systems in the 1980s with a common **European Train Control System** (ETCS). It lasted until 2002, when Switzerland put the world's first commercial ETCS route (Level 2, see below) between Zofingen and Sempach into operation (N.N. 2002a). In Germany, the first commercial route with ETCS (Level 2) was opened at the end of 2017 with the completion of the high-speed connection Munich–Berlin between Ebersfeld/Bavaria and Halle (Saale) and Leipzig respectively (Drescher and Feldwisch 2017).

Apart from the fact that many national variations of ETCS have emerged, there are three fundamental, downward-compatible levels of the system. In terms of functionality, level 1 corresponds to an intermittent automatic train control with stationary signals. It is possible to replace them with signal displays on the control panel in the driver's cabin (Fig. 4).

ETCS level 2 corresponds to a continuous train control. Technically, the permanent connection to the traffic control center is enabled by a permanent and secure wireless communication through **GSM-R**. The train is localized using stationary tags (Eurobalises) supplemented by on board odometers. Train operation still happens with fixed track intervals, i.e., blocks (Fig. 5).

[6] LZB—Linienförmige Zugbeeinflussung.

[7] Deutsche Bundesbahn s the West German predecessor of Deutsche Bahn AG.

[8] AFB—Automatische Fahr-/Brremssteuerung.

Fig. 3 Start of a light rail track equipped with LZB (Photo C. Schindler)

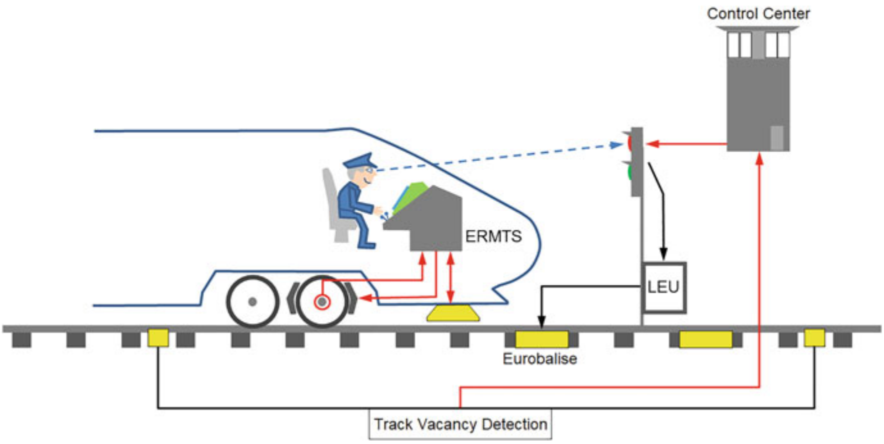


Fig. 4 Operating principle of ETCS level 1 (Sketch: A. Daniel, © IFS, RWTH)

With PZB and LZB systems as well as ETCS levels 1 and 2, the track vacancy detection, i.e., the information that a section of the route (block) is no longer occupied by the train ahead and can thus be available for a next train, is reported

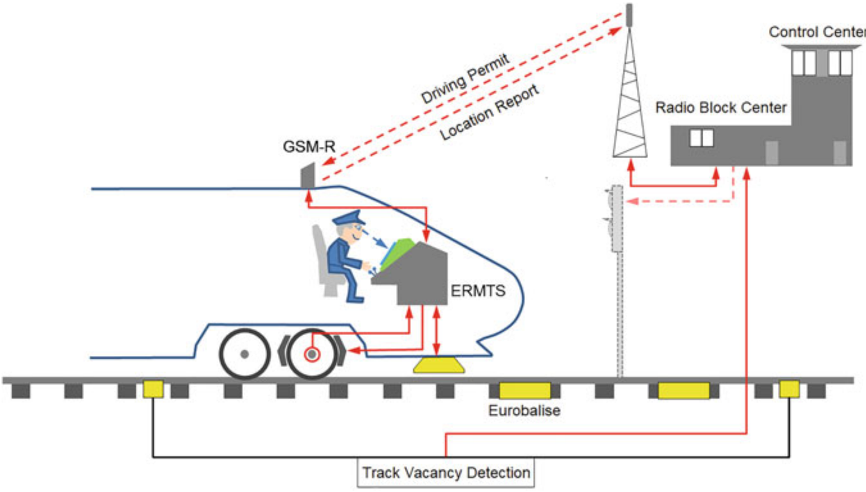


Fig. 5 Operating principle of ETCS level 2 (Sketch: A. Daniel, © IFS, RWTH Aachen)

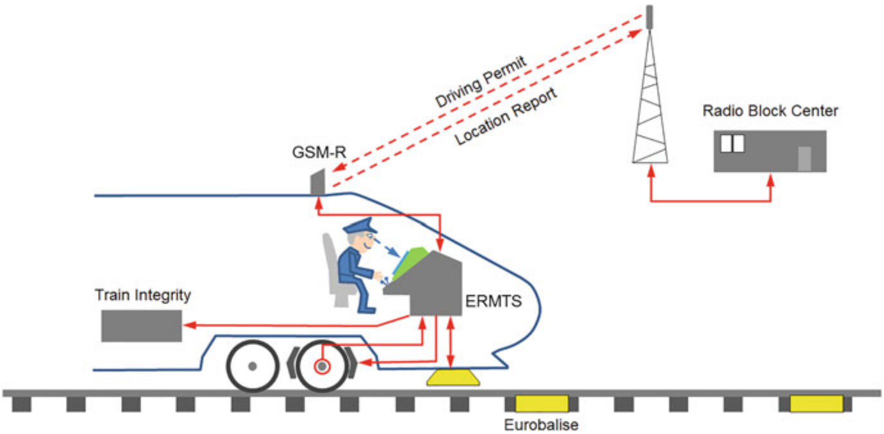


Fig. 6 Operating principle of ETCS level 3 (Sketch: A. Daniel, © IFS, RWTH Aachen)

via axle counter equipment to the Train Control Center. This also ensures that a single wagon or wagon group has not separated (**train integrity monitoring**).

Only with level 3 of ETCS, the stationary signals as well as fixed track sections are abandoned. As with level 2, the localization is via Eurobalises and odometers. The train integrity monitoring is now performed on the vehicle (Fig. 6). Since the positions of all trains in the area of accountability of one control center are always known and the control center controls their speed, it is now possible to drive at an absolute braking distance. That is, a succeeding train can follow the preceding train at the distance it needs to come to a standstill with a full braking (Pachl 2008).

Currently, with PZB and LZB, the data transfer between the vehicle and the control center takes place via hardware wiring on the track. The so-called train radio is employed merely for communication between the rail traffic controller and the train driver. The wireless communication GSM-R, which is used with ETCS, is based on the second generation of wireless mobile telecommunications technology 2G (the first digital generation of mobile communications). While ETCS is just recently being slowly introduced in most countries—in some smaller European countries, above all Switzerland and Luxembourg, there are at least concrete plans for implementation and completion for a network-wide expansion—the wireless mobile telecommunications technology is already much more advanced. Nowadays, the 4th generation (4G also called LTE) is the standard here and the licenses for the next generation **5G** were auctioned in Germany in 2019 (Dornis and Wulf 2019). SBB is therefore considering the implementation of further ETCS expansion with a more advanced train wireless communications technology (N.N. 2018). At DB, it is initially planned to launch ETCS with GSM-R, but the so-called Future Railway Mobile Communication System (FRMCS) is in progress as well, including the options offered by the 4G and 5G mobile wireless communication standards.

With mainline railways driving without the external train control is only allowed up to a maximum speed of 50 km/h, because in this case the train driver can still brake safely before reaching a recognized obstacle. By comparing the braking distances of different types of rail vehicles with that of an automotive vehicle, it is observed that only the **tram** comes close to that braking power. This is because trams and light rail vehicles are equipped with an extra brake that is independent of the wheel/rail adhesion, normally an electromagnetic track brake. Together with the wheel brakes (regenerative or dynamic brakes and mechanical brakes), they must enable average decelerations of at least 2.73 m/s^2 from an initial driving speed of 70 km/h to standstill and thus be capable to brake within 69 m (BOStrab 1987). **Subway** vehicles already need twice as much of that braking distance and regional railcars about three times (Fig. 7).

However, subway and metro systems have the advantage of having no interfaces to other type of transport opposite to trams and light rail vehicles, which operate extensively in road traffic, also as opposed to the railways with their freely accessible tracks as well as the open platforms and level crossings. Except for the platform, which is the intended and necessary interface to the passenger, subway vehicles operate on an independent railway track that is in the tunnel or elevated. The few non-elevated routes above the ground are protected from unauthorized trespassers by fences. There are no single-track routes, so that only protections for flank and succeeding operations must be guaranteed. It is similar to the so-called **People Mover** systems, normally smaller short-distance transport systems, which often use track guidance technologies different from the classic wheel/rail technology. In addition, they often operate only on one line on a double-track route. Examples in Germany are suspended monorail systems such as the Wuppertal suspension railway and the Sky Train in Düsseldorf airport as well as upright standing people movers such as the airport trains in Munich and Frankfurt. Except for the first mentioned system, they are vehicles with pneumatic or solid rubber tires that are guided over

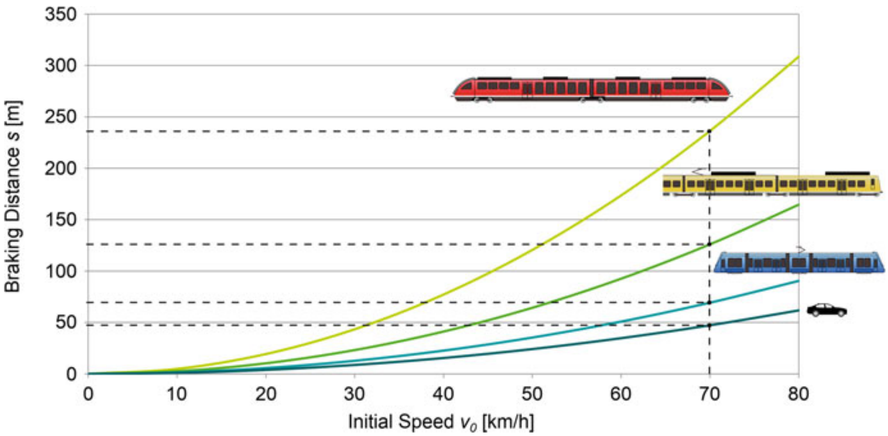


Fig. 7 Braking distances of short-distance railway vehicles (© IFS, RWTH Aachen)

horizontally arranged lateral rollers (Trummer and Rappe 2008; Hondius 2015; Kuschinski 2003; N.N. 2016a).

Reasonably, the first implementations of fully automatic and unattended driving (GoA 4) took place on subways and people movers. The world's first GoA 4 system is the Port Island Line in Kobe, Japan, which opened in 1981. Europe followed in 1984 with the VAL system in Lille, France. Both systems use conventional pneumatic tires for load support and lateral rollers for guidance. The first GoA 4 system with classic steel wheel/steel rail technology is the Sky Train in Vancouver, Canada, which came into operation in 1985. A study by the UITP (Malla 2013) shows how strong the trend towards fully automated, unattended metro and people mover systems is. Subsequently, such systems existed in seven cities worldwide until 1990, 48 lines in 32 cities in 2013, and nowadays in 2018, according to Wikipedia, 70 systems in 59 cities (Wikipedia 2018a). The same source lists additional nine fully automatic metro systems with train attendants, i.e., GoA 3 systems.

It is therefore clear that fully automated, driverless operation in metro and people mover systems has been state of the art for decades. Half of all systems are controlled inductively via LZB such as line cables or via radio communication called Communication Based Train Control CBTC similar to ETCS level 3. The only critical interface, the platform, in most systems is protected by platform screen doors that open only when the train is at a standstill and close again before departure. To align the vehicle doors with the platform doors, a precise target braking of the train is necessary (Malla 2013).

The advantage of fully automated subway transport is the possibility of accurate following of a given driving profile and thus adhering to the timetable correctly. This is a requirement for extremely short headways. By application of platform screen doors, the risk of accidents caused by entering and exiting passengers is also minimized.

Currently, the world's only commercial **driverless mainline railway** system is operated by the mining group Rio Tinto in north-west Australia. Since 2017, the approximately 2400-m-long ore trains have been operating at speeds of 70 km/h to 80 km/h on the 100 km to 500-km-long routes between the inland mines and the loading terminals on the coast. The trains are controlled and monitored via GSM-R from the control center in Perth, 1500 km away. Level crossings are monitored remotely via video and obstacle detection. There, the existing trackside train protection remained (Randelhoff 2018).

Based on the above remarks, it is realized that the rail transport system is currently quite far developed concerning the attributes "automatic" and "networked." Trams in road traffic and mainline vehicles in shunting operate extensively on sight and therefore are controlled only by the driver or shunter, respectively. Railway vehicles in accordance with rules such as the German EBO are at least controlled and monitored from the outside depending on location, but it does not yet correspond to the definition of GoA 1, which exists only from continuous automatic train control (LZB, ETCS levels 2 and 3). With the subway, the step from semi-automatic driving in GoA 1 via GoA 2 to the fully automatic operation in GoA 3 or GoA 4 is relatively simple. In the case of train disruption with GoA 4, the repair services must come from the outside to the fault location, which is more easily done in an urban transport network than in long-distance transport.

The rail vehicles are networked by their control center via continuous automatic train control (LZB) or GSM-R (ETCS). Additionally, there exists an audio connection via radio communications from the train driver to the control center and to other train drivers. However, the other Industry 4.0 attributes "self-organizing" and "self-optimizing" cannot be found in this context yet.

1.2 Automatic vs. Autonomous Driving

The previous subsection was exclusively about automatic driving. This operating concept requires a control center, in which all information about the locations and current driving speeds of the trains are available within the range of the area of responsibility. Based on this information, the timetables, and the current traffic situation (delays, technical malfunctions, etc.), it is decided centrally which train should be granted the driving permit at a time via the signaling technology and if applicable, by radio communication. In case of loss of the signals or the communication with the control center, the train must not continue to drive. The automatic driving is therefore completely dependent on information and signals from the outside.

There have been attempts for years in the automotive industry as well for driverless transportation. However, since the tradition and infrastructure of control and safety technologies are absent, a completely different approach is chosen. The vehicle must be capable of moving independently without information from the outside and merely based on its on-board sensors and **artificial intelligence**. Only the start time and destination are specified from the outside, possibly by the passenger. This type of driverless transportation is called autonomous.[9]

Autonomous driving is therefore a substitute for on-sight driving. The on-board sensor system replaces the driver's senses—mainly the visual perception. If a sensor were capable of reliably detect an obstacle at a distance of up to 3000 m from a moving train, a **high-speed train** could theoretically also operate autonomously. Since there are no long-range sensors that work reliably in all weather and lighting conditions and there is rarely a clear visibility of the aforementioned long distance, this form of autonomous driving fails for most types of rail transportation. In addition, a braking distance of 1000 m for freight trains with maximum speed of up to 100 km/h and passenger trains up to 160 km/h cannot be monitored reliably using on-board sensors.

Nevertheless, the technology of autonomous driving arises in areas where driving is still currently on sight. That would be the tram in urban road traffic, the mainline trains in the stations, and any freight transport in closed areas, where it is shunted with low driving speeds. However, the route must be set in the case of rail vehicles, which can be done either by a control center or individually from the vehicle at the relevant switch, as is usual with trams and in shunting operations (Girnau et al. 2007; Peiser 2015).

1.3 Research and Development

In Germany, Frederich had carried out a research about possibilities of autonomization of rail transport in the late 1980s and early 1990s. To this end, he casted doubt on the ordinary rail transport with long trains that operate according to a fixed schedule and proposed individual motorized driverless units.

His SST[10] version (**automatic signal-guided traction unit**) still operates relatively conventionally, but in mixed traffic with ordinary trains. An SST recognizes its route, the data of which are stored in the on-board computer, and its location. Additionally, it is aware of the speed limits at each point as well as of the location of signals. Principally, it stops in front of every signal, "speaks" to it via radio communication, and waits for the permission to continue driving. The SST system was successfully tested at RWTH Aachen University and from 1996 onwards, approved in a pilot operation in the inter-company transport of VW between Salzgitter and Wolfsburg (Frederich 1992; Frederich and Lege 1996; Molle 1998).

---

[9] In the automotive industry, the terms "automatic," "automated," and "autonomous" are often used interchangeably. The Society of Automotive Engineers SAE defines the levels of vehicle autonomy with partially, conditionally, highly, and fully automatic, although it means autonomous indeed (Gasser et al. 2012; ADAC 2018).

[10] SST—Selbsttätiges Signalgeführtes Triebfahrzeug.

Fig. 8 "Aachener Rail Shuttle (ARS)" study by the Institute for Rail Vehicles and Transport Systems at RWTH Aachen University (Picture by: B. Schiefer, © IFS, RWTH Aachen)

The further development of the SST led to **self-organizing freight transport** (SOG[11]), which does not require signaling technology or a timetable, but is based on single transportation of small, autonomous, and motorized units. The vehicles set their own routes via radio-controlled switches after entering the destination by an operator. The switches exchange information about the vehicles' locations and driving speeds with the nearby vehicles with the same equipment via radio communication. Additionally, the units organize and optimize their own route in "coordination" with the other autonomous units (Frederich 1994; Frederich and Lege 1996). The SOG concept thus has all the properties of Industry 4.0 long before this idea arose. Concept studies on the possibilities of driverless rail transport are currently carried out again at the author's institute at RWTH Aachen University. In contrast to Frederich, the focus is less on freight transport, but on passenger transport in rural areas. Small driverless units with electric motors are anticipated that will undertake the shuttle service to the larger train stations on secondary lines. This **rail bus transportation** can operate either as an SST in the conventional mixed traffic or as an SOG in single operation (Fig. 8).

With the **Aachener Rail Shuttle** (ARS), the passengers should be transported sufficiently quickly, which requires maximum driving speeds of approx. 100 km/h. On-board sensors of different technologies with the corresponding range and resolution should be able to stop securely with the mentioned speed by means of signal fusion. Networking with the infrastructure is considered as well, e.g., with a fixed level crossing control, which should make it unnecessary to drive slowly at those points. On main lines, there is the possibility that several ARS independently come together to form a train and operate at a short sensor distance (electronic coupling).

In the considerations of Frederich and Vallée about the SOG mentioned above, the topic of obstacle detection did not appear. They assumed that every vehicle is aware of the location of other vehicles at any moment. Vallée has developed an extensive conflict strategy on that (Vallée 2002).

[11] SOG—Selbstorganisierender Güterverkehr.

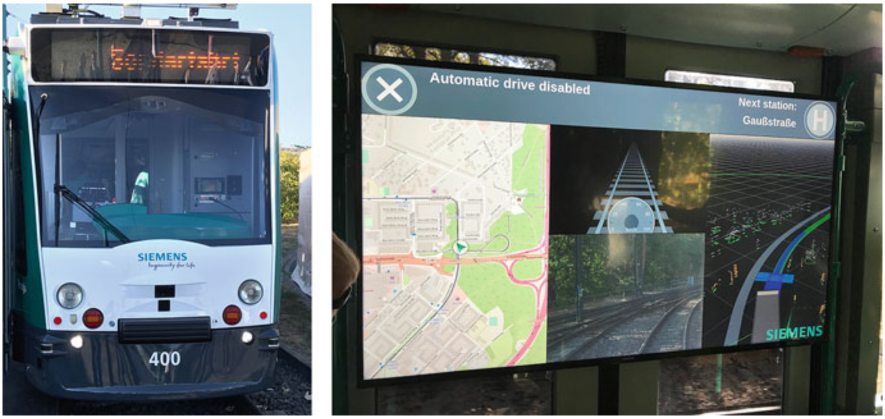


Fig. 9 Autonomous tram in demonstration mode in Potsdam (left), monitoring screen (right), (Photos: C. Schindler)

A project with the same strategic focus is currently being tested by a spin-off from DLR at the Harz narrow gauge railways. In this project, a system originating from the aviation sector for localization of all aircrafts in a surrounding area is adapted to the train system. Key technologies here are the combination and fusion of several localization technologies including all current satellite navigation systems and the state-of-the-art radio technology for communication. The investment costs for such a system with allegedly higher security than conventional control and safety technology would amount to approximately 20% of the costs for PZB systems (Haas 2018).

German Railways (DB) is also currently working on further automation of its traction units. In this regard, their **advanced driver assistance systems** ADAS, which gives the train driver real-time information on **energy-optimized driving** and current information on the route and timetable as well as arrival and departure time predictions, schedule deviations, and vehicle locations, has been adjusted to enable automatic driving in combination with the existing train control systems. The driver's former tasks, such as route knowledge and route monitoring are replaced by route data saved in the on-board computer and sensors. Moreover, starting, driving, and braking functions are performed by ADAS 4.0 through the automatic control of the driving/braking lever. The system is not intended to replace but merely to relieve the driver and further increase the safety and availability of vehicle operation (Claus and Ammoser 2018).

Like the automotive industry, the tram sector deals with enabling its vehicles to drive autonomously in the road traffic. Driver assistance systems for **obstacle detection** and **collision avoidance** are already on the market (Rüffer 2017; Rüffer 2018). A current industrial research project addresses the autonomous tram. Siemens has equipped the pilot series vehicle of the COMBINO type with sensors and artificial intelligence and presented it to the public as an autonomously operating vehicle at Innotrans in September 2018 (Fig. 9, left) (Bihn 2018). So far, the vehicle has only been operated on separate railway track sections, where the traffic situation is less complex than on the street. Interfaces with other road users, such as automotive vehicles, cyclists, and pedestrians are limited to the crosswalks and level crossings. However, it must always be considered that a pedestrian crosses the track at any point. During the test drive, the vehicle position on the route, the current driving speed, the field of view from the driver's cabin, the current sensor range, and the instantaneous braking distance to a standstill could be followed on a monitor (Fig. 9, right).

2 Driverless or Unattended Shunting

In addition to the line service, called main run, shunting processes are of great importance in rail freight transport. They are labor-intensive, partly dangerous for the workforce, and includes much idle time. These are the main reasons for the very low average transportation speeds in this segment. Shunting denotes (German Wikipedia 2018b):

- coupling and uncoupling (forming) trains
- transferring a group of wagons or individual cars to another train station
- bypassing a train turning at the railhead with the locomotive (terminal headshunt)
- moving individual traction units to and away from the trains inside the train station track area
- the supply and collection of railway wagons at loading facilities such as loading docks and loading ramps
- the supply and collection of wagons in sidings, e.g., inside an industrial plant
- the transfer of traction units, wagons, and wagon groups to and from workshops and holding areas

A common attribute of all these operations is a very low driving speed. A number of them are performed in closed areas, e.g., at freight stations or factory premises. There are hardly any interfaces to other modes of transport (if there are, they can be easily controlled by own-account transport regulations) nor is there the possibility of unauthorized access to the track. For this reason, while shunting, it is driven almost entirely on sight, in which a so-called shunting attendant is allowed to operate the locomotive via a wireless remote control instead of the driver in the driver's cabin. The attendant monitors the shunting process from the outside.

2.1 State of the Art

Currently, train formation and decoupling is almost exclusively a process of rail freight transport in Europe. Passenger trains operate largely as fixed train units, regardless of whether they are multiple units that cannot be (de-)coupled in the

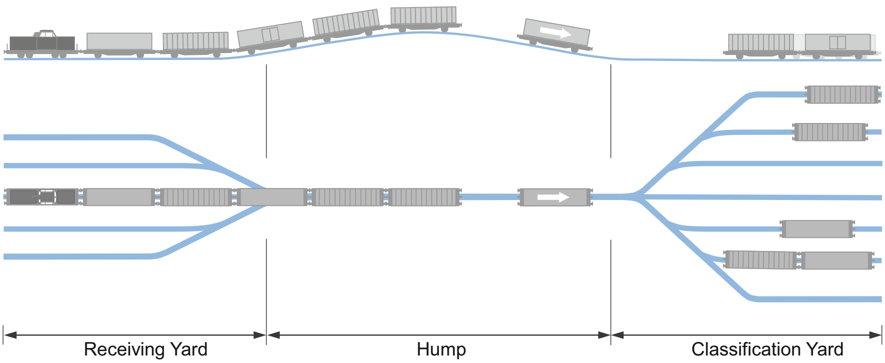


Fig. 10 Schematic representation of a hump (Picture by: A. Daniel, © IFS, RWTH Aachen)

operation or locomotive-hauled trains with individual wagons. With freight trains, a distinction is made between block trains and single wagon transport.

Merely with the latter, trains are always configured anew and relatively freely. At smaller marshalling yards, it is done by pushing or pulling the wagons or wagon groups out of the track, on which the arriving trains and wagon groups have been parked, and pushing (pulling) them into the track, in which the train is being reassembled. The arriving wagon groups, that are to be divided into new trains, must be also uncoupled beforehand, which is still done manually at least in Europe, where the screw coupling predominates. The uncoupled single wagons (or wagon groups) are pushed via the side buffers, while automatic shunting couplers are used mostly for the pulling, which, attached to the locomotive, grip the coupling hook of the wagon screw-coupling controlled by the driver and then release it again (Janicki and Reinhard 2008).

Large marshalling yards have a **hump** onto which the single wagons are pushed up as a group (Fig. 10). Then, they move independently downwards due to the gravity towards the classification yard, in which the tracks are sorted using automatically changing switches, so that each track develops the configuration of a new train. It is important to push the wagon groups that are not coupled onto the hump at a speed that is low enough to enable the switches to change after each wagon. Currently, it is controlled by computers. Additionally, the braking of the wagons in the classification yard is performed today automatically with trackside retarders (Berndt 2001).

The wagons of the newly assembled trains must then be coupled manually. Before departure, the wagon inspector walks along the train to detect possible damages or defects. In addition, the performance of the brakes at standstill must be checked together with the traction unit driver (**brake test**).

Thus, despite the automation of some steps of the process, the train formation with help of a hump is not realized without human intervention. The most advanced operation is currently at the Russian marshalling yard Luzhskaya, where it is automatically performed and controlled from the control center (Smagin and

Popov 2018). This operation is facilitated by the **automatic central buffer coupling,** which is predominant in freight transport in Russia as well as in nearly all major railway countries outside Europe, which enables automated coupling and uncoupling. A temporary solution is a facility for the automatic disconnection of freight wagons with a screw coupling. It was developed at the University of Applied Sciences Upper Austria in Wels and is currently being tested at the marshalling yard in Linz (Egger et al. 2018).

All other shunting processes mentioned at the beginning, including own-account transports, require currently locomotive drivers or wireless-controlled locomotives—Luzhskaya is the exception.

At least the coupling process may improve in the future. In September 2020, the transport ministers of the European Union decided to introduce the so-called Digital Automatic (Center-) Coupling (DAK[12]) until 2030, the year when they want to achieve a market share in freight transport with rail of 30%. The innovation in comparison to existing automated center couplings is the cable coupling, which provides electricity and data connection between the wagons.

This allows not only automatic coupling, but also train integrity supervision without wayside installations, automated brake test per remote, condition monitoring through on-board sensors, data transmitters, and finally, freight train lengths of up to 1000 m due to a better brake capacity by immediate pneumatic braking initiated by an instant electric brake signal and a higher mechanical strength of the coupler compared with the existing screw-coupler (N.N. 2020; Feindert 2020).

Generally, about 450,000 existing freight wagons in Europe are to be equipped with the DAK. How this shall be done will be decided by the EU until 2022. Switzerland and Germany are among the first movers and do research and trials on the DAK (N.N. 2019a; Stötzel and Vallentin 2020).

2.2 Research and Development

Most shunting tasks offer ideal conditions for a driverless operation. For this reason, many application-oriented research projects focus on automatic or autonomous shunting.

As early as 2002, a consortium of Siemens AG, RWTH Aachen, and TU Braunschweig presented the **CargoMover**, a diesel-powered freight vehicle that operated fully autonomously in shunting operations. Radar and laser sensors as well as a black and white video camera made it possible to monitor the space in front of the vehicle and to detect the distance and speed of existing objects. Up to a driving speed of 30 km/h, the vehicle was able to come to a standstill at a specified distance from an obstacle. However, the CargoMover should have been able to operate on the rail network independently and perform door-to-door transportation (rail connection

[12] DAK—Digitale Automatische Kupplung.

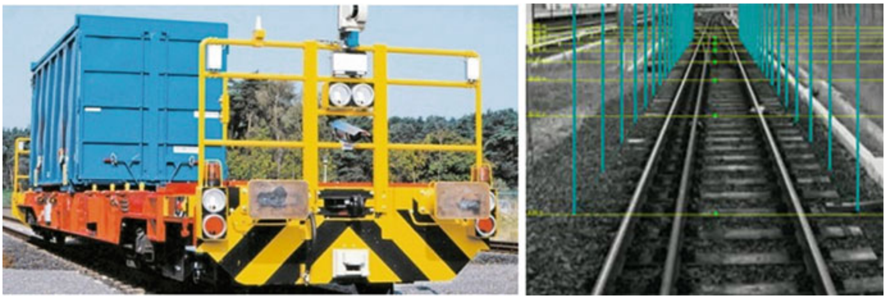


Fig. 11 Autonomous CargoMover in Siemens test center Wildenrath (© IFS, RWTH Aachen)

to rail connection) up to a distance of 150 km at a maximum speed of 80 km/h. Therefore, at the Siemens test center in Wegberg-Wildenrath, the track and vehicle were equipped with ETCS Level 2 in the early 2000s (Fig. 11) (N.N. 2002b).

In another project named Flex Cargo Rail, the idea of a motorized freight wagon was further developed. The concept envisaged that the motorized freight wagons would provide the "**last mile**" service from the marshalling yard to the final destination (or vice versa) independently, i.e., autonomously and with their own power. At the marshalling yard, they would independently couple themselves to full trains, which would be hauled by locomotives in the main run (Daniel 2007).

Today, this idea can be pursued that the single wagons—certainly with electric motors—run uncoupled at a very short distance behind each other, but communicatively networked as a train with "electronic couplings." The term "**platooning**" has been established in the automotive sector. Its advantage would be that expensive couplings could be dispensed and the wagons moving closely behind one another would still be valid as a train with respect to operation, i.e., train control technology. See also sect. 1.3, Aachener Rail Shuttle.

In 2018, a shunting locomotive of class 296 equipped like the CargoMover was presented to the public. A consortium of Deutsche Bahn, a medium-sized company, and the Technical University of Applied Sciences, Nuremberg, equipped the locomotive with sensors for obstacle detection (video, thermal image, laser scanner), GPS signal reception for localization, and an automatic control system in the project "Fully automatic push-locomotive VAL 2020." The goal was to enable it to move autonomously toward a group of freight wagons and push them onto the hump of a marshalling yard. It is able to detect possible obstacles between itself and the wagon group and to stop in front of them independently (Bergmeister 2018).

Another concept for driverless shunting is pursued by a consortium of DB Systel, the road-rail vehicle manufacturer Zwiehoff, and RWTH Aachen University. Instead of locomotives, shunting is performed with small and unmanned **road-rail vehicle**s. The operation of these small tractors is state of the art and they are capable of shunting wagon groups with up to 250 t total weight by means of wireless remote control. The battery electric vehicle has a velocity of up to 10 km/h and is deployed especially in rail yards and factory premises (Fig. 12).

Fig. 12 Driverless road-rail shunting vehicle (Photo: H.-S. Jung, © IFS, RWTH Aachen)

Since the vehicle not only must be able to push the wagons or wagon groups, but also to pull them, it requires an automatic shunting coupling, which besides the state of the art must detect and notify whether the (de)coupling process was successful. Furthermore, the vehicle is equipped with sensors for obstacle detection. It must also be able to distinguish between a disturbing obstacle and a wagon, to which it should slowly approach for coupling or pushing. In contrast to pushing onto the hump, pushing a wagon or a wagon group along a shunting track poses a major safety challenge, since the track in front of the first wagon must be monitored for obstacles. The localization should be carried out with inexpensive industrial RFID tags attached to the track. The vehicles receive information about the location of the wagons to be shunted from a control center. The followings are calculated knowing the position of all wagons and shunting units:

- which shunting unit is free and can execute the shunting tasks best (fastest, most energy-efficient, etc.)
- which track should the shunting vehicle use to drive to the wagons and move with them to its destination, and
- which switches must be adjusted for that and how.

The disposition unit transmits the driving command to the vehicle and adjusts the track(s) accordingly (Knapmöller and Pritsching 2018).

Alternatively, a solution like Frederich's SOG would also be possible, in which the shunting units store the (limited) track network in the on-board computer and search the way to their destination themselves. In this case, only the order "which wagon (group) and from where to where" is received from the control center.

The localization could also be carried out by satellite-borne positioning via GPS or via the European **satellite navigation system** Galileo as in the project “Galileo Online: Go!” funded by the Federal Ministry of Economic Affairs and Energy. This project could demonstrate that even localization with track-selective accuracy is possible by satellite (Zweigel et al. 2018).

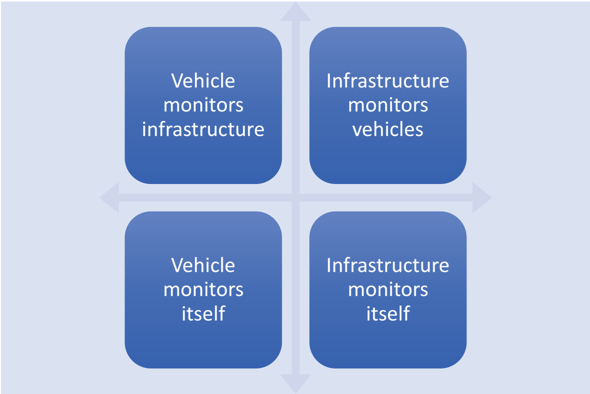
Another challenge for an autonomous or automatic shunting vehicle is to correctly determine the braking distance with a wagon group being trailed or pushed, which strongly depends on their weight. This problem is addressed by Franzen et al. in a paper about the development of a shunting assistance system RANGierASSistent by the locomotive manufacturer Reuschling and the Ruhr University Bochum. Even if autonomous driving is not the goal, obstacle detection and automatic driving and braking should still be installed on a shunting locomotive as in the VAL 2020 project (Franzen et al. 2017).

3 Monitoring via Sensors and Networking

Increasing the reliability of vehicles and infrastructure is a primary goal of all railway companies to accomplish the tasks of track and shunting services on time. It is important to have an exact knowledge of the technical components condition continuously. DB follows a so-called four-quadrant strategy as part of its current activities for **condition monitoring** (Schulte-Werning et al. 2017) (Fig. 13).

Sensors are required to monitor the condition of technical equipment. In the simplest case, they transmit the measured data to the driver or preferably to a data acquisition center (via telematics). There, the data are analyzed. If there is a deviation from the target behavior or target condition of the monitored system, actions are initiated. This can be maintenance and repair measures on the vehicle in the depot or on the infrastructure on the line up to warning the driver to reduce speed for example (Fig. 14).

Fig. 13 Condition monitoring strategy of Deutsche Bahn (Figure IFS, RWTH Aachen)



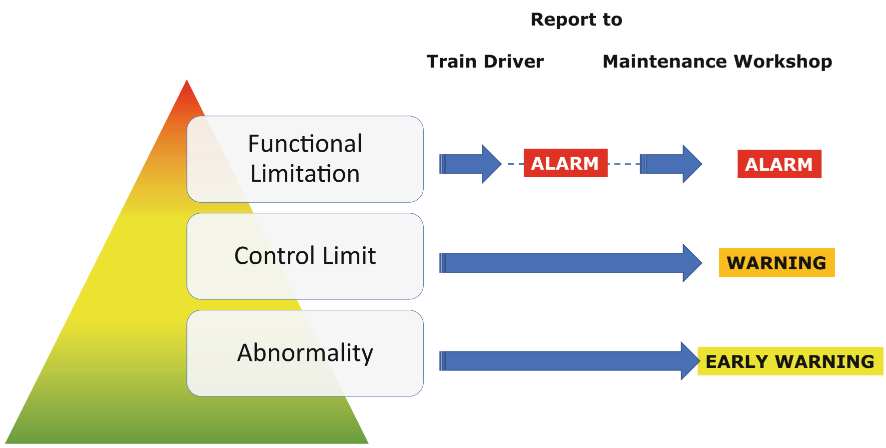


Fig. 14 Hierarchical monitoring and reporting concept [Figure IFS, RWTH Aachen, acc. to Müller and Sunder 2011]

In addition, it is important to have the exact current location of the vehicles. Particularly, the SBB expects a more precise control of the train movements with the potential of increasing the track capacity as well as conservation of infrastructure facilities, e.g., track vacancy detection system. The precise localization is carried out via a continuous on-board positioning system of the train in contrast to the infrastructure-based localization that is common today, cf. (Rytz and Mandour 2018).

In contrast to passenger trains, freight trains are freely configured as required, i.e., according to the tasks of goods transportation and the type of goods. The trains perform a so-called wagon management, which means they have the possible locations of their wagons at least in their own area. However, the exact positions of the rail cars, e.g., on the customer's factory premises or in a large marshalling yard is not known. This is aggravated by the fact that the freight wagons have no energy aboard, which means that precise localization is not possible, one of the draw-backs that will be cured with the introduction of the Digital Automated Coupling (DAK), refer to Sect. 2.2.

Another goal of on-board sensors and networking with the operations center is train integrity monitoring that is independent of the infrastructure, cf. Sect. 1.1. It is also not possible for freight trains today due to the lack of energy on freight wagons but can also be solved with the DAK.

3.1 Condition Monitoring in Passenger Transport

After the Eschede disaster, in which a broken tire caused a series of malfunctions and lead to 101 fatalities, increased efforts were made to monitor the wheelsets with sensors, especially those of high-speed trains.

Wheelset monitoring is still carried out today for ICE 1 and 2 by local trackside sensors that measure the dynamic forces between wheel and rail caused by concentricity deviations of the wheels.

Nowadays, rail vehicles in Europe must be approved according to the Technical Specifications for Interoperability (TSI). Accordingly, at least the following are required (TSI 2014):

- a so-called roll monitoring, which controls whether a wheelset is locked and reports it to the driver, and
- a wheelset bearing monitoring, which for the vehicles with maximum speed of less than 250 km/h can be installed on the track as well.

For the first time, the ICE 3 (series 403/406) was equipped with a wheelset-locking monitoring system and a driving stability monitoring system (instability detection) (Müller and Sunder 2011).

The 2nd generation ICE 3 (series 407, at Siemens AG "Velaro D"), which went into operation at Deutsche Bahn in 2016, has an on-board monitoring system for the wheelset bearing temperature, a redundant roll monitoring system, and a driving stability monitoring system for the bogies. In addition, it optionally has a temperature monitoring system for the engine and gearbox bearings and a damper monitoring system (Steuger 2009).

Currently, wheelset bearing manufacturers provide condition-monitoring systems for rail vehicle bogies, which can be retrofitted as well. In 2011, there was a report on the implementation of a monitoring system by a wheelset manufacturer at the Barcelona Metro. The conditions of wheelset, gearbox, and engine mounts are monitored via temperature and acceleration sensors. Furthermore, gearbox oil temperature and wheelset accelerations are also recorded. The data are saved in the on-board memory, where it can be viewed at any time via a UMTS interface. The system triggers an alarm signal if the specified limit values are exceeded (Kure and Martinez 2011).

Many rail vehicles are equipped with rubber/metal springs, especially in the primary suspension. Rubber has the tendency to creep and to lose its elasticity, which is at least uncomfortable for the passengers but can also become dangerous in terms of derailment potential. Therefore, rubber spring producers are developing sensor equipped rubber/metal springs to supervise the function of the suspension system (Trelleborg 2019)

Monitoring of important train components via sensors has been state of the art for a long time. For example, the crucial functions of doors, air conditioning systems, transformers, traction motors, etc. are monitored and reported to the central control unit (Schulte-Werning et al. 2015).

3.2 Freight Wagon 4.0

Due to the prevailing low technical level, the introduction of innovations in the sense of Industry 4.0 in freight transport is of high priority particularly in the case of freight wagons. Due to the absence of old systems, the **freight wagon 4.0** offers by far the greatest potential. The only problem is that this segment has the greatest cost pressure.

The SBB and the DB Cargo, the latter in the consortium with Europe's largest freight wagon leasing company VTG, are currently working flat out to prepare for the implementation of innovations on freight wagons. The project **5L** ("Low-noise, Light, Powerful, Logistics-capable, and Life-cycle-cost-oriented"[13]), which is funded by the Swiss federal offices for the environment and transport, as well as the project "innovative freight wagon" funded by the German Federal Ministry of Transport and Digital Infrastructure (BMVI), focused on equipping freight wagons with innovative technology that is already available on the market.

In both projects, the wagons were equipped with automatic central buffer couplings. In addition, they have battery-powered telematics devices to which various sensors can be connected.

Measurement of the following variables is currently state of the art:

- current position via satellite navigation technology (mostly GPS at present)
- load weight and its distribution over force measurement sensors
- mileage via telemetry
- shock detection via accelerometers (to estimate whether the wagon structure was overloaded due to excessive shunting impacts with the aim of allocating the repair costs according to the cause)
- derailment detector via accelerometers (Buchmeyer et al. 2008).

In the **innovative freight wagon** project, a digital brake indicator was also employed in the vehicle and by telematics in the locomotive or on the tablet-PC of the wagon inspector, which shows whether the brake is applied or released and in which position (G or P) the brake lever is (Fig. 15). Additionally, the wagons are equipped with RFID tags for better identification and with an electrical control of the brake valve (EP brake) for direct responsiveness (Mues and Galdiks 2017; Klocksin et al. 2018).

In the SBB parallel project "5L," 150 refrigerated wagons were equipped with condition monitoring technology, in which in addition to the abovementioned measurement variables, the temperature and humidity in the cooling chamber and the door closure were monitored (Fig. 16).

The recording and transmission devices provide the power for themselves and the sensors from an integrated battery, which has an operating life of up to 6 years in the case mentioned above. The prerequisite is that the data is not measured and transmitted continuously, but at tolerable time intervals (Mues 2017).

[13] In German, 5L stands for Leise, Leicht, Laufstark, Logistikfähig, Life-Cycle-Cost-orientiert.

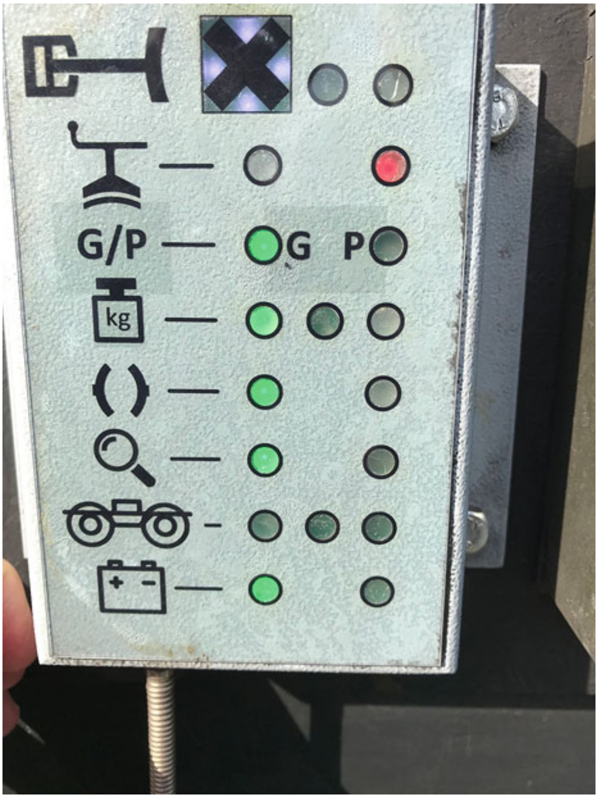


Fig. 15 Brake indicator and telematics equipment in "innovative freight wagons" (Photo: C. Schindler)

Another possibility for energy generation and storage is to install the data acquisition and transmission devices on the axle bearings and to generate electricity using an integrated generator. A battery stores it so that it can be measured and evaluated even when it is at standstill (PJM 2018).

Depending on the energy consumption, electricity can also be generated from the vehicle vibrations via solar elements or linear generators. Shi et al. from TU Berlin provide a good overview (Shi et al. 2018).

Another related concept is suggested by Schmidt et al. from Aachen University of Applied Sciences. Their **intelligent freight wagon** has its own generator and battery-backed power supply, a sensor system for condition monitoring of rotating components, WLAN for communication with the adjacent wagons, wireless train integrity monitoring, and mobile communications for data transmission to the cloud. The pressures in the brake pipe and cylinder are monitored as well (Schmidt et al. 2018).

Instead of WLAN, short-range signal transmission via high-frequency wireless technology would also be possible. It was demonstrated in a BMBF-funded project from RWTH Aachen, TU Kaiserslautern, DLR, and two medium-sized equipment suppliers (Sand and Bürkle 2018).

Fig. 16 Digital train monitoring with wireless transmission equipment (© Bosch Engineering GmbH)

Further considerations of DB Cargo include the visual remote diagnosis of vehicles via camera systems that automatically detect damage or anomalies in the wagons (Thomas et al. 2017). They could be permanently installed or transported to the vehicles by **drones**.

3.3 Infrastructure-Side Vehicle Monitoring

While on-board monitoring of critical components on the vehicle is performed continuously, a system for trackside monitoring (**Wayside Train Monitoring System** WTMS) can only measure pointwise. Therefore, it has the advantage that one measuring unit can record data several times at one vehicle or train, e.g., at every wheelset. Furthermore, the measurement can be carried out for all trains passing through the measuring equipment.

The state of the art includes the so-called **hotbox detectors** (HOA), often in combination with **brake-block detectors** (FBOA), which measure the temperatures of wheelset bearings, wheel tires, and brake disks via infrared sensors and send an alarm signal to the rail traffic controller if a limit value is exceeded (Fig. 17). In the case of internally mounted wheelsets, which are increasingly being used for reasons of weight reduction, e.g., the trailer bogies of the ICE 4, an appropriate equipment is still missing (Hohnstädt and Pachnicke 1992; Eisenbrand 2011).

Other trackside train monitoring devices are:

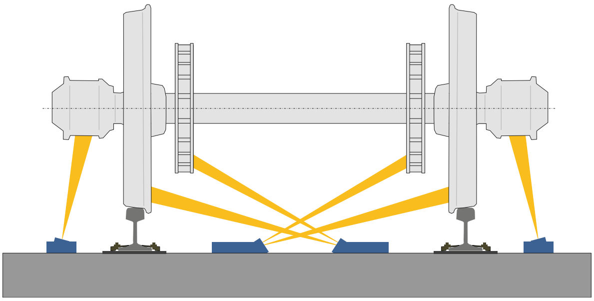


Fig. 17 Combined hotbox and brake-block detectors [Figure: A. Daniel acc. to Eisenbrand 2011, © IFS, RWTH Aachen]

- **Wheel force measuring systems**, which use strain gauges attached to the rail web to determine whether wagons are overloaded or intensively asymmetrically loaded. They also detect gross wheel damage, such as larger flat spots or chipping.
- **Pantograph monitoring systems**, which work optically and/or by measuring the uplift or the contact pressure of the overhead line (indirectly via the overhead wire tension).
- **Loading-gauge monitoring systems,** which optically detect if vehicle antennas or cargo in freight transport protrude from the permitted vehicle loading gauge by laser scanner, light or infrared photography and even measure trains in three dimensions (Vouillamoz and Munter 2011; Bochetti et al. 2011).

The systems, called checkpoints, can be found continuously, particularly in the network of the Swiss Federal Railways SBB (Nietlispach and Frey 2011).

DB is acting in this direction as well. Furthermore, there is concentricity measuring equipment for the wheels (mainly ICE, refer to Sect. 3.1) as well as future laser-optical wheel profile measurement and acoustic wheel bearing monitoring (Schulte-Werning et al. 2015).

3.4 Vehicle-Side Condition Monitoring of Infrastructure

As part of its four-quadrant strategy mentioned above, DB is increasingly using the accelerometers integrated in the bogies of standard vehicles to monitor the track position quality as well (Schulte-Werning et al. 2015).

One of the drivers' tasks is the observation of the track and if available, the parallel track. In this way, he can alert the rail traffic controller about any obstacles that protrude into the loading gauge, so that it can be eliminated. Since the task of track monitoring applies to driverless vehicles as well, it is reasonable to automate it. In the BMVI-funded project ZuG from the Federal Railway Authority (EBA[14]), Deutsche Bahn, a Fraunhofer Institute, the University of Stuttgart, and a

[14] EBA—Eisenbahnbundesamt.

medium-sized software company, digital track monitoring via **image recognition** and evaluation has been developed and tested. The data can be used simultaneously for the maintenance of the infrastructure, which can detect if trees are growing into the loading gauge, or there is a possibility of collapsing, or if the overhead line is sagging too far (Salander et al. 2018).

The Technical University of Munich is conducting another project for the automated detection of infrastructure elements together with a local service provider and a subsidiary of TÜV Süd. In this project, methods of **machine learning** based on **neural networks** for image recognition of infrastructure elements are investigated (Genc et al. 2018).

3.5 Model-Based Condition Monitoring and Digital Twin

Not all data of interest can be determined directly, because it is not desirable to place so many sensors in the vehicle, or the position at which measurements are to be carried out is not accessible, or a sensor is not safe there. In this case, the method of model-based diagnosis helps. The condition data are measured at a suitable and accessible location and fed into a real-time computer model of a specific vehicle system, the behavior of which cannot be measured at the point of interest. The author could not find any implementation or research work on this topic in the rail sector. A scientific work from the field of commercial vehicle technology is mentioned, which similarly could be applied easily to the rail vehicle sector (Engelhardt et al. 2010).

This approach can be characterized as one application for a digital twin, an expression that was initially introduced in the context of product lifecycle management (PLM) in the aircraft industry (Grieves 2005). It contains computer models of the respective technical system and is able to image its characteristics and behavior by simulations and data analysis algorithms. The difference to a pure computer simulation is the fact that the model or virtual system is continuously or at least from time to time fed with measured data from the real world system as an input and can react to the output from the computational analyses. German railways has the vision to model its whole railway system as a digital twin (Rees 2020).

3.6 Usage of Condition Monitoring Data

Currently, the state of the art goes far beyond the mere monitoring of components by comparing the actual values with target values and their thresholds. Particularly in maintenance, data is increasingly being used for service life predictions (**data mining**).

The maintenance strategies of all major railway companies are moving away from the fixed terms of **preventive maintenance** to **condition-based maintenance**, in which the maintenance expert uses the condition data to estimate when a component should be maintained, repaired, or replaced.

The increasing number of condition monitoring data can be used in conjunction with downstream evaluation, learning, and prediction methods to estimate the service life until the failure of the monitored component much more exactly and autonomously and to use its service life optimally. The strategy of **predictive maintenance** is only feasible by acquisition and transmission of digitized sensor data, which are processed by intelligent algorithms so that they provide information about the latest possible maintenance date as reliably as possible before the component failure. In this context, DB refers to as advanced early warning time. Now, this time has to be adjusted to the most suitable date for taking the concerned vehicles out of service for maintenance and to the occupancy schedule in the workshop. It is reasonable to carry out maintenance works that are temporally related, so that the vehicle is not taken out of operation several times in a short period. In addition to the reduced workload in the maintenance workshop and savings in spare parts and fuel, the greatest advantage of preventive maintenance is that fewer components fail unexpectedly during the operation and thus compromise the latter.

The machine-based data processing can be performed in different ways. The easiest way is the condition assessment and prediction based on statistical parameters. However, it is usually only possible with simple correlations between the measured variable or the statistical coefficients derived from it, and the component condition. It could be, for example, the correlation between the vertical accelerations at the axle bearing and the presence of a flat spot on the wheel surface. Sometimes, frequency analysis of the signal can also be used to find a correlation, e.g., with the wheel rotational speed or the frequency of the first bending mode shape of the wheelset axle. Brundisch et al. report in this context on fault diagnosis possibilities at a wheelset bearing (Brundisch et al. 2016).

Today, intelligent methods of failure prediction make use of various options of machine learning. These methods are based on the principle that a programmed algorithm can learn from so-called training data, from which functions can be developed to describe the relationship between the (processed) sensor data and the condition of a component. The better and more numerous the training data are, the more precise the characteristic function will be, which is mostly outputted as a regression with respect to the data. Training data can be sensor data that either are related to data of technical characteristics, e.g., lateral accelerations at the wheelset in relation to the measurable wear condition of the wheel profile or associated with experts' assessment data.

If no or too little training data are available, it is also possible to generate driving tests results by computer simulations. The company Knorr and TU Berlin are performing this as part of the further development of their derailment detector (Friesen et al. 2017). This approach can be included into the digital twin philosophy. Herewith, and with suitable models of the behavior of certain components of the trains included, their operation time-dependent changing conditions can be predicted by these models as well.

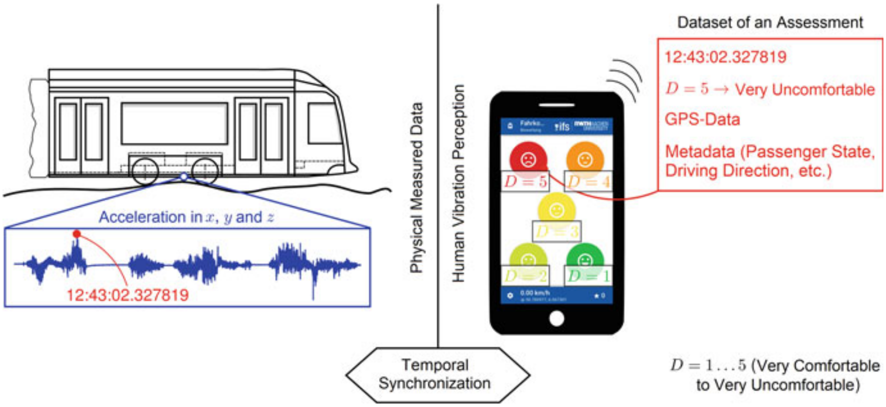


Fig. 18 Simultaneous recording of measurement data and assessment values (Sketch: T. Bettinger, © IFS, RWTH Aachen)

Linke et al. provide a good insight into predictive maintenance at DB without using this term, as well as Bobsien et al. (Linke et al. 2018; Bobsien et al. 2018).

Not only the operators, but increasingly also the manufacturers collect condition data and analyze it. On the one hand, it helps them to detect the weaknesses of their products and to perform improvement measures on the existing product or at the latest on its successor. On the other hand, they increasingly undertake the maintenance for the operators and thus become maintenance specialists themselves (Emmelheinz 2018).

Bettinger provides an example for the acquisition of objective sensor data in connection with subjective assessment data. In a research project at RWTH Aachen University funded by several vehicle manufacturers and the Association of German Transport Companies (VDV), he recorded the accelerations in the car bodies of trams and had test persons to assess the ride comfort during the run at certain time intervals via a smartphone application. Then, he found a relationship between the subjective assessment values and the objectively measured accelerations using methods of machine learning and developed a new method for assessment of ride comfort in rail vehicles. By simultaneous localization via GPS and the driving speed, he converted the measured lateral accelerations and rotational speeds around the vertical axis into track dimensions (straight line, curve radii) and thus correlated to the assessed vibrational comfort of certain points on the track (Fig. 18) (Bettinger et al. 2018).

3.7 *Mobile Sensing*

A more thorough usage of smartphones is emerging from current projects for so-called mobile sensing, the condition monitoring via smartphone measurements

Fig. 19 Sensors in modern smartphones (© V. Brundisch, Bombardier Transportation)

in the vehicle. Modern smartphones have several built-in sensors for acceleration and angular rate measurement, light intensity and proximity measurement, and a magnetic field sensor (compass). In addition, they have a camera and microphone, GPS receiver, as well as mobile communication network-WLAN- and Bluetooth interfaces. Brundisch mentions several simple potential uses, such as vibration measurements, flat spot detection, and measurement of curvature and presents exemplary measurement results (Fig. 19). Simple analysis applications, such as frequency analysis (FFT) and sound intensity evaluation, etc. are available for every smartphone user on the market (Brundisch 2018).

Another project is concerned with the acceleration measurement in the vehicle and processing to obtain the speed, which is already possible for every smartphone user via commercially available apps, with the track dimensions (straight lines and curves) via the angular rate sensor, and detection of switches via the acceleration sensor as well. The so-called Station Buzzer is also of interest, which notifies the passengers if their destination has been reached via a smartphone signal. For this purpose, the smartphone can be connected via Bluetooth with station beacons (Litzellachner 2018).

Berijanian et al. measured the acceleration signals in a passenger train with different medium price smartphones and high-accuracy sensors connected to a professional data acquisition system. Based on the measurement system analysis and the ride comfort evaluations derived from the acceleration data, they compared the smartphone accelerometers with the high-accuracy sensors. They found out that despite of differences in the absolute measuring results, smartphones are suitable for measurements in different vehicles or driving situations (Berijanian et al. 2020).

Brundisch considers the users of mobile sensing as well. He divides them into experts, employees, and passengers. Measurements by experts do not differ from

Fig. 20 Defined position of a smartphone for measurements (photo: M. Berijanian, © IFS, RWTH Aachen)

other measurement groups that use professional devices. However, experts cannot be permanently on site, in contrast to the employees, who can perform measurements in the service with their smartphone once instructed and require little or no additional action. To "collect" a large amount of data, it would be desirable to ask the passengers to measure with their smartphones. For this purpose, their consent must be obtained, and they must be instructed in the use of their devices during the measurement. For example, acceleration measurements can only be rationally carried out if the smartphone is rigidly connected to the vehicle and has a defined orientation, e.g., lying on a shelf (Fig. 20) (Brundisch 2018).

4 Ticketing and Passenger Information

As defined at the beginning, the idea of Industry 4.0 is not only the networking of machines with each other, but also the support of people via better and most accurately timed information (cf. Jaspernite and Niggemann 2012). In particular, the interface to the passenger still leaves many desired options in the public transport. The ticket-purchase process and the associated pricing as well as passenger information before and during the trip have a great potential for improvement through digitalization.

4.1 Ticketing

In addition to the possible options of purchasing a ticket in a travel agency, at the ticket counter of the transport company, or with a ticket machine, almost every transport company now offers an electronic ticket via the company's internet portal or a corresponding smartphone application. These apps are becoming more popular and are developing into powerful tools. Most of them provide, e.g., travel information beyond their tariff zone as well. However, no ticket can be purchased, and it is only possible via the app or web page of the corresponding tariff zone. The problem is that each application is structured differently, and the passengers must find their way again every time (Muth 2018).

The members of the Association of German Transport Companies (VDV) have realized this and are jointly developing a program that can be used to buy tickets from any app of a German operator or tariff association also in other tariff zones in Germany (Wortmann 2018).

The same problem also applies to long-distance cross-border transport. While, e.g., the route Cologne-London can be booked via the Deutsche Bahn app (DB to Brussels, then Eurostar to London), the price information for a Cologne-Paris trip with Thalis still cannot be requested. Superior portals for long-distance travel in Europe are not able to find all travel options and offer only the flight or the bus for many destinations and not the train alternative. The reason could be that the individual transport companies have difficulties in making their existing customer base available to third parties (Haban et al. 2018).

The considerable differences in the tariff systems of the transport associations are still particularly challenging. Therefore, it is not easy for passengers to find the cheapest ticket in a foreign city. An interesting solution was tested as part of a pilot project funded by the Ministry of Transport of North Rhine-Westphalia named nextTicket in the tariff zone of the Rhine-Ruhr transport association (VRR). Using the smartphone app of the same name, the user who has registered once, reserves a ticket that he receives online including the QR code. The system calculates the travel chain and the cheapest fare. If the app is used several times, the system accumulates all trips over a month and retrospectively calculates the lowest price for all trips made in the month. Multiple tickets, 24-hour and weekend tariffs are also considered (Merten et al. 2018).

4.2 Passenger Information

Passenger information is an everlasting nuisance for using public transport. Although the abovementioned smartphone applications that also show delays have accomplished a great deal, much remains undone. Not every passenger looks at his or her smartphone continuously, and even those who do, do not constantly look at the information about their train. Thus, vocal announcement and display information will still be necessary for a long time. In order that the content is understandable with the vocal announcement, **acoustic comprehensibility** is often absent due to the background noise at the platform. To improve it, the Fraunhofer Institute for Digital Media Technology in Ilmenau is working on a software that partially amplifies or weakens the loudspeaker announcements depending on the background noise (N.N. 2016b).

There are large shortcomings regarding the electronic display panels, at least for the DB. Most of the rather expensive boards can only display a few lines of information. Rough information about train delays is given in moving text, while in urban public transport the display of the exact-to-the-minute arrival time of a vehicle has long been state of the art.

Another problem is the **wagon order**. On the platforms in Germany, still a paper sheet in a display provides information about the order in which the numbered wagons arrive at the platform. For operational reasons, the train often arrives in the opposite order. The passenger is then informed of that change by a loudspeaker announcement and if necessary, on the electronic display. However, the usual display in platforms of Germany can only show the order of the 1st class wagon group with respect to the 2nd class group, but not the position of the individual wagons. Examples from France and the Netherlands show that it could be done better (Wikipedia 2017).

Even if the passengers know the current wagon order of their train, it is not completely helpful. The position of the wagon on the platform side is only given roughly. In Germany, 400-m-long platforms are generally divided into six zones (A-F). Since the train does not stop at a defined point, but in an interval of at least plus/minus one wagon length, a more precise position information makes little sense. Higher automation of the train could achieve precise **target braking**, as is already the case with automatic subways today. More specifically, they must stop in such a way that the positions of the vehicle doors match those of the platform screen doors. It has the advantage for the passengers that they can stand at the marked spots of the platform at the door sides to be able to enter quickly. The middle area is left free to allow the passengers to exit the train quickly (Fig. 21).

This measure would not be sufficient for long-distance transport. The passengers could find their wagons immediately. However, since the doors are at the ends of the approx. 26-m-long wagons, they may have queued at the wrong end and have to walk the longer way to their reserved seat through the narrow aisle inside the train (often with luggage). On the way, the passengers in the similar situation approach them in the reverse direction. It would be helpful to let the passengers know where their seat is in the wagon, which would be possible using clearly visible displays on the wagon walls and/or via a smartphone app. Another, possibly simpler way to avoid congestions on the long-distance train would be to reserve the two doors respectively for the entering and exiting passengers and to mark them accordingly.

The feature of actively guiding the passengers to their seats would also include the possibility to optimize the passenger flow and the seating distribution in a way to, e.g., minimize close personal contacts in the train that is to be avoided due to dangers of infections, i.e., in the time of COVID-19 pandemic. Stillfried et al. created a

Fig. 21 Entry and exit conditions on the platform of the Hong Kong subway (Photo: C. Schindler)

simulation on the best strategy of boarding and seating and propose for a passenger coach with only end doors to position the passengers with the longest distance of travel in the middle of the wagon and to put the short distance travelers closer to the doors (von Stillfried et al. 2020).

5 Conclusion

Research, initiation, and implementation of technologies that contribute to the further development of rail transport technology in the sense of Industry 4.0 are at the height. While the railway system has been better automated and networked than the individual private transport for many decades, the automotive sector is now catching up. Especially in automation, the track-guided transport has systematic advantages that should be expanded as fast as possible to get more passengers and goods onto the rails. In particular, automatically operating trains are expected to be more punctual. But, for example in Germany, this can only one true if the number of construction sites on the track network decreases and the bottlenecks at the network nodes (the reason for many construction sites) are eliminated. Furthermore, passengers can be allured to the environmentally friendly rail transport by driverless small rail busses, which for example run every quarter of an hour instead of every hour even on secondary tracks in rural areas.

Although it is rarely addressed, one of the goals of automation is to make train operations more independent of staff shortages, as train cancellations repeatedly occur because the train driver does not appear on duty due to illness. In some cases, the entire lines had to be stopped over weeks due to this problem, refer to, e.g., Wunder (2019). There is a great lack of young-generation applicants not only among the train drivers, but also especially with the shunting staff (Scherer 2017; N.N. 2019b).

The current possibilities of digitization and networking support the maintenance of vehicles and infrastructure as well, which will result in less technical train cancellations. Unfortunately, the problems that occur due to about 800 suicides alone on the German rail system per year have not yet been solved. Possibly, a more intensive track monitoring by drones could timely locate people who come too close to the tracks. DB is already using this technology in vegetation monitoring along the track (Busse 2017).

Furthermore, at least in Germany, there are no additional tracks after years of investment backlog, especially in the metropolitan areas. The investments in the trackside hardware will be relatively lower, since trackside signals and wiring will be eliminated through on-board signaling and secure wireless train control.

In rail-guided road traffic, the technologies that will be established in automobile transport will have to be inevitably adapted. If it is in fact possible that cars can and will be allowed to drive autonomously in cities, trams will also have to possess this technology to be further deployed. The urban transport must therefore be seen and networked as a whole to work autonomously (Fig. 22).

The fastest successes with the introduction of Industry 4.0 features are likely to be achieved in freight transport. By on-board intelligence and energy, the freight wagon of the nineteenth century will become that of the twenty-first century. Especially in the light of the fine dust and CO_2 debates, rail freight transport has great potential for growth. The long-time overdue introduction of the Digital Automatic Center-Coupling will be a big step into that direction.

The reliability of the technical systems deployed in rail transport will increase through on-board detection, networking, and machine learning, which will have a positive effect on the punctuality of rail traffic. The progress in this field is so rapid that it is difficult to distinguish between research and application.

For further information on the topic of rail transport 4.0 or digitization in rail transport, please refer to the book by Dagmar Rees (Rees 2018).

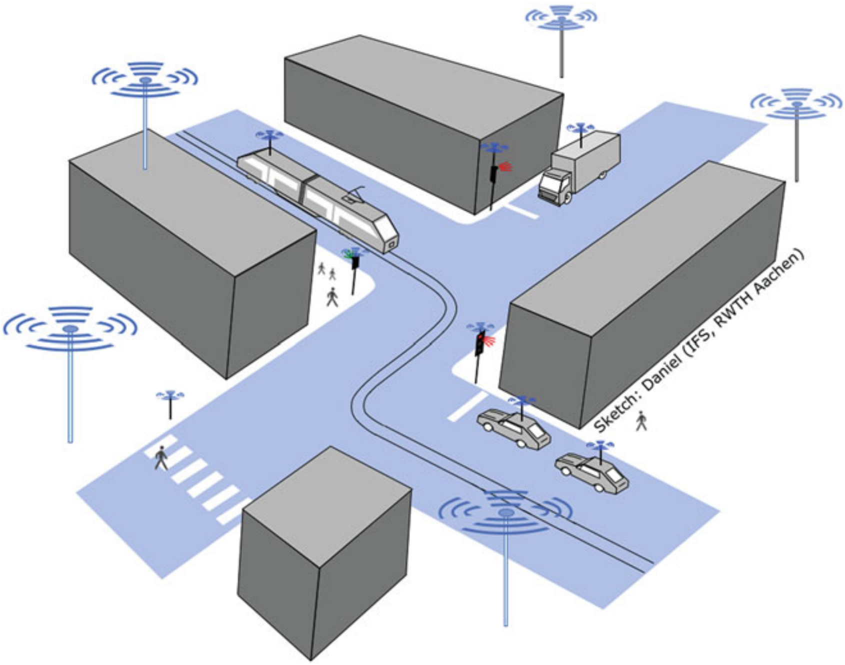


Fig. 22 Scenario of a networked autonomous urban transport (Sketch: A. Daniel, © IFS, RWTH Aachen)

References

ADAC (2018) Autonomes Fahren: Die 5 Stufen zum selbstfahrenden Auto, ADAC 7.11.2018. https://www.adac.de/rund-ums-fahrzeug/autonomes-fahren/autonomes-fahren-5-stufen/. Last access 7 Jan 2019

Bergmeister A (2018) CNA Sonderpreis für den Demonstrator der ersten automatisch fahrenden Rangierlok in Deutschland, Press release of the TH Nürnberg from July 25, 2018. https://www.th-nuernberg.de/fileadmin/newsdaten/Pressemitteilungen/2018/2018_07_Juli/PM_45_TH_Nürnberg_CNA_Sonderpreis.pdf. Latest access 29 Dec 2018

Berijanian M, Hampel F, Leibner P, Schindler C (2020) Accuracy of smartphone accelerometers for evaluation of rail vehicles ride comfort. ETR – Eisenbahntechn. Rundschau. 69 Int. Ed. pp 29–33

Berndt T (2001) Eisenbahngüterverkehr, 1st edn, Teubner Verlag, Stuttgart, p 103 ff

Bettinger T, Leibner P, Schwickert M, Schindler C (2018) Schwingungsbewertungsverfahren - Historie und aktuell laufende Entwicklung methodischer Ansätze. In: 16. Internationale Schienenfahrzeugtagung Rad Schiene - Tagungsband 2018. DVV MediaVerlag, Hamburg, pp 116–118

Bihn F (2018) Autonom fahrende Straßenbahn in Potsdam, In: InnoTrans: Bahnmarkt wächst weiter weltweit jährlich um 2,7 Prozent, Der Nahverkehr 12, p 19

Bobsien S, Niggemann H, Schulze H (2018) Digitale Instandhaltungsstrategie für produktiven Güterverkehr. ETR – Eisenbahntechn. Rundschau 5:54–59

Bochetti G, Mazzino N, Lancia A (2011) TCCS – Train Conformity Check System. RTR Special 2011, pp 45–52

BOStrab (1987) Verordnung über den Bau und Betrieb von Straßenbahnen, Bundesamt für Justiz. www.gesetze-im-internet.de, last change 16.12.2016. Last access 8 Dec 2018
Brundisch V (2018) Zustandsüberwachung mit dem Smartphone, In: Digitalisierung im Schienenverkehr. Sonderheft ZEVrail 142
Brundisch V, Breemersch M, Niu Y (2016) Condition monitoring for rolling stock axle bearings. ZEVrail 140(4):124–128
Buchmeyer G, Walter M, Schmidt S (2008) Der Entgleisungsdetektor EDT 101, EI(4), pp 28–30
Bundesministerium für Bildung und Forschung (BMBF) Hrsg. (2014) Die neue Hightech-Strategie - Innovationen für Deutschland, August 2014, Publikationsversand der Bundesrepublik Deutschland, Rostock
Bundesministerium für Verkehr und digitale Infrastruktur (BMVI) Hrsg. (2016) Zukunftsforum Schiene Digital – 5-Punkte-Strategie. 22.06.2016. Event Consult - Gesellschaft für Messe- und Congress-Management mbH, Berlin
Bundesministerium für Verkehr und digitale Infrastruktur (BMVI) Hrsg. (2017) Masterplan Schienengüterverkehr. June 2017, Berlin
Busse A (2017) Dem Quantum TRON entgeht nichts. EI – Eisenbahningenieur 3:24–26
Claus S, Ammoser H (2018) Digitalisierung der Bahn mit FASSI 4.0, ATZextra – Sonderheft für die VDI-FVT 2018/2019, pp 30–32
Daniel A (2007) Flex-Cargo-Rail: Flexibles System für den Einzelwagenverkehr – Innovatives Konzept setzt auf selbstfahrende Güterwagen. Güterbahnen 6(3):23–27
Deutsche Bahn AG (2015) DB 4.0, https://www.youtube.com/watch?v=2mMc7urGl9I. Last access 7 Dec 2018
Dornis V, Wulf V (2019) Das Wichtigste zum Ergebnis der 5G-Auktion. Süddeutsche Zeitung 13.6.2019. https://www.sueddeutsche.de/wirtschaft/5g-auktion-deutschland-fragen-antworten-1.4484909. Last access 8 Nov 2020
Drescher O, Feldwisch W (Hrsg.) (2017) Neu- und Ausbaustrecke Nürnberg – Berlin (VDE 8) – Das größte Bahnprojekt Deutschlands, ETR-Spezial, December 2017
EBO (1967) Eisenbahn-Bau- und Betriebsordnung (EBO), www.gesetze-im-interne.de, Bundesamt für Justiz, last change 26.07.2017. Last access 8 Dec 2018
Egger M, Zellner C, Stadlmann B et al (2018) Automatisches Tennen von Güterwagen mit Schraubenkupplung. ETR – Eisenbahntechn. Runschau 9:136–140
Eisenbrand E (2011) Hot box detection in European Railway Networks. RTR Special 2011, pp 16–25
Emmelheinz J (2018) 100 Prozent Verfügbarkeit durch Digitale Services. In: Digitalisierung im Schienenverkehr, Sonderheft ZEVrail 142, pp 32–36
Engelhardt L, Bitsch G, Schindler C (2010) Online condition monitoring based on real-time system simulation. In: Proceedings of 1. Commercial Vehicle Symposium Kaiserslautern, pp 529–538. Shaker Verlag, Aachen
Feindert J (2020) Digitale Automatische Kupplung DAK – eine Chance für den Schienengüterverkehr Europas. EI – Eisenbahningenieur 8:13–15
Franzen J, Pinders U, Schreiber W, Kuhlenkötter B (2017) Erhöhung der Sicherheit und Wirtschaftlichkeit des Rangierens durch ein innovatives Assistenzsystem. ETR – Eisenbahntechn. Rundschau 6:43–46
Frederich F (1992) Eisenbahn, quo vadis? Betrachtungen, nicht nur zur Technik. ZEV+DET Glas Ann 116(8/9):260–269
Frederich F (1994) Chaos als Konzept – Schienengütertransport mit selbstorganisierenden Fahrzeugen. ZEV+DET Glas Ann 118(7):329–338
Frederich F, Lege B (1996) SST und SOG, Neuerungen für den Schienen-Güterverkehr. ETR – Eisenbahntechn. Rundschau 45(10):611–616
Friesen U, Herden M-O, Kreisel N, Herrmann T, Götz G, Dongfang S (2017) Lauftechniküberwachung: Erweiterung der Entgleisungsdetektoren für Anwendungen auf festen Fahrbahnen. ZEVrail 141(9):340–345

Gasser TM et al (2012) Rechtsfolgen zunehmender Fahrzeugautomatisierung, BAST – Forschung kompakt 11, Bundesanstalt für Straßenwesen
Genc C, Harmsen F, Köhler T (2018) Automatisierte Erkennung von Infrastrukturelementen mit neuronalen Netzen. Signal+Draht 110(9):48–55
Girnau G, Krüger F et al (2007) Fahrwege der Bahnen. Alba Fachverl, Düsseldorf, pp 256/258
Grieves M (2005) Product lifecycle management: the new paradigm for enterprises. Int J Prod Dev (IJPD) 2(1-2):71–84
Haas K-G (2018) Mehr Kollisionsschutz mit dem Infrastrukturlosen Train Collision Avoidance System (TrainCAS). Signal+Draht 110(9):76–82
Haban C, Köhler A, Merkl D, Pessier R, Sommerfeld T (2018) Servicekette “Informieren – Buche – Bezahlen – Fahren” aus einer Hand. Der Nahverkehr 12:40–43
Haigermoser A (2002) Schienenfahrzeuge, Lecture notes for the course 310.214 at TU Graz, edition 07.2002, p 60, unpublished
Härdi R (2018) Umfassend und ganzheitlich, Interview durch Rees D. ETR – Eisenbahntechnische Rundschau, April 2018, pp 9–11
Hohnstädt K, Pachnicke J (1992) Heißläuferortungsanlagen für Schnellfahrstrecken der Deutschen Bundesbahn, Eisenbahningenieurkalender 1992. Tetzlaff Verl., Hamburg, pp 347–357
Hondius (2015) Die Wuppertaler Schwebebahn erneuert ihren Wagenpark. Stadtverkehr 60(9): 12–17
Janicki J, Reinhard H (2008) Schienenfahrzeugtechnik, 2nd edn. Bahn Fachverlag, Heidelberg, p 124
Jaspernite J, Niggemann O (2012) Systemkomplexität in der Automation beherrschen. ATP Edition September 2012, pp 36–44
Kagermann H, Lukas WD, Walster W (2011) Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. Industriellen Revolution. VDI Nachrichten No. 13, VDI, Düsseldorf
Kagermann H, Walster W, Helbig J (2012) Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Vorabversion des Abschlussberichts des Arbeitskreises Industrie 4.0. Forschungsunion Wirtschaft und Wissenschaft. www.forschungsonion.de. 2.10.2012. Büro der Forschungsunion im Stifterverband für die Deutsche Wissenschaft e.V., Berlin
Knapmöller A, Pritsching C (2018) Pilotprojekt automatisches Rangieren – die Zukunft des Rangierens. Signal+Draht 110(6):23–27
Klocksin J, Turge R, Eckel A, Frenzel J (2018) BMVI-Projekt “Innovatver Güterwagen” liefert erste Ergebnisse und startet Betriebserprobung. ETR – Eisenbahntechn. Rundschau 6:47–53
Kreimeyer A, Lanza G, Weber T, Wittenstein M (2012) Empfehlungen zum Zukunftsprojekt »Nachhaltig bewegt, energieeffizient mobil«. Forschungsunion Wirtschaft und Wissenschaft begleiten die Hightech Strategie. www.forschungsunion.de. February 2012. Büro der Forschungsunion im Stifterverband für die Deutsche Wissenschaft e. V., Berlin
Kure G, Martinez V (2011) Die Zustandsüberwachung von Schienenfahrzeugen in der Praxis. http://evolution.skf.com/de/applied-condition-monitoring-in-railways1/, 19.11.2011. Latest access 3 Jan 2019
Kuschinski N (2003) Im Schutz des hohlen Balkens. Straßenbahn Magazin 1:52–58
Linke P, Jaroszewski D, Hocke C (2018) Ausfallprognose als Baustein der Condition Based Maintenance im Instandhaltungssystem eines EMU. ETR – Eisenbahntechn. Rundschau 11: 56–60
Litzellachner A (2018) Mobile Sensing im SPNV. ZEVrail 142(8):298–302
Malla R (2013) The trend to automation. In: 4th Automated metro seminar, London, 6 September 2013
Masse J-P (2018) SNCF’s deep dive into digital. Railway Gazette Int 11:29–30
Merten F, Binder M, Mentz H-J, Levert J (2018) Testlauf für elektronischen Tarif. Der Nahverkehr 12:35–39
Möbius B, Wolff O, Flege D, Stoffregen M, Westenberger P (2017) Innovationsprogramm Schiene 4.0. November 2017, VDB, Berlin

Molle P (1998) Die Erprobung Selbsttätiger Signalgeführter Triebfahrzeuge (SST) im Güterverkehr der Deutschen Bahn AG. ETR – Eisenbahntechn. Rundschau 47(7):443–446

Mues J (2017) Die Güterbahn der Zukunft – automatisiert und digitalisiert. ETR – Eisenbahntechn. Rundschau 7(8):71–73

Mues J, Galdiks J-E (2017) Die Güterwagen der Zukunft rollen bereits. ZEVrail 141(6/7):239–241

Müller L, Sunder R (2011) Überwachung und Diagnose in Hochgeschwindigkeitszügen (ÜDF). ETR – Eisenbahntechn. Rundschau 1(2):10–17

Muth F (2018) Das Tarif-Abitur wird bald überflüssig, In: ÖPNV-Report Deutschland 2018/2019. DVV-Media VerlagHamburg, pp 20–25

N.N (2002a) Führerstandssignalisierung Zofingen – Sempach in Betrieb. Eisenbahnrevue Int 6: 276–277

N.N (2002b) Siemens stellt fahrerlosen Güterzug vor. https://www.verkehrsrundschau.de/nachrichten/siemens-stellt-fahrerlosen-gueterzug-vor-50155.html, 19.12.2002. Latest access 29 Dec 2018

N.N (2016a) Münchens fahrerlose Airport-U-Bahn – Mit 30 km/h in die Zukunft, Spiegel Online 22.2.2016. http://www.spiegel.de/reise/staedte/flughafen-muenchen-fahrerlose-u-bahn-nimmt-betrieb-auf-a-1078694.html. Last access 7 Jan 2019

N.N (2016b) Verständlich: Software verbessert demnächst Lautsprecherdurchsagen. https://www.ingenieur.de/technik/fachbereiche/verkehr/verstaendlich-software-verbessert-demnaechst-lautsprecherdurchsagen/, published on 4.2.2016, 7:55 am. Latest access 20 Jan 2019

N.N (2018) Beyond ETCS. Railway Gazette Int 174(12):24

N.N (2019a) Probleme mit der automatischen Kupplung von SBB Cargo. Eisenbahn-Revue 11:596

N.N (2019b) DB Cargo chronisch unterbesetzt, SWR, https://www.swr.de/swraktuell/rheinland-pfalz/mainz/DB-Cargo-chronisch-unterbesetzt,db-cargo-unterbesetzt-100.html, 29.1.2019. Latest access 18 Feb 2019

N.N (2020) Digitale Automatische Kupplung – Güterverkehr mit Zukunft. Allianz pro Schiene. https://www.allianz-pro-schiene.de/themen/gueterverkehr/digitale-automatische-kupplung/. Latest access 11 Nov 2020

Nietlispach U, Frey M (2011) Consistently networked WTMS at SBB. RTR Special 2011, pp 60–64

Pachl J (2004) Systemtechnik des Schienenverkehrs, 4th edn. Teubner Verl, Stuttgart

Pachl J (2008) Betriebsführung – Zugbeeinflussung, In: Handbuch Das System Bahn, 1. Aufl. 2008, DVV-Media Verl. Hamburg, pp 533–536

Pachl J (2017) Betriebliche Randbedingungen für autonomes Fahren auf der Schiene, Deine Bahn September 2017, pp 11–19

Peiser S (2015) Effizienter Technikeinsatz mit EOW und Rangierstellwerken. EI – Eisenbahningenieur 4:34–37

PJM (2018). https://waggontracker.com/index.php?id=14, Webpage of PJM Company. Latest access 2 Jan 2019

Poré J (2007) ERTMS/ETCS – Erfahrungen und Ausblicke. Signal+Draht. 99(10):34–40

Randelhoff M (2018) Rio Tinto setzt zum Eisenerztransport in Australien auf fahrerlose Güterzüge. https://www.zukunft-mobilitaet.net/8377/zukunft/rio-tinto-autohaul-autonomer-gueterzug/, last change June 6, 2018. Last access 27 Dec 2018

Rees D (2018) Digitalisierung in Mobilität und Verkehr, 1st edn. PMC Media House GmbH Hamburg

Rees D (2020) Ein digitales Abbild der DB. Interview mit Rolf Härdi. ETR – Eisenbantechn. Rundschau 11:44–46

Rüffer M (2017) Perspektive Autonomes Fahren im schienengebundenen ÖPNV. Deine Bahn 3: 28–35

Rüffer M (2018) Assistiertes, automatisiertes und autonomes Fahren im ÖPNV. Deine Bahn SPEZIAL 9:16–21

Rytz B, Mandour T (2018) Smartrail 4.0 – Gemeinsamer Aufbruch in die digitale Bahnproduktion. ETR – Eisenbahntechn. Rundschau 10:53–55

Salander C, Strobel T, Tavakoli H (2018) Projekt Zustandsüberwachung des Gleisumfeldes (ZuG). Deine Bahn 1:6–11

Sand S, Bürkle D (2018) Sichere Datenübertragung zwischen Zügen mit dem 5G-Mobilfunksystem. https://www.dlr.de/dlr/desktopdefault.aspx/tabid-10081/151_read-30590//usetemplate-print/, last changed on 8.11.2018. Latest access 2 Jan 2019

Scherer B (2017) Lokführer-Nachwuchs dringend gesucht, Saarbrücker Zeitung. https://www.saarbruecker-zeitung.de/wirtschaft/sz-wirtschaft/lokfuehrer-nachwuchs-dringend-gesucht_aid-6883559, 27.11.2017. Latest access 18 Feb 2019

Schmidt B, Enning M, Pfaff R (2018) Güterwagen 4.0 – Der Güterwagen für das Internet der Dinge, Teil 3: Einführungsszenarien für aktive, kommunikative Güterwagen, ETR – Eisenbahntechn. Rundschau 5, pp 60–63

Schulte-Werning B et al (2015) Monitoring und Zustandsorientierte Instandhaltung von Schienenfahrzeugen und –fahrweg mittels Mustererkennung in Ereignisdaten, Presentation by German railways (DB), Modern Rail Vehicles conference (Tagung Moderne Schienenfahrzeuge), Graz, 6.4.2015

Schulte-Werning B, Mösken H, Linke P (2017) DB's advances in condition-based maintenance and failure forecasting for rolling stock. https://www.globalrailwayreview.com/article/71378/db-condition-based-maintenance/, 16.3.2017. Latest access 2 Jan 2019

Shi D, Hecht M et al (2018) Cargo condition monitoring based on autonomous self-powered wireless sensor network. ZEVrail 142(8):290–297

Smagin Y, Popov P (2018) Technologie und Prinzipen des Betriebs lokführerloser Rangierlokomotiven auf dem Rangierbahnhof Luzhskaya. Signal+Draht 110(12):30–38

Steuger M (2009) Velaro – kundenorientierte Weiterentwicklung eines Hochgeschwindigkeitszuges. ZEVrail 133(10):414–425

Stötzel S, Vallentin F (2020) Die Digitale Automatische Kupplung (DAK). Faktenblatt der DB AG. www.deutschebahn.com/presse. 2020

Thomas M, Stöffler F, Häger S (2017) Techlok treibt digitale Transformation, In: Güterbahnen 4.0, Rail Business Spezial 1, pp 28–30

Trelleborg (2019) ConeX3 – 3D Absolutwegmesssystem. Supply Chain Management II, p 104

Trummer G, Rappe J (2008) RUBIN: Die erste fahrerlose U-Bahn im Mischbetrieb. ZEVrail Glas. Ann. 132(9):347–352

TSI (2014) Verordnung (EU) Nr. 1302/2014 der Kommission vom 18. Nov. 2014 über die technische Spezifikation für die Interoperabilität des Teilsystems "Fahrzeuge – Lokomotiven und Personenwagen" des Eisenbahnsystems der Europäischen Union, Amtsblatt L 356/228 der Europäischen Union (Official Journal of the European Union) from 12.12.2014

UITP (2011) Press Kit Metro automation facts, figures and trends. https://www.uitp.org/sites/default/files/Metro%20automation%20-%20facts%20and%20figures.pdf. Last access 07 Feb 2019

Vallée H (2002) Konzept zur fahrzeuggestützten Kollisionsvermeidung im selbstorganisierenden Güterverkehr, Thesis from RWTH Aachen, VDI Fortschritt-Berichte, Reihe 12 Verkehrstechnik/Fahrzeugtechnik, No. 495

von Stillfried A, Schneider P, Rakowitsch M, Schindler C (2020) Möglichkeiten zur Minimierung des Infektionsrisikos von SARS-COV-2 im Schienenpersonenverkehr. ZEVrail 144(8): 310–314

Vouillamoz J, Munter C (2011) Profile and antenna detection system at Heustrich. RTR Special 2011, pp 39–44

Wikipedia (2017). https://de.wikipedia.org/wiki/Wagenstandanzeiger, last edited 6.7.2017. Latest access 21 Jan 2019

Wikipedia (2018a) List of automated urban metro subway systems. https://en.wikipedia.org/wiki/List_of_automated_urban_metro_subway_systems, last edited 29.12.2018. Last access 17 Jan 2019

Wikipedia (2018b) Rangieren. https://de.wikipedia.org/wiki/Rangieren, last edited 15.12.2018. Last access 28 Dec 2018

Wortmann I (2018) “Wir benötigen schnellere Planungsprozesse”, Interview mit dem neuen VDV-Präsidenten Ingo Wortman. Der Nahverkehr 12:28–31

Wunder O (2019) Zahlreiche Zugausfälle: Kranke Lokführer bremsen die Bahn aus, tag24. https://www.tag24.de/nachrichten/rb81-zug-faellt-aus-bahn-lokfuehrer-krank-hamburg-ahrensburg-verkehr-958176, 2.2.2019. Latest access 18 Feb 2019

Zweigel R, Trzukowsky A, Lin J, Abel D (2018) Autonomes Rangieren im Förderprojekt Galileo Online: GO! EI – Eisenbahningenieur 10:44–47

The Industry 4.0-Concept Within Aerospace

Eike Stumpf

Contents

The aerospace sector has a reputation for exploring feasible technology options first, i.e., to foster innovation and for being an early adopter. However, this follows in most cases an incremental development path. In contrast to the public perception, the aerospace sector overall is rather conservative to limit uncertainties and to remain on the safe side. This strategy is essential for sustainable business, as profit margins in aerospace are small and strict safety requirements drive the design and operation. The novel opportunities connected to Industry 4.0 are widely investigated by the stakeholders but undergo at the same time a critical review for securing relevant added value. In fact, several technologies and approaches associated with Industry 4.0 have already made it into application within the aerospace sector.

1 Introduction

The term **"Industry 4.0"** initially merely covered the potentials that arise from digital transformation and real-time communication within production [Davies 2015, p.1]. This special focus on production disregards the technological developments within other technology fields and does not give adequate consideration to the

E. Stumpf (✉)
Institute of Aerospace Systems, RWTH Aachen University, Aachen, Germany
e-mail: stumpf@ilr.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_33

area-wide technical transformation process that is observed since the dawn of the third millennium. Thus, the term “Industry 4.0” is preferably understood in a broader sense, covering the complete lifecycle of a product, respectively of a service, and involves the aspects of all relevant stakeholders. Rarely in industrial history a technological transformation process was so profound: potentially by Industry 4.0 not only the way a product or service is created does change, but furthermore, the way it is used or applied as well as the system context of the product or service.

Industry 4.0 as a concept comes with high aspirations. On the one hand, it reaches out for a substantial performance increase with respect to product or service value creation; on the other hand, it frames the necessary innovation. In that sense, it can be understood as a Fourth Industrial Revolution, following the preceding three: steam engine ca. 1780 / mass production ca. 1870 / computer technology ca. 1970.

The first building block of Industry 4.0 is **“Digitalization.”** Digitalization in literal sense is purely the transfer of data, information, and knowledge into a digital form and the respective storage within a suitable storage medium. In turn, this allows for access and control respectively manipulation of digital contents. Exactly this usage and processing of digital contents make up the **“Digital Transformation.”** The next level of development is reached by comprehensively interconnecting the virtual elements among each other within a network, and furthermore, to embed real elements (“**Cyber-Physical Systems**”). To a certain degree, this allows to overcome natural restrictions imposed by space and time:

- The fast-forward “time-capsule effect” can be seen in the possibility to simulate fully functioning prototypes in their prospective environment already in early design stages and to experience and assess these prototypes in the virtual world, e.g., in a **3D cave** or with the help of a **virtual reality (VR) headset**. Similarly, time can be turned back, e.g., if using virtual scenarios to reconstruct the exact sequence of a crash.
- The attainable independence of space becomes apparent when analyzing industries with globally distributed design and production sites. Teams around the globe work on a central data set in a cloud, with always the next team in western direction taking over the shift according to the time zone. This could mean that the European site, and six to nine hours later the American site, continues the work when the colleagues in Asia or India have completed their business day.

For the digital transformation to build up momentum, certain prerequisites had to be fulfilled:

- Fast and inexpensive broadband connections (cable-based or wireless)
- Inexpensive IT hardware/computing power and data storage capacity (local, cloud-based, etc.)
- Universal standards (protocols, interfaces, data, and software compatibility, etc.)
- Suitable software (modelling environment and language, databases, data analytics, artificial intelligence, etc.)

- Well trained staff (via training courses in industry and implementation in university curricula)
- Reasonable and profitable applications

Especially to enable profitable applications, further technological and business-related developments were needed to unfold in combination with the digital transformation and real-time communication of system elements, e.g.:

- Inexpensive miniaturized sensors and electronic parts
- Inexpensive and fast additive layered manufacturing/3D-printing techniques
- Inexpensive and ergonomic hardware for augmented/virtual reality
- Increased integration level in technical systems
- Demand for diversified/individual products and services

Clearly, the Industry 4.0-transformation proceeds within the aerospace sector, but the observed development resembles rather an evolution than a revolution. Indeed, the digital transformation in aerospace took off more than four decades ago. The process, however, is driven purely by a favorable cost-benefit relation. Thus, today it cannot be concluded whether the aspects of Industry 4.0 will be implemented across the aerospace sector: while the global booking systems are fully digital and automated, the aircraft and satellite production is not. Aircraft, launch vehicles, and satellites are predominantly manually assembled since robots and automated assembly lines need a large lot size and large through-put to be cost effective. Respective production rates are not yet reached in the aerospace sector.

The changes brought about by Industry 4.0 are so complex that this article cannot treat the topic exhaustively. Instead, opportunities opened up by Industry 4.0 (future ones and those already implemented) are illustrated with the help of a limited number of examples from the national and European context.

2 Implementation Examples of Industry 4.0-Concepts Within the Aerospace Sector

The actual and future implementation of technologies and approaches from the Industry 4.0-context differs quite substantially within different parts of the aerospace sector, and with respect to the different stakeholders and the different phases of the lifecycle.

2.1 Research and Development

In research, the areas of pure fundamental research and application-oriented fundamental research have to be differentiated from applied research [Meyer-Krahmer and Schmoch 2004, p. 3]. Both forms of fundamental research (equal to technology readiness level 1) are characterized by hypotheses, creative approaches for problem solving, and build-up of knowledge and understanding. This phase is hardly formally structured and therefore, it is less adequate for application of classical Industry 4.0-approaches.

In contrast, the phase of applied research (equal to technology readiness level 2–4) is predestined for Industry 4.0-approaches. Here, the prospective application and implementation of a product or service is in the focus and requirements, design space, and applicable regulations are known to some extent. Appropriately, one of the central Industry 4.0-technologies, namely the **Digital Thread**, is employed [Risse 2018, p. 17]. The digital tread combines on a chronological and virtual basis the subsequent development steps of a product or service following the lifecycle, if possible, from cradle to grave. The advantage is that subsequent steps can access the volume of data, information and knowledge already built up in prior steps. Different from today this allows for avoiding losses of information at the interfaces between applied research and industry and between industrial units during development (→ fix to observed "over the fence"-mentality of individual units in current development sequence). To support the realization of a digital thread a technology description, as comprehensive as possible, is needed within a development environment or—in a simple case—in a flexible data structure. Initially, this description may be rough and the data structure scarcely populated, but over time, it evolves into a complete description, i.e., the **"Digital Twin."** In some applications, no exact twin is needed. In these cases, a simplified form, i.e., the **"Digital Shadow,"** may be employed. Such kind of data structure, the Common Parametric Aircraft Configuration Schema (CPACS) [Jepsen 2021], is developed at the German Aerospace Center (DLR). Based on XML (Extensible Markup Language), CPACS can easily be extended according to user preferences. CPACS is about to evolve into a quasi-standard in aircraft design.

Increasingly complex to handle are XML-based structures if the aim is to incorporate models on top of simply carrying and administrating parameter values. To cope with this, the concept of Model-Based Systems Engineering (**MBSE**) has been introduced by INCOSE in 2007 [Estefan 2007, p. 14]. Application of MBSE is since growing steadily. MBSE is essential if, e.g., the digital thread is used to contribute to certification ("virtual flight testing") as developed within the DLR research project Digital-X and subsequent projects since 2012 [Kroll et al. 2016, p. 1].

In the context of applied research and development, the virtual assessment of technical solutions within the prospective system environment has been proven beneficial [Stumpf et al. 2011, p. 9]. The single technology performance is usually overrated in lab conditions. Each integration step into the next higher level (e.g., aircraft component → aircraft vehicle → air transportation system) causes net performance losses due to interaction with other system elements and the system environment. To be close to optimal in the final integrated condition, a virtual integration is required upfront in the early research and design phases that allow for a transparency and bookkeeping of performance parameters und thus, enable robust and reliable decision making concerning the preferable technology options

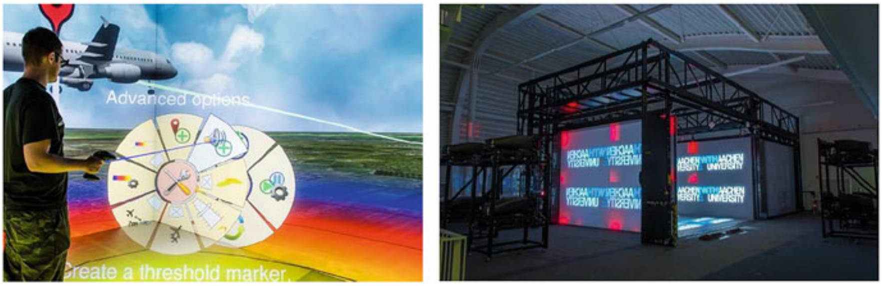


Fig. 1 Virtual Visualization and Auralisation (left), in aixCAVE of RWTH Aachen University (right) [Institute of Aerospace Systems, RWTH Aachen University]

and attributes. Due to inherent uncertainty, the assessment of solutions is usually not done based on absolute values, but on relative comparisons to reference configurations. The underlying hypothesis is that with such procedure the method-based errors cancel out. To promote this procedure the Central Reference Aircraft System (CeRAS, ceras.ilr.rwth-aachen.de) has been developed and maintained at the Institute of Aerospace Systems of RWTH Aachen University [Risse et al. 2016, p. 121]. CeRAS provides to the community comprehensive, fully documented open-access aircraft designs, verified by industry.

If dynamic processes are targeted by research and development, progressively simulation environments are involved. The analysis of complex helix-shaped flight trajectories for limiting noise annoyance on ground to restricted areas for example has been done with extensive simulations, including 3D-visualizations and binaural auralizations in the 3D-cave aixCAVE of RWTH Aachen University, see Fig. 1 [Sahai et al. 2016, p. 24].

Similarly, technologies are developed and qualified in space applications. With the help of **"Simulation-Based Design"** (subsequent development beyond Model-Based Design) the approach and docking manoeuver of the ATV-module (Automated Transfer Vehicle) to the international space station ISS has been optimized and qualified at the Institute of Human-Machine-Interaction of RWTH Aachen University [Thieling and Roßmann 2018, p. 202], see Fig. 2.

Another already implemented technological approach from the Industry 4.0-context is the simultaneous parallel usage of experimental and numerical methods. Among others, this is used for optimizing test setup and test execution in wind tunnels as well as to close temporal or spatial data gaps by employing the best suited approach (experimental or numerical).

2.2 Production

All aerospace companies try to cope with the challenges and opportunities of Industry 4.0. Accordingly, almost all companies have established the position of a

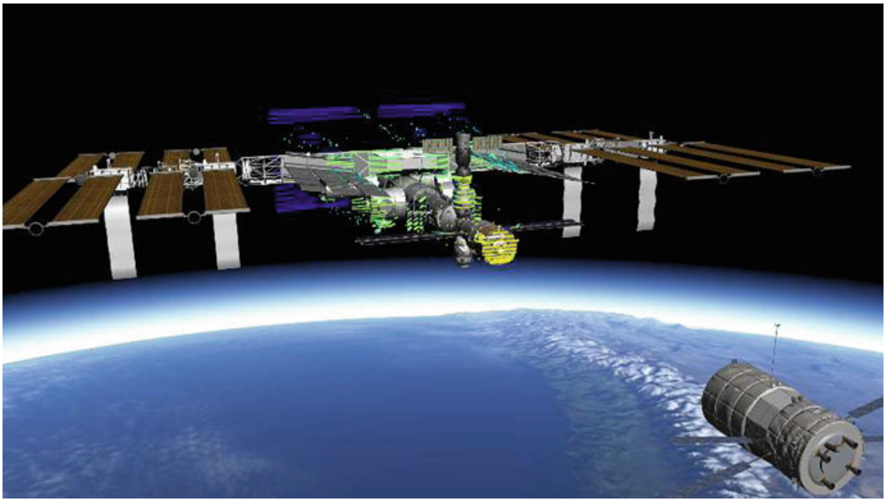


Fig. 2 Approach of Automated Transfer Vehicle to International Space Station [Thieling]

Digital Transformation Officer and they commence at times more internal digital transformation projects for testing and evaluating the new concept than can be effectively monitored and analyzed. Emphasis is put on the implementation of cyber-physical systems together with other key elements:

- Digital thread, digital twin/shadow
- Augmented reality and improved human-machine interaction, e.g., by collaborative industrial robots (**Cobots**)
- Automatic and flexible adaption of production
- Optimization of production with the help of interconnecting all relevant elements, the so-called **Internet of Things**, smart parts and products ("Self-Awareness"), Big Data, and additive manufacturing
- IT-based realization of individual customer-specific requirements
- New forms of services, business cases, and production chains

Within aerospace production, digitalization is no longer entirely new. Trailblazer in this respect has been the aeroengine company Pratt & Whitney. Already in 2002, they decided jointly with IBM and Dassault Aviation to transform the development and production process of future jet engines and to successively establish the digital thread and digital twin.

Problems often arise due to an incoherent digital thread and a non-exact digital twin. Especially transnational enterprises, such as Airbus S.A.S, mainly focusing on integrating contributions from a huge supplier network, risk to struggle to secure coherence and compatibility at all times. The usage of two different CATIA releases has, e.g., been the root cause for the delayed delivery of the first A380 aircraft. A perfectly fitting construction in the digital world, turned out to be faulty in the real world. Due to undiscovered differences within the two used CATIA releases, cable wires on the upper deck did not meet as planned, a few inches were missing. The pragmatic solution of introducing an extra cable segment would have involved increasing the number of connectors and thereby, reducing system reliability. As certification was granted for an architecture exhibiting one connector only, therefore, a substantial re-design had to be done.

In the meantime, powerful commercial development software products have been introduced that combine, e.g., CAD (Computer Aided Design), structural design and PDM (Product Data Management) in one single IT environment. With the help of such environment (here: Siemens NX) the Falcon-9 family of launchers has been developed by SpaceX. SpaceX currently dominates the market of orbital launchers due to unrivaled low transportation costs. The SpaceX launch system has been brought into operation in very short time following the "Fail Fast, Fail Smart"--strategy commonly adopted by start-ups. For this approach to pay off, it was essential to reliably and quickly feedback the lessons learned from failures. Exactly this is the strength of modern development environments and PDM systems, that among others are also offered tailor-made for aerospace application by SAP.

Traditionally aerospace production relies on manual assembly. Due to the low production volume an automation is profitable only in selected application steps, but not for large parts of the assembly line as known, e.g., from automobile industry. Compared to the production infrastructure needed to enable the output of car factories the all-renewed production infrastructure (in 2017) at Stelia in Méaulte, producing all Airbus front fuselage sections, is minor. In total 35 industrial robots and automatic riveting machines achieve a theoretical production rate of 60 front fuselage sections per month. However, the pre COVID-19 production goal of 63 Airbus A320 aircraft per month had to be considered a quantum leap for aeronautics. Few, if any, have expected this compelling market success. Initially, the production rate of the Airbus A320 has been envisioned to reach eight per month after ramp-up.

Within space industry, equally no large production volume is obtained. The production goal of 16 Falcon 9 and ten Falcon Heavy per year as postulated by SpaceX in 2015 cannot be considered a mass production and would not justify high automation levels either. This situation could change if finally, Greg Wyler succeeds with his plans to provide telecommunication and Internet access in rural parts of the world. Large parts of developing countries, particularly the equatorial African region and Sahara Desert, are currently not served ("Digital Divide"). The telecommunication and Internet supply is planned to be delivered by a huge fleet of small low-orbit satellites. The first attempt with the company O3B (which stands for "Other 3 Billion") in 2007 was no commercial success. The second attempt with the company OneWeb was in a better position due to vital partners such as Coca Cola and Airbus, but entered bankruptcy in 2020. The post-bankruptcy company still aims at finally operating a fleet of 650 small satellites.

Apart from exceptions like OneWeb, eventually justifying mass production with high automation levels, the current development trend is rather realizing synergies between human workers and machines. At Airbus the possibility to establish **Cobots**, i.e., the combination of industrial robots and humans in the same workspace, is continuously investigated. The product quality may be furthermore improved by implementing **Augmented Reality**. For time-consuming, complex production steps, such as to arrange harnesses or final assembly checks, the use of specific tablets or glasses that overlay information and/or instructions in the field of vision have proven to be beneficial. These devices allow for an efficient check of the assembly result. In current aircraft several 100 km of electric cables are installed. These cables are arranged manually on a special workbench or pin board and bound together to form harnesses. This work is prone to errors and requires substantial rework if done wrong. Similar in terms of time needed to execute the task, is the check of ten thousands of brackets installed within the aircraft structure. For such kind of application, the Airbus subsidiary Testia developed the smart augmented reality system named MiRA. This allows for efficient component scanning and error detection. Within the Airbus A380 program, the time needed for final inspection was reduced from 300 h to 60 h. Overall, the fraction of incorrectly mounted or missing brackets, detected in later production steps, has been brought down by 40% [Frigo et al. 2016, p. 125]. MiRA has been introduced also at Spirit AeroSystems in the US.

Another approach from the Industry 4.0-context is the usage of flexible assembly lines und processes. Since the beginning of the jet age, the fuselage diameter has been chosen identical for as many aircraft programs as possible (compare identical fuselage diameter of Boeing B727, B737, B757, or Airbus A318, A319, A320, A321) to use the economy of scale effect associated with large numbers of common parts and similar jigs and tools. However, until recently it was not possible to produce components from different aircraft programs in the same production line. Since 2016 this is realized at Airbus with the “Mixed Model Line” producing both wings for the Airbus A330 and A350. Required is a digital control of the production line, usually facilitated via an “**Internet of Things**” (IoT) approach. IoT allows to control any item (tools, machines, facilities, materials, products, etc.) and process individually and enables real-time communication [Arntz et al. 2016, p. 2]. In addition, this brings about a substantial efficiency increase within the supply chain and material stocking.

If the production process is monitored with a multitude of sensors and this data is stored together with the quantitative quality check results for each part individually, then, recorded data from the operational phase can be correlated to feed back this information to the design and production engineers for iteratively improving product quality and for adapting the product for its real working environment, if needed. This fine tuning of the production process together with adaptation of software components of the product opens a new dimension of agility with respect to continuous “ex post” product optimization. This is of high value if operational requirements or conditions turn out to differ from the initial guess or change over time, as is likely to be the case in aeronautical products with a lifetime expectation of at least 25 years. De facto, components might increase performance via downstream learning processes and adaptation enabled by Industry 4.0, compared to the initial specification done by the design engineer (“Better than Engineered”).

Airbus developed a dedicated system to collect data during operation: The data storage and transmission module FOMAX enables the number of continuously monitored and stored aircraft parameters to be increased, e.g., for an Airbus A320 from 400 to 24,000. Furthermore, Airbus established the platform SKYWISE to enable transmission of huge amount of data to the ground and to offer predictive maintenance services to the aircraft operators. Operators that provide their data to SKYWISE for processing receive big data or data analytics services at reduced prices.

Notably high is the extra design freedom for components produced by **additive layered manufacturing** techniques or 3D-printing. These production techniques allow for free-form components that could not be manufactured before, using the classical techniques. Coolant ducts or hydraulic tubes can now be routed in an arbitrary way, which helps reducing weight and achieving compact volume components. The amount of material debris decreases as well. Thus, in Germany the companies Airbus and Liebherr are very active in this field. Problems arise due to different metal material alloys. According to Peter Sander from Airbus [Fischer 2018], only 10 out of 30 currently employed aluminum alloys may be used in combination with additive manufacturing. Spare part provision may largely benefit from additive manufacturing. If one day 3D printers are affordable and globally available, spare parts could be produced on demand anywhere in the world, thereby rendering the nowadays existing complex spare part logistics unnecessary. In 2018, only 10 out of 310,000 Airbus spare parts were eligible to be re-produced with the help of additive manufacturing [Fischer 2018], but these numbers steadily increase. 3D-printing of plastic instead of metallic components is far easier, thus, according to a case study of the World Economic Forum already 1000 printed plastic parts can be found in an Airbus A350 [World Economic Forum 2017, p. 22].

The added product agility enabled by Industry 4.0 allows manufacturers to offer a high number of variants to satisfy specific customer demands beyond the already diverse configuration portfolio in a cost-effective manner. Leadtime may as well be reduced, thus, manufacturers can respond to the market quicker than before.

Smart components offer business potential to the manufacturers in the after sales market. By not sharing details or giving access to the digital core of the components to the MRO companies (Maintenance, Repair and Overhaul) the manufacturer can reserve the vast part of value creation for himself. Classical MRO companies are thereby mainly reduced to a simple installation function with low potential revenue.

Digital transformation lowers the barriers to entering the aerospace market by reducing complexity of design, development, production and certification at the same time. This explains the observed hype of start-ups in the space sector in the context of “New Space” and the innumerable air taxi projects. Even university spin-offs enter the market of air mobility for passengers and goods. Some can draw on broad experience in related fields. In case of the e.SAT GmbH of the RWTH Aachen University with participation of FH Aachen (see Fig. 3), this is the existing knowledge from decades of applied research as well as substantial lessons learned from development, fund-raising and market launch of the ground-based electric cars e.Go and Streetscooter.

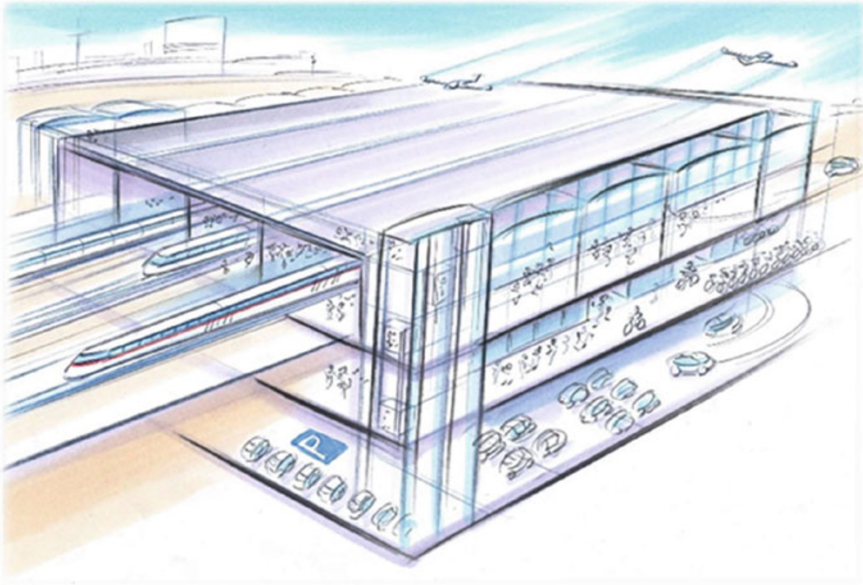


Fig. 3 Concept for eSAT-“Silent Air Taxi”-configuration [eSAT]

2.3 Air and Space Operation

The aforementioned Industry 4.0-solutions are applied in either way in air and space operation.

First, this refers to big data in combination with data analytics. Huge amount of data is collected, e.g., during earth observation. For a large variety of applications, such as weather prediction, climate research, tsunami warning, precision farming, etc., data sets are recorded with the help of satellites and high-altitude drones and are post-processed on ground. Data analytics is employed as well as machine learning algorithms to perform pattern recognition and to derive conclusions. Problematic in this respect is the data storage and data handling of huge amount of data. Airbus has made it a business with the Airbus Defence & Space Cloud Service [Airbus 2014, p. 2].

Spacecraft, satellites and aircraft are long-lasting commodities. Their time of operation often exceeds 25 years. If using these commodities as sensor platforms, e.g., equipped with pressure and temperature sensors or strain gauges at neuralgic locations of the structure and automatic data storage, easily, huge databases may be assembled. The dimension of such data sets becomes apparent if considering the fact that a single Boeing 737 aircraft compiles 20 TB of data per hour in flight [Badea et al. 2018, p. 19]. Via data analytics techniques non-trivial patterns can be recognized, i.e., complex interactions are identified. This may happen individually in each industrial unit that operates a vehicle or can be done on larger scale on commercial platforms, such as Airbus SKYWISE. Data are a resource per se: they can be used for information generation, be exchanged for services, or be directly sold to interested parties.

Collected operational data do allow for drawing conclusions on the state of systems as any wear or defect can be seen in an altered operational behavior and fluctuations in the sensor plots. This serves as basis for predictive maintenance techniques that precisely predict failures and help to safely keep up operation before. Unnecessary cycle-, flight hour- or calendar-based maintenance events may be avoided this way. For this purpose, in most cases digital twins are required that compile, apart from individual geometrical information, the full component history. This comprises of (1) individual production details, (2) history of load cycles, (3) all uncritical (but maybe collectively relevant) events, and (4) all failure cases.

From the collected data equally emission values, atmospheric conditions and flight details can be correlated to train/optimize the autopilot, engine control or flight planning in the airline operation center for similar flights. As described in the previous subchapter, operational data can be fed back to the design and production departments and there, being used for iterative modification, resp. fine-tuning of the serial production. These examples however, relate to a posteriori analysis of data stored onboard in flight. By incorporating IoT in the future, data is supposed to be continuously transferred to the ground for analysis, and delivering—quasi in real-time—insight and instructions for optimizing the flight.

Within jet engines, the approach of **Cyber-Physical Systems** is realized since decades. To control the conditions in the combustion chamber of a jet engine, valid information on pressure and temperature is required. Sensors within the combustion chamber or in close proximity would not sustain the harsh conditions, e.g., up to 1600 °C, for a long time. Thus, a virtual representation of the engine is set up via a real-time simulation based on a simple thermodynamic model. This engine model is integrated into the engine control module FADEC (Full Authority Digital Engine Control). The required simulation accuracy of the virtual engine model is reached by continuous calibration, based on sensor input from the real engine measured at less critical locations. A similar cyber-physical system based on the combination of parallel real-time simulation and real measurements is found at airports. Here, passenger flows in terminals are controlled, based on data fusion of real passenger behavior measurements and simulated passenger behavior, employing agent-based models or models based on the hydrodynamic analogon of passenger movement.

The airline industry was pioneering digitalization. Thirty years back, Lufthansa alone employed almost 1000 staff members to match the revenues with collected paper tickets [Brützel 2017, p. 2]. Since 1 July 2008, no printed tickets are accepted any more. Instead, all tickets are distributed in electronic form (ETIX) and bookings are processed via the global distribution systems, e.g., Amadeus or Sabre. The next development step will be a comprehensive individualization of the services of an airline or travel agency for the customers. The organization IATA (International Airline Transport Association) lied the foundations for this with the introduction of the NDC standard (New Distribution Capabilities). With it, Internet travel agencies and low-cost as well as legacy carriers can better categorize individual passenger data and offer tailor-made products. The passenger is placed in the focus ("Passenger Journey/Experience") and is involved more closely by offering a large selection of choices. Often travel segments that are not covered by air and additional services are included in the package ("Door-to-Door"). The next big transformation process in travel industry is likely to happen in this context [Brützel 2017, p. 6]: enterprises such as Google, Amazon, eBay, etc. draw profound knowledge with respect to demand and preferences of their customers from analyzing queries. With this knowledge and their core competence of Internet-based business they push into the market of Internet travel agencies and global distribution systems. In response, the large global distribution systems have established leading Internet travel agencies themselves (e.g., Amadeus: Opodo, Sabre: Expedia/Orbitz) to cope with the market pressure of the Internet giants for now.

Airports use smart baggage tags with integrated RFID-chips (Radio Frequency Identification). First was Hong Kong International Airport to introduce this technology in 2005 for baggage handling and matching the actually boarded passengers with the individual baggage pieces. A similar approach was considered for localizing missing passengers within the terminal premises with the help of RFID chips printed on boarding passes. Many flights are delayed due to passengers boarding the aircraft late. In aviation a delay is defined as deviation of more than 15 minutes from the flight plan. This delay propagates through the air transportation system and affects other traffic and the operation at the destination airport. However, the technology of RFID chips on boarding passes remained in the trial stage as meanwhile, printed boarding passes are phased out. Passengers by now could be localized more effectively and even be contacted based on their mobile phones. This would be in the interest of airport operators because this additionally might allow to place commercials or advertise products from the retail area. In contrast to the airlines, the airport operators aim at keeping the passengers in the retail area as long as possible as the shops contribute more than half to the overall airport revenue.

The air traffic management is modernized both in the US with NextGen and in Europe with SESAR. Several approaches from the Industry 4.0-context are applied. A crucial component of the novel architecture is a central data pool, the so-called System Wide Information Management (SWIM). This provides a comprehensive set of information to all stakeholders (pilots, air traffic control, airline operation centers, airport operation centers, air navigation service providers, meteorology service providers und military operation centers). This information set comprises of 4D-trajectories, weather information, airport status, air space capacity and aircraft positions, etc. Thus, all stakeholders have an equal level of knowledge and decisions can be made quicker and more efficient.

As final example, the **automation** in air space operation might serve. Since the early days space operation is highly automated and flight operation is partly automated. Industry 4.0-related technology will successively support the development to refrain from placing human astronauts in space. Substantial costs and environmental damage is caused by the current necessity to launch humans into orbit for scientific or maintenance tasks. With reliable full automation and agile robots this can be avoided.

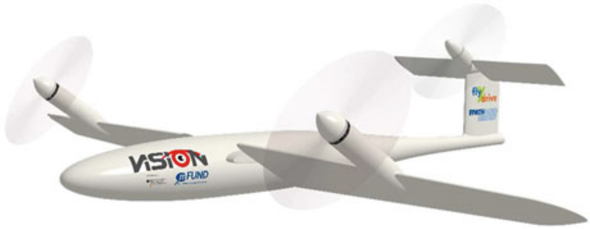


Fig. 4 Unmanned Aerial System Project of RWTH Aachen [Institute for Flight System Dynamics, RWTH Aachen University]

With the potential to establish unmanned aerial systems (UAS, see Fig. 4) as mass market, once more the zero-pilot operation of passenger aircraft is discussed. The necessary technology has been ready for quite a while, but public acceptance cannot be taken for granted. Slowly this situation is changing as partly higher automation levels are introduced in the automobile sector. This steadily leads to a general familiarization with this novel operational concept. A regular operation of autonomous passenger airplanes and air taxis can be expected in the near future.

Notably different, compared to the past, is the new certification strategy: for larger aircraft a complete, strict set of regulations is applicable to a wide range of aircraft sizes, and airworthiness has to be proven for the complete flight envelope. If this strategy would be equally applied to UAS, a certification would in many cases be impossible. Thus instead, UAS are certified for specific missions only, based on SORA (Specific Operations Risk Assessment). With this flexibilization EASA (European Aviation Safety Agency) paves the way for introducing this new form of mobility by digitalization.

3 Challenges of Industry 4.0 in the Aerospace Sector

The Industry 4.0-transformation causes large investment costs for the aerospace sector. According to a survey done by PwC in 2014 [Koch et al. 2014, p.7] German enterprises were planning to invest 3.3% of their total revenue to realize the digital transformation. This accumulates to 50% of total investment volume. In return, a production efficiency increase of 18% was expected.

The cumulative benefit expressed as increased efficiency, reduced costs, smaller lead times, increased agility and better customer demand satisfaction has not yet been assessed in a quantitative manner. However, the multitude of examples given in the preceding chapters proves that the advantage of Industry 4.0 within the aerospace sector has been recognized.

Within the initial euphoria with respect to Industry 4.0-potential, the dimension of the upcoming transformation effort could only be judged rudimentary. In fact, apart from amortization several challenges from very different fields lie ahead:

- Data quality: The final result is in any case largely dominated by available input data. Thus, data quality has to be secured on highest level. This relates among others to:
  - Sensor quality
  - Fidelity of simulation
  - Sustainable (consistent) maintenance of digital twins
  - Effective cleaning up of large data sets
- Control/Verification: Findings obtained by data analytics can seldom be verified, but usually only be checked for plausibility. Similarly, automated systems, especially when involving artificial intelligence, may develop self-induced momentum that defies effective control, if no precautions are taken in advance.
- Standardization: The access and usage of common data sets, the interconnection and encryption require standardized protocols, interfaces, data, and software compatibility to secure data exchange.
- Frequency Bands: The number of radio frequencies is limited. These frequencies are assigned proactively by the German Federal Network Agency. A future alternative may be general data transfer by light (cable-based and wireless).
- Energy Demand: The multitude of sensors, IT hardware, robotics, data transfer systems, etc. within the Industry 4.0-context bring about an electric power consumption not to be underrated (operation and cooling). This causes a substantial complexity with respect to the energy supply.
- Security: The Industry 4.0-technologies open up a large field of possible, systematic manipulations and represent as such an overall not yet covered security risk:
  - Cyber security tries to catch up with the threats of malware, spyware and virus attacks as well as jamming or spoofing, etc. For navigation and communication Galileo now offers PRS (Public Regulated Service), i.e., a service with security level similar to those used by agencies.
  - Similarly, the dependency on satellite-based services is increasing: large amount of data is transferred via Satcom, all aircraft and flying platforms use GNSS (satellite-based navigation) for safe flight operation, most drones and the future air taxis cannot be operated without available GNSS, geofencing (virtual closing of specific air spaces) requires GNSS, etc.
  - All electronic devices are prone to strong electromagnetic radiation. The disturbances range from undesired, but tolerable electromagnetic interference of systems to the extreme case of tactical nuclear weapons that may destroy any electronic device by electromagnetic pulse in a certain radius.
- Intellectual Property Rights: The joint and possibly inter-organizational work on digital contents and the easy access directly raise questions concerning the permanent protection of intellectual property rights. Sometimes, it is even troublesome to constitute a legitimate claim.

Fortunate circumstances have placed the Industry 4.0-transformation at the beginning into a period of steady growth and a huge production backlog in the aeronautic industry and into an already ongoing transformation in space industry, namely “New Space.” For the new players in the aerospace sectors—be it in the orbital launch or micro-satellite business or advancing urban air mobility—the Industry 4.0-transformation is without a doubt beneficial. Overall, the digital transformation catapults aerospace to new levels of process speeds, performance, cost- and eco-efficiency, transparency, and participation for:

- A now feasible focus on efficient unmanned space operations
- A sustainable, passenger-centered air transportation system

Meanwhile, the pandemic struck and changed the game. These are unprecedented times. Empirical knowledge no longer holds, the once high level of predictability of the aerospace sector is gone. At this point in time only scenarios describing the future exist. Scenarios in itself are consistent. That is to say, they are possible. But scenarios do not come with an information on likelihood with respect to their occurrence. Thus, the extreme level of uncertainty remains, but still, preparation for the future is mandatory. It is to be hoped that the current situation is taken as a chance, that industry finds itself in “reset”-mode. If so, the current situation might act as a catalyst and speeds up the overdue digital transformation that in turn will help to re-build the aerospace business.

References

Airbus (2014) Intelligence – Cloud Services. Broschure, https://www.intelligence-airbusds.com/files/pmedia/public/r20805_9_int_044_cloudservices_en_low.pdf, Accessed: 15.3.2021

Arntz M, Gregory T, Lehmer F, Matthes B, Zierahn U (2016) Arbeitswelt 4.0 - Stand der Digitalisierung in Deutschland – Dienstleister haben die Nase vorn. IAB Report 22/2016, pp 1–8

Badea VE, Zamfiroiu A, Boncea R (2018) Big data in the aerospace industry. Informatica Economică 22(1):17–24

Brützel C (2017) Digitalisierung und Disruptive Innovation in der Airline-Industrie. Gastbeitrag Airliners, http://www.airliners.de/digitalisierung-disruptive-innovation-airline-industrie-aviation-management/41243, Accessed 15 Mar 2021

Davies R (2015) Industry 4.0 – Digitalisation for productivity and growth. Members’ Research Service PE 568.337 EN, Briefing, pp 1–10

Estefan JA (2007) Survey of model-based systems engineering (MBSE) methodologies. INCOSE MBSE Focus Group 25(8):1–70, http://www.omgsysml.org/MBSE_Methodology_Survey_RevB.pdf. Accessed 15 Mar 2021

Fischer J (2018) Airbus-Werkstatt: Ersatzteile aus dem 3-D-Drucker. Interview in Weser Report 10. October 2018, 2018, https://weserreport.de/2018/10/wirtschaft/airbus-werkstatt-ersatzteile-aus-dem-3-d-drucker/. Accessed 15 Mar 2021

Frigo MA, da Silva ECC, Barbosa GF (2016) Augmented reality in aerospace manufacturing: a review. J Ind Intell Inform 4(2):125–130

Jepsen J (2021) CPACS – A Common Language for Aircraft Design. http://code.google.com/p/cpacs/. Accessed 15 Mar 2021

Koch V, Kuge S, Geissbauer R, Schrauf S (2014) "Industrie 4.0 – Chancen und Herausforderungen der vierten industriellen Revolution. PwC, Seiten 1-52, https://www.strategyand.pwc.com/media/file/Industrie-4-0.pdf, Accessed 15 Mar 2021
Kroll N, Abu-Zurayk M, Dimitrov D, Franz T, Führer T, Gerhold T, Görtz S, Heinrich R, Ilic C, Jepsen J, Jägersküpper J, Kruse M, Krumbein A, Langer S, Liu D, Liepelt R, Reimer L, Ritter M, Schwöppe A, Scherer J, Spiering F, Thormann R, Togiti V, Vollmer D, Wendisch J-H (2016) DLR Project Digital-X: towards virtual aircraft design and flight testing based on high-fidelity methods. CEAS Aeronaut J 7(1):3–27
Meyer-Krahmer F, Schmoch U (2004) Zur Bedeutung der angewandten Grundlagenforschung – Ergebnisse einer schriftlichen Befragung bei Fraunhofer-Instituten. Fraunhofer ISI Discussion Papers, Innovation and System Policy Analysis, No. 4/2004, ISSN 1612–1430, pp 1–18
Risse K (2018) Das virtuelle Flugzeug – Digitalisierung der Luftfahrt. DGLR Luft- und Raumfahrt 3:15–17
Risse K, Schäfer K, Schültke F, Stumpf E (2016) Central Reference Aircraft data System (CeRAS) for research community. CEAS Aeronaut J 7(1):121–133
Sahai AK, Wefers F, Pick S, Stumpf E, Vorländer M, Kuhlen T (2016) Interactive simulation of aircraft noise in aural and visual virtual environments. Appl Acoust 101:24–38
Stumpf E, Langhans S, Weiss M, Sun X, Gollnick V (2011) A methodology for holistic air transport system analysis. In: 11th AIAA ATIO Conference, AIAA 2011-7010. Virginia Beach, pp 1–11
Thieling J, Roßmann J (2018) Highly-scalable and generalized sensor structures for efficient physically-based simulation of multi-modal sensor networks. In: Proceedings of 12th International Conference on Sensing Technology (ICST 2018), pp 202–207
World Economic Forum (2017) Digital transformation initiative – aviation, travel and tourism industry. White Paper:1–40

BIM and the Digitalization in the Construction Industry

Jörg Blankenbach and Ralf Becker

Contents

1 Introduction

The **digital change** is a global megatrend that leads to disruptive changes in many areas of society. In industry, this transformation process is often described by the synonym "4th industrial revolution" after mass production, electricity, and

J. Blankenbach (✉) · R. Becker
RWTH Aachen University, Geodetic Institute and Chair for Computing in Civil Engineering & Geoinformation Systems, Aachen, Germany
e-mail: blankenbach@gia.rwth-aachen.de; becker@gia.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_34

computers (Frick 2017). **Digitalization** is currently a key challenge in many sectors, whereas the **construction industry** (AECO industry) in comparison, has a need to catch up, especially in the European countries (Fig. 1).

Digitalization in the **construction industry** is primarily linked with a novel collaborative working method, ***Building Information Modeling (BIM)***. BIM promises not only a higher planning quality for construction projects, e.g., in terms of schedules and costs reliability, but also an increase in efficiency in management, information selection and presentation/visualization as well as exchange of construction-related information during the entire **construction life cycle**.

Planning, construction, and operation are highly interdisciplinary tasks of various disciplines. Optimal value creation therefore requires a seamless **digital process chain** and a seamless exchange of information between all parties. The use of integrated **digital models**, the application of open standard as well as the data management in suitable open shared data environments throughout the entire building life cycle are central building blocks for fully exploiting the benefits of collaborative work.

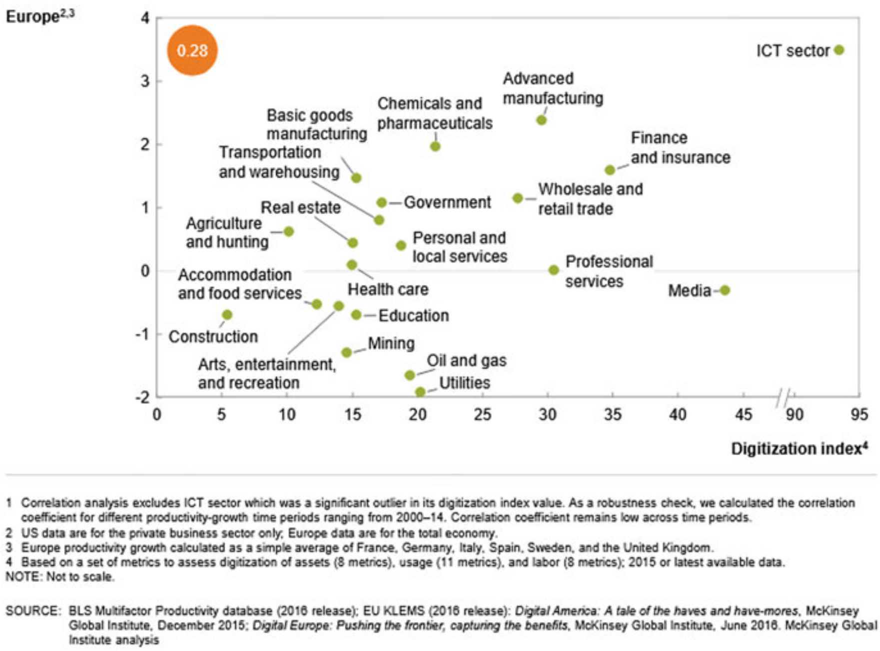


Fig. 1 Digitization Index in Europe (Source: Remes et al. 2018)

2 BIM: Building Information Modeling

The BIM method is not fundamentally new and has been a topic of discussion for many years, especially in science (e.g., van Nederveen and Tolman 1992). The U.S. National Institute of Building Sciences (NBIM) described BIM in 2007 as a product, collaborative process, and facility lifecycle management requirement (NBIM 2007). The EU BIM Task Group thinks of BIM as “digital construction” which combines the use of 3D computer modeling with whole life asset and project information to improve collaboration, coordination and decision-making when delivering and operating public assets (EU BIM Task Group 2017).

2.1 Definition

Many countries found their own definition corresponding to the first mentioned by van Nederveen and Tilman. Currently, NBIM defines: BIM “is a digital representation of physical and functional characteristics of a facility. As such, it serves as a shared knowledge resource for information about a facility, forming a reliable basis for decisions during its lifecycle from inception onward” (NBIM 2015). UK defines “BIM or Building Information Modeling is a process for creating and managing information on a construction project across the project life cycle” (NBS 2016a). The former German Federal Ministry of Transport and Digital Infrastructure (BMVI), today Federal Ministry for Digital and Transport (BMDV) defines BIM as “a **collaborative work** method that creates and uses digital models of an asset as a basis for the consistent generation and management of information and data relevant to the asset’s life cycle as well as for the sharing or passing on of such information and data between the participants for further processing by way of transparent communication” (BMVI 2015). In all cases, the focus is a jointly used, central

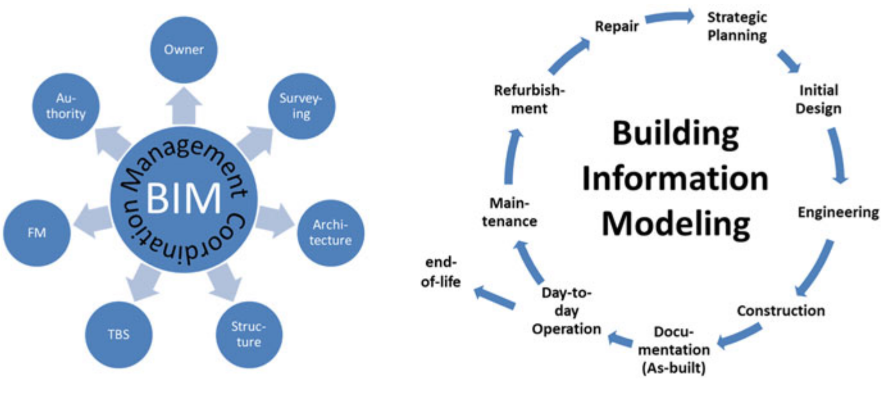


Fig. 2 left: BIM and Building Trades, right: BIM in building life cycle according to ISO 19650

data base in the form of a digital, holistic, and uniform **building database** (Fig. 2 left).

It contains the **digital models** where data are already supplied by all parties involved during the planning and construction phase as well as being used and maintained throughout the entire **life cycle** up to the refurbishment or deconstruction (Fig. 2 right) (Becker et al. 2018).

Thus, the method **BIM** addresses not only the planning and construction phase but also operation and management (e.g., facility or maintenance management). BIM is associated with the expectation that coordination between the technical disciplines will be improved through significantly higher transparency. Inconsistencies and errors should be detected and corrected in early planning phases.

2.2 Use Cases

Although the application of BIM addresses the entire life cycle, in practice BIM was initially introduced primarily in the design and planning phases of new buildings. Many countries around the world started initiatives and road maps to establish BIM in their AEC processes. Already in 2003, the US General Services Administration (GSA) formulated the National 3D-4D BIM Program (GSA 2020). The UK government started its Digital Built Britain strategy in 2011 (Cabinet Office 2011). In Germany, based on the **Road Map for Digital Design and Construction** and the **Masterplan Bauen 4.0** (BMVI 2015, 2017) BIM is now increasingly used in infrastructure construction, i.e., roads, railways, bridges, and waterways. Internationally, this is particularly evident in the intensive activities to expand standardization in the area of data modeling and data exchange (see Sect. 5) (buildingSmart 2020). The **use cases** of BIM in the building life cycle are diverse and increasingly extend beyond the planning and construction phase to the operation phase, including:

Planning:

- As-is data acquisition
- Visualization and derivation of sections and plans
- Component lists / bill of materials, area, volume, and **quantity determinations**, technical specification
- Model-based cost calculation and time management
- **Coordination and clash detection**
- Simulations and variant studies

Construction:

- Construction (progress) documentation/control
- Defect management and documentation
- Creation of execution plans, site management, and logistics

Handover/commissioning/approval:

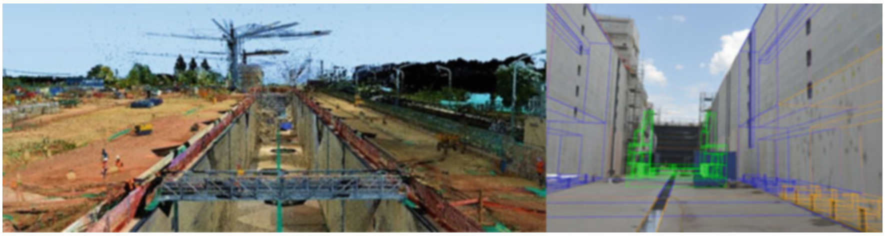


Fig. 3 New Trier lock: **left:** Colored point cloud of a laser scanner, **right:** digital as-built model

- As-built documentation
- Billing of construction works

Operation phase:

- As-is documentation
- (**Computer Aided) Facility Management** (CAFM)
- **Maintenance Management**
- Deconstruction/dismantling planning/recycling

The use of BIM in the **operation phase** requires a documentation model that reflects the actual state of the structure, i.e., an existing planning model must be checked against reality (**as-built**). If the model is non-existent (e.g., in the case of **construction in existing buildings**), the digital model must first be created by means of a suitable reality capturing and modeling technique (**as-is**). An example for the application of BIM in the life cycle is shown in Fig. 3. To enable a cross-checking with the planning and to be able to use it as a **Digital Twin** of the actually built condition during the operating phase, e.g., for asset management, an as-built model of a newly constructed infrastructure (ship lock) was derived from laser scanner and image data.

Currently, the advantages of BIM are mainly seen in more complex structures. This is also expressed by the fact that the legal requirements for the use of the BIM method in various countries are linked to **minimum construction volumes**, e.g., in Denmark from 2.7 million euros (Goltermann and Grube 2017), in Singapore the obligation to electronically submit construction documents in IFC for public projects since 2004, (Khemlani 2005), and since 2015 for tendering of more than 5000 m^2 of floor space (Singh 2017). In Germany, in the area of building construction for public buildings with a construction cost volume of 5 million euros or more, the use of BIM must be examined to determine whether it makes sense (BMUB 2017). It remains to be seen to what extent BIM will become more widespread in companies as its advantages are also established for less complex objects.

2.3 Overall and Aspect Models: Collaborative Work

When working with BIM, the involved disciplines create **aspect models** for their trades using specialized software. According to Egger et al. (2013), Reif (2017), Tulke (2018), DB Station&Service AG (2019), the following are among the aspect models:

- Environment model (terrain model, vicinity from a city model)
- Stock model (topography, construction and technical assets, surveying, terrain model)
- Mass model (urban classification)
- (Official) site plan for the building application
- Architectural model (in various level of development) (Fig. 4)
- Structural model (in various level of development)
- Technical building services models (TBS) (in various level of development and disciplines)
- Site equipment model
- Construction process model (4D model)
- Construction and assembly model
- Construction handover or documentation model
- As-built model
- CAFM model

In addition to the above-mentioned models, supplementary aspect or sub models, e.g., from other participating disciplines, can be added. If technical, functional and commercial aspects such as **time and cost management** etc. are introduced in

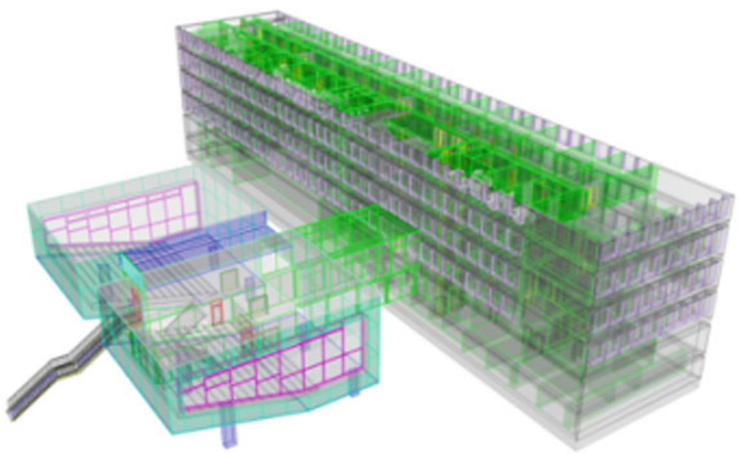


Fig. 4 BIM architecture model (as-is) of the faculty building for Civil Engineering of the RWTH Aachen University

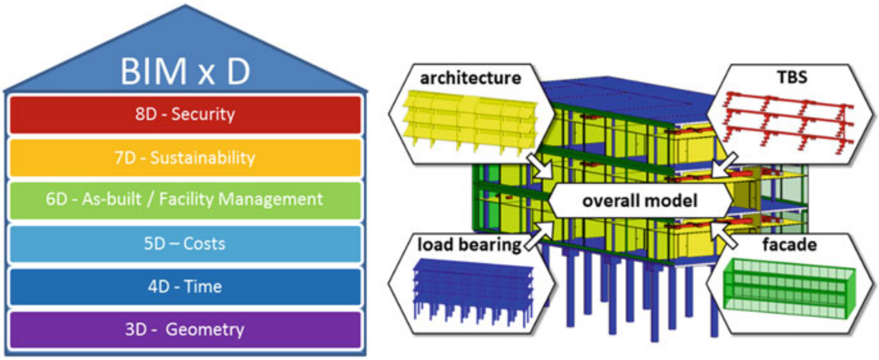


Fig. 5 **left:** BIM xD (following Smith 2014, Eastman et al. 2011), **right**: BIM—overall and aspect models

addition to the structural properties of a building, a multidimensional BIM is created (Fig. 5, left).

For collaborative work, the specialized aspect models are combined into a superordinate **overall model** (also referred to as **coordination model**) by means of suitable data interfaces (Fig. 5, right). Only in an overall model software-supported model checking (e.g., clash detection) can be used to detect conflicts between the planning of the various trades (e.g., between the technical building equipment and structural models) at an early stage and to assign remediation tasks to the aspect model planning.

The distinction between **closed/open** and **little/big** BIM and their permutations has become established for the degree of collaborative action. Here, the term "closed" stands for the use of only one single software as opposed to "open" for the use of software from different manufacturers using open standards for data exchange. "Little" refers to the use of BIM only for individual trades as opposed to the use of BIM from several or ideally all disciplines ("big").

2.4 BIM Implementation Process

The introduction of BIM goes along with paradigm shift in construction practice that will not succeed abruptly. The introduction of BIM is therefore a process of gradually increasing digitization of data, data storage, data exchange and (collaborative) processes. The way of working is changing from the use of analogue plans to file-based data exchange and the use of data clouds. Similarly, the classic, often still two-dimensional, CAD-based working method is changing to object-based work in 3D. The transition to digital exchange and digital processes requires the introduction of open standards for data exchange, the introduction of **component libraries** and the standardization of processes. In many cases, e.g., by the British BIM Industry

Working Group, the introduction of BIM is broken down into so-called **maturity levels** (BIS 2011).

3 Modeling Paradigm

Traditionally, for design and planning of constructions two-dimensional plans and sections are utilized. In many cases the plans are still “drawn” with simple lines in CAD systems, which is the digital equivalent to drawing on paper. Different meanings—semantics—are assigned to the lines through coloring, line types and line thicknesses. For additional differentiation, e.g., of different trades, layers are created. Further semantics are documented in the (CAD) drawings by labelling. These have—apart from the placement—no relation to the object.

The modeling paradigm of the BIM method is completely contrary to this because the focus is not on the drawing but on the **component** as basic information carrier. The component carries all design information as well as physical, functional, and descriptive properties (cf. Sect. 3.1). The entirety of the **objects** forms the building information model. By including organization and time schedules as well as costs it becomes a 4D or 5D **building model**. Such a building model enables a variety of analyses and simulations such as the creation of specified material lists, cost planning or mass or volume calculations (see use cases in Sect. 2.2).

3.1 Object-Oriented Modeling

Each component or object has properties. In addition to the shape, defined by the **three-dimensional geometry**, these properties include descriptive attributes as semantic information, relations to other components (neighborhood, group membership) and if applicable display information such as color or texture (Fig. 6).

Thus, the focus of the modeling is not on the basic geometric elements such as points, lines, or surfaces but on the components such as walls, floors, windows, and doors. The geometry is only one object property of many, commonly represented by a parametrically described solid body. Figure 7 shows, as an example, the properties of a selected building component in a BIM authoring software.

3.2 Level of Development and Accuracy

Planning of a building is a multi-stage process starting with the preliminary or conceptual design planning and ending with the fabrication planning, which is characterized by the increase of detail in each stage. In BIM, the term **Level of Development** (LOD) of a building model has been created for this purpose. The

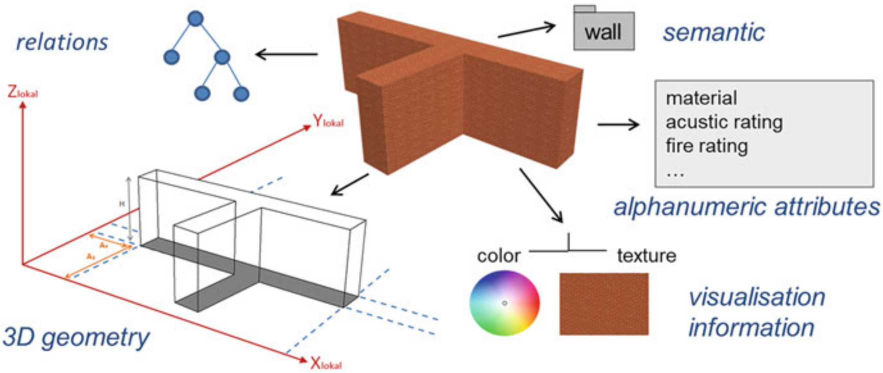


Fig. 6 Semantic data models——objects with attributes and relations

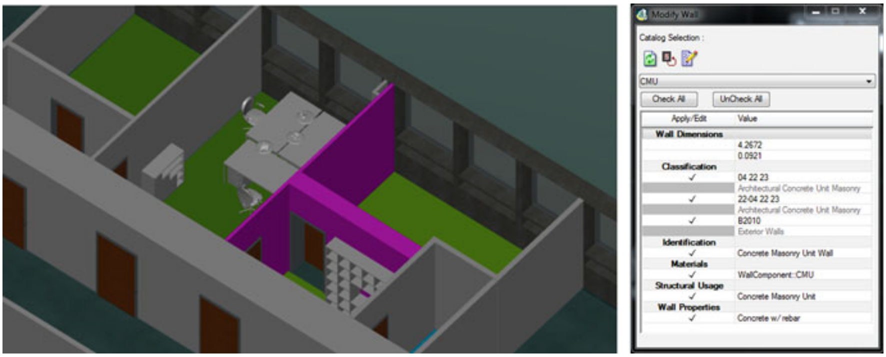


Fig. 7 “Wall” component with descriptive properties in the BIM software (here: Bentley AECOsim Building Designer)

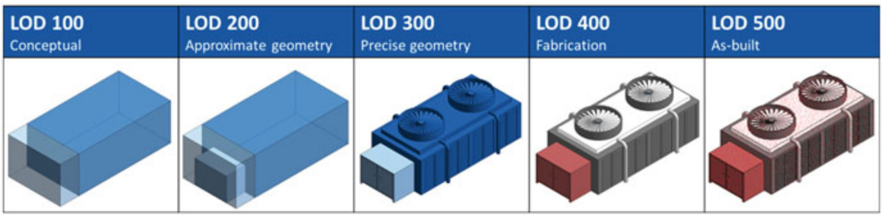


Fig. 8 LOD (according to NATSPEC 2013)

LOD usually describes the degree of completion in five to six main degrees from a purely symbolic representation in a conceptual planning phase to as-built modeling (Fig. 8) (cf. Egger et al. 2013; NATSPEC 2013; BIMFORUM 2016). The LOD can be divided into a geometric (the **Level of Geometry**, LOG) and a semantic part (**Level of Information**, LOI). Examples of LOG and basic LOI can be found in one of the most recognized specifications, the US LOD specification (BIMFORUM 2016).

European standardization (CEN) defined the **Level of Information Need** (LOIN) (EN 17412-1:2020). While the above-mentioned LOD refers to the level of development as a property of the model, the LOIN defines the level and extend of information that is required to fulfil a specific purpose throughout the life cycle of built assets. In addition to the LOG and LOI, LOIN also contains requirements for documentation (e.g., of an object).

To specify the accuracy of an as-built (LOD 500) or as-is building documentation, the USIBD **Level of Accuracy** Specification Guide (Level; LOA) was published in addition to the LOD by the U.S. Institute of Building Documentation (USIBD 2016). Corresponding in content to the specifications in the German standard DIN18710-Ingenieurvermessung, it knows five classes of measurement accuracy. LOA summarizes the measurement accuracy to be achieved by the used data acquisition technologies and—to be considered separately—the geometric accuracy of the modeling process.

4 Processes

BIM is not only a model, nor only the modeling process, but above all a method for optimizing collaboration in digital processes. Collaboration requires agreements on the timing of the interaction between the different actors, the **work processes**. To develop, coordinate and define them clearly and bindingly for all participants is an essential part of BIM.

4.1 Information Management

The advantages of the BIM method are based on the consistent **collaborative work** with digital data. Therefore, clear and unambiguous regulations for the work processes associated with the BIM method are needed. They have been developed for several years and transferred into standards and guidelines (see Sect. 5):

At the beginning of a project, the so-called **Employer's Information Requirements** (EIR), in which the client defines his requirements (level of development, data formats, coordinate systems used and project reference point, necessary preliminary measurements) to potential contractors, e.g., in the tender. During the contract award process, the **BIM Execution Plan** (BEP) is developed and agreed between the client and the contractor. It describes the process for producing the required data and defines all necessary roles, functions, procedures, interfaces, interactions, and the technologies used, e.g., how often and when planning meetings and merging of the aspect models with collision checks take place or which parts of the planning have to be delivered by when and in which level of detail. The **Master Information**

Delivery Plan (MIDP), which summarizes the information to be delivered, is the final step in the allocation process.

With the **Information Delivery Manua**l (IDM), buildingSmart has developed a framework for the uniform description of processes and data transfer points. Part of the IDM are the **Model View Definitions** (MVD), i.e., the definition of the subset of the data model, which must be transferred in particular in which geometrical representation (ISO 29481-1). The international standards of ISO should be transposed into national law by the individual states.

EIR, BEP, and MIDP finally become part of the contractual agreements between the client and the customer. An exemplary representation of these work processes from the UK is shown in Fig. 9.

4.2 Common Data Environment

In addition to defining work processes, **collaborative work** also requires a central data room. The creation or at least provision of data and information should therefore take place in a **Common Data Environment** (CDE) (ISO 19650), in which the various aspect models are also fused. The information contained in the CDE should be understandable by all participants. This requires coordination regarding

- information formats
- delivery formats
- structure of the information model
- the means of structuring and classifying data and
- attribute names for metadata

Ideally, the CDE is not limited to the planning and construction process of buildings with the Project Information Model (PIM), but also extends to the operational phase with the Asset Information Model (AIM), in line with the objective of BIM as a method for the entire life cycle of buildings. Here, the handover between PIM and AIM is of crucial importance, since only with a comparison and, if necessary, adjustment of the planning model (as-planned) to the actually built situation (as-built), a meaningful further use of the model during the operating phase can succeed.

By merging the different aspects models into one complete overall model in the CDE, geometric and technical checks are possible to detect **clashes** between the different partial or aspect models. These checks should be carried out as early as possible and regularly to minimize the costs of necessary rescheduling. Above all, as far as possible, collisions should not be detected only in the construction phase, so that associated conversions are avoided. The transmission of collision information to the actors involved should be standardized to ensure a smooth digital and software-independent exchange. For this purpose, the **BIM Collaboration Format** (BCF) is now being established to communicate model-based issues. In addition to visualization, it contains standardized collection of textual information, a snapshot of the

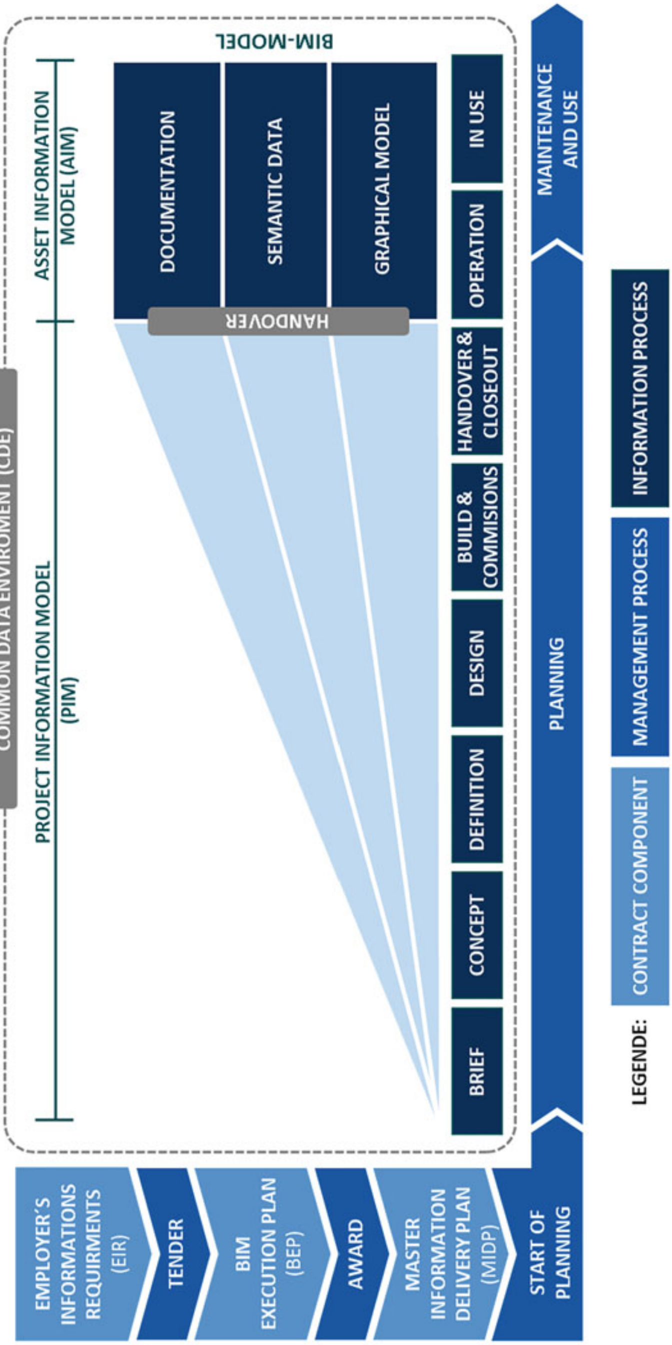


Fig. 9 BIM—The process (adapted from BSI 2013)

topic and the camera position as well as the view direction of the camera (BCF 2020).

4.3 Interoperability and Domain-Specific Data

Planning, construction, and operation of structures require the consideration of additional specialized data and information from other domains. In many cases, this data is not BIM-based or other forms of data storage are established and are still further pursued. With the open modeling and exchange standard of the **Industry Foundation Classes (IFC)** (see also Sect. 5.1), a new standard has been created and is currently being further developed to enable collaborative work with BIM in the sense of so-called "big open" BIM. Nevertheless, even if existing component libraries are included in the standardization process for IFC and continue to be included, exchange formats and component libraries cannot be dispensed with. To achieve the necessary interoperability, syntactic as well as semantic and process-related aspects must be considered.

An example is the integration of geospatial data that is usually managed in **geographic information systems** (GIS): Especially when planning large-scale infrastructure structures, information about the subsoil, the terrain heights and topography or the existing building stock are also required. Syntactically, however, data from BIM and GIS differ, e.g., in the modeling approach and associated data exchange formats, e.g., IFC (see also Sect. 5.1) and **City Geography Markup Language (CityGML)** (CityGML 2020; Becker et al. 2017). Semantically, the linked elements or object classes and their attributes must be assigned to each other or transferred to each other. In terms of processes, different coordinate systems must be considered, e.g., for spatially referenced data, as is often the case when linking BIM and GIS. In this case, the data must be converted or transformed into the other coordinate system (Becker et al. 2020). To ensure interoperability, fundamentally different approaches can be pursued, which Hijazi and Donaubauer (2017) describe using the example of BIM-GIS integration:

- **Model transformation**: Transformation of the domain-specific model (e.g., CityGML) by conversion and mapping to the target system (e.g., IFC) or vice versa. Model transformation represents the most direct path to interoperability and is the most frequently followed in practice. Problematic are different modeling degrees or modeling depths, which lead to loss of information or information gaps.
- Introduction of a unified model: A **superordinate model** contains the models of both worlds, e.g., IFC and CityGML (El-Mekawy et al. 2012). Thus, transformations are avoided, and all information is preserved. The main advantage is that the bi-directional transformation between the two worlds is possible without reduction or loss of information granularity.

- Linking: The data remain in the original data structures and are linked on the application level on a case-by-case basis in the sense of a **multi-model.** The linking can take place on the process level, e.g., using web services, or on the data level, e.g., using semantic web technologies.

Work on linking domain-specific object catalogues with BIM through model transformation of semantics (classes as well as attributes) has already been or is currently being carried out. Interfaces such as the GAEB format or CAFM Connect (CAFM Connect 2020) should be mentioned as examples. Component libraries to be linked are, e.g., the ones for road and traffic of Germany (OKSTRA 2020) or The Netherlands (RWS 2020) (Beetz and Borrmann 2018).

5 Standardization and Norms

Norms and standards are essential in many domains, including the construction industry. While **norms** have a legally binding character, standards and specifications are also developed by voluntarily cooperating committees and established in practice through multiple use.

In the German construction industry, primarily the **Association of German Engineers** (VDI) and the German-speaking group of the non-profit organization **buildingSmart** are driving forward standardization for BIM. If standards are to be transformed into norms, this is the task of the **German Institute for Standardization** (DIN). For international coordination, VDI and DIN have formed mirror committees to the standardization committees of CEN and ISO responsible at EU and international level. Pre-normative standardization in the field of BIM is being promoted internationally in particular by buildingSmart, an organization that operates both internationally and in national groups.

5.1 Industry Foundation Classes (IFC)

In the context of BIM, standards are particularly important for data description, modeling and **data exchange.** As long as all parties involved exchange data solely with purely pictorial representation formats (analog or pdf format) or work with the same software, standards for data exchange are not necessary. In the latter case, transfer in proprietary manufacturer formats is sufficient. With the increasing demand for collaborative work, as required by the BIM method, and the associated exchange of digital vector and semantic data, even between different software, data exchange standards are indispensable. The standards should be as software independent as possible, i.e., open. The open standard IFC (IFC 2020) has now been established for the BIM method. Since the BIM method began in building construction, first IFC models were developed for it and have since been standardized as ISO

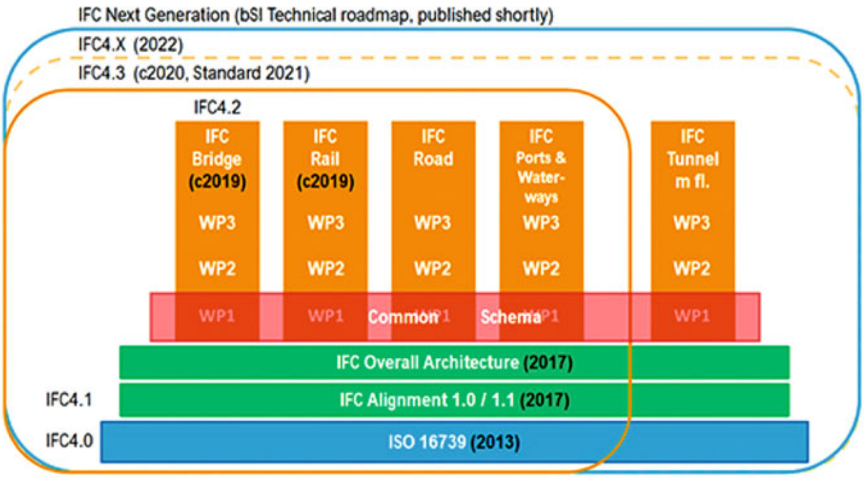


Fig. 10 IFC Infrastructure extensions (Source: https://www.buildingsmart.org/standards/calls-for-participation/ifcroad/)

16379 in IFC 2x3 and IFC4 versions at international level and transposed into national law in many countries. Based on the standards for buildings, common extensions are currently being created for infrastructure types such as bridges, rail and road, ports and waterways, for example, for routing and earthworks. Specific definitions for individual infrastructure types are under development based on these standards (buildingSmart 2020) (Fig. 10). Currently they are published as candidate standard (IFC4.3 RC2) for bridges, rail, road, and ports and waterways.

Standardization also includes a uniform use of technical terms. Only by using them can semantic information in the form of attributes be unambiguously translated into each other. In the field of IFC these definitions are called "**buildingSmart Data Dictionary**" (bsDD) (bsDD 2020).

5.2 *Collaborative Work*

A central aspect of working with BIM is **collaborative cooperation.** This requires coordination and definition of processes of cooperation between the parties involved (see Sect. 4). For this reason, standardized procedures have been and are being developed in this respect as well, which are incorporated into national and international standardization procedures. Internationally, the UK was a pioneer and has introduced national developments into international discussion and standardization through the buildingSmart organization. In the meantime, the resulting **ISO standard** ISO 19650 for organization of information about construction works (information management using BIM) has been transferred via the **EU standard** (EN ISO

19650) into different European law, e.g., the German national standardization as DIN EN ISO 19650. It outlines concepts and principles for the successful management of information at a stage of maturity that can be described as "BIM according to ISO 19650." It is applicable to the entire life cycle of a built asset, including strategic and detailed planning and design, construction, operation, maintenance, refurbishment, repair, and end-of-life (ISO 19650).

For collaborative work, a uniform and as simple as possible approach to data exchange is equally important. The structure for classification of information as well as object-based information exchange should therefore be standardized according to the rules of ISO 12006-2 and ISO 12006-3.

In addition, on national stage guidelines, as well as recommendations for actions and sample documents are produced by working groups on administrative, scientific, or private sector, e.g., the Centre for Digital Built Britain in the UK, the BIM4INFRA2020 working group commissioned by the German Ministry BMVI or The Association of German Engineers (VDI) (VDI directives 2552).

6 Status of Introduction

The process of introducing BIM is being pursued in many countries around the world. Figure 11 gives an overview of the current situation.

In Germany, as in many other nations, a phased plan for the gradual introduction of BIM has been published. The German "Road Map for Digital Design and Construction" (BMVI 2015) has three levels. In the first two stages, first, experiences with the method BIM were to be collected, legal questions were to be examined, recommendations, guidelines, and patterns for assignment and processing were created, and requirements for uniform data structures and database concepts were identified. Since level 3 started in 2020, infrastructure projects of the federal government are to be planned using BIM. The roadmap was continued in a "Masterplan Bauen 4.0" (BMVI 2017) and the establishment of a competence centre BIM Deutschland (BIM Deutschland 2021). It is the declared goal of the current federal government to promote the use of the BIM method in Germany in infrastructure and building construction. Although the German Road Map primarily addresses infrastructure construction, the BIM method is currently being vigorously promoted in building construction in Germany as well.

6.1 International

Scandinavia, the UK, the US, and Australia are particularly advanced (McAuley et al. 2017; NBS 2016b; McGraw Hill 2014). During this development, guidelines have emerged, sometimes already in a second or third version, which make the use of BIM mandatory for construction projects—possibly depending on the construction

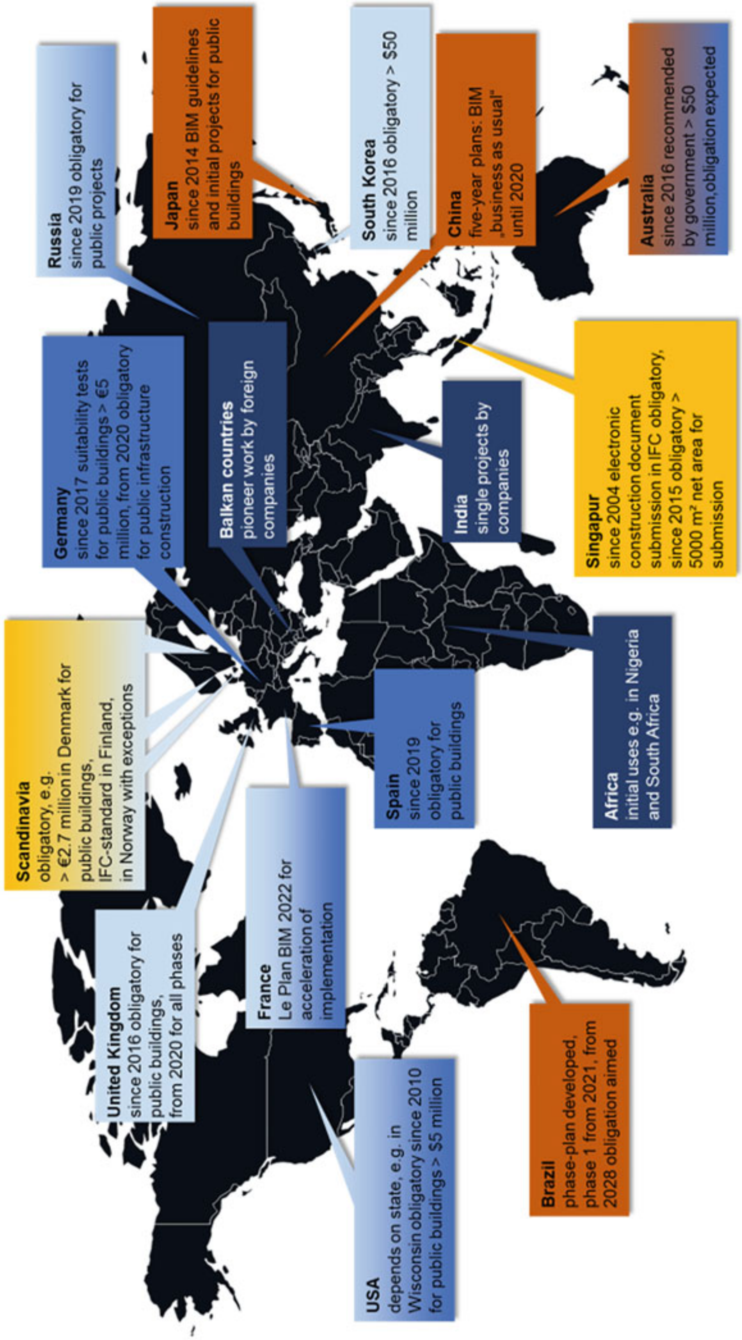


Fig. 11 Exemplary presentation of the international distribution of BIM (status as of 2019)

volume or client (May 2015; Tulke 2017). An example of this is the BS 1192 series of British guidelines, which now covers almost the entire objective of BIM, from the design and construction phase (PAS 1192-2:2013) to asset management (PAS 1192-3), safety-related information (PAS 1192-5:2018), and damage information (PAS 1192-6:2018).

6.2 European Union

On the recommendation of the European Parliament, a directive on public procurement law to promote the "use of alternative electronic means of communication" (by which BIM is meant) in publicly financed construction and infrastructure projects was published in 2014 (EU 2014). At the EU level, the EU BIM Task Group, with the participation of Germany, is working towards a common and coordinated European approach and has published a handbook on the introduction of BIM (EU BIM Task Group 2017).

7 Conclusion

The **digital transformation** is in full swing in many countries around the world. The construction and real estate industries are not exempt from this and even have some catching up to do. An essential part of digitalization in the construction industry is the **BIM** method. It is increasingly becoming an integral part of construction and real estate management processes. It changes the way data is handled (object-oriented modeling) and the type of collaboration (cooperation), because the basis of BIM is the cooperative digital working method. Thus, BIM also brings about new job profiles and roles such as BIM Manager and BIM Coordinator. They take on tasks such as responsible management or coordination of BIM processes and models, definition of requirements in EIR, BEP, etc. (see also Sect. 4) as well as performing quality audits (Egger et al. 2013). The standardization of data models required for collaboration as well as processes is progressing rapidly. The application of the BIM method as a component of digital transformation in AECO industry and real estate management promises to increase efficiency and avoid errors as well as reduce costs.

While the use of BIM is currently mainly aimed at the planning phase, in the future the focus will also be on other life cycle phases, especially the operating phase. For this purpose, the BIM model must become a core part of a so-called "**Digital Twin**" of the real construction. Linked to this is the need for the model to be constantly updated to reflect reality. For this purpose, both classical methods and new technical developments such as the **Internet of Things** (IoT) can be used to transfer information as up-to-date as possible and subsequently draw conclusions and make decisions using the information from the BIM model. New dimensions of use open up, for example, through techniques of **augmented reality**, in which

Fig. 12 Augmented Reality and BIM

models or model information are virtually superimposed on reality in suitable end devices (Fig. 12). Finally, construction processes must be rethought in the context of **integral**, collaborative **planning** in the model and increasing automation.

Acknowledgement We thank for the kind permission to reproduce several figures from other publications (references are given in the figure legends).

References

BCF (2020). https://technical.buildingsmart.org/standards/bcf/

Becker R, Falk V, Hoenen S, Loges S, Stumm S, Blankenbach J, Brell-Cokcan S, Hildebrandt L, Vallée D (2018) BIM – Towards the entire lifecycle. Int J Sustain Dev Plan 13(01):84–95

Becker R, Kaden, R, Blankenbach J (2017): Building Information Modeling (BIM) – neue Perspektiven für Geodäten // Building Information Modeling (BIM) – new perspectives for Geodesists. gis.Business Heft 5/2017 S. 50-57.Wichmann Verlag. ISSN 1869-9286.

Becker R, Clemen C, Wunderlich T (2020) BIM in der Ingenieurvermessung. In: Leitfaden Geodäsie und BIM, Herausgeber: Kaden R; Clemen C, Seuß R, Blankenbach J, Becker R, Eichhorn A, Donaubauer A, Kolbe T, Gruber U, DVW Gesellschaft für Geodäsie, Geoinformation und Landmanagement e.V., Runder Tisch GIS e.V., Version 2.1 2020, DVW – Merkblatt 11-2020, S. 110-119, ISBN 978-3-95786-268-6 (Printversion)

Beetz J, Borrmann A (2018) Benefits and limitations of Linked Dataapproaches for road modeling and data exchange. In: EG-ICE 2018: Advanced Computing Strategies for Engineering. pp 245–261

BIM Deutschland (2021). https://bimdeutschland.de/

BIM4INFRA2020 (2018) Arbeitsgemeinschaft BIM4INFRA2020. https://bim4infra.de/

BIMFORUM (2016) Level of Development Specification, 2016. http://bimforum.org/lod/

BIS (2011) A report for the Government Construction Client Group – Building Information Modelling (BIM) Working Party Strategy Paper. http://www.bimtaskgroup.org/wp-content/uploads/2012/03/BIS-BIM-strategy-Report.pdf

BMUB (2017) Bundesministerium für Umwelt, Naturschutz. Bau und Reaktorsicherheit Runderlass vom 16(01):2017

BMVI (2015) Road Map for Digital Design and Construction, 2015. https://www.bmvi.de/SharedDocs/EN/publications/road-map-for-digital-design-and-construction.pdf?__blob=publicationFile

BMVI (2017) Masterplan Bauen 4.0. https://www.bmvi.de/SharedDocs/DE/Anlage/DG/Digitales/strategiepapier-masterplan-bauen.html
bsDD (2020) buildingSMART Data Dictionary Service, http://bsdd.buildingsmart.org/
BSI (2013) PAS 1192-2:2013. Specification for information management for the capital/delivery phase of construction projects using building information modelling. URL: http://bim-level2.org/en/standards/
buildingSmart (2020) Infrastructure Room https://www.buildingsmart.org/standards/rooms/infrastructure/
Cabinet Office (2011) Government Construction Strategy
CAFM Connect (2020). https://www.cafm-connect.org/
CityGML (2020). https://www.opengeospatial.org/standards/citygml
DB Station&Service AG (2019): BIM–Vorgaben - BIM-Methodik Digitales Planen und Bauen. https://www1.deutschebahn.com/resource/blob/1786332/054e18a2bb1da993bd37f8680c39706c/Vorgaben-zur-Anwendung-der-BIM-Methodik-data.pdf
DIN (2018). https://www.din.de/de/mitwirken/normenausschuesse/nabau/projekte/wdc-proj:din21:259218776
Eastman C, Teicholz P, Sacks R, Liston K (2011) BIM handbook: a guide to building information modeling for owners, managers, designers, engineers and contractors, 2nd edn. John Wiley & Sons, New Jersey
Egger M, Hausknecht K, Liebich T, Przyblo, J (2013) BIM-Leitfaden für Deutschland – Information und Ratgeber – Endbericht, Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) im Bundesamt für Bauwesen und Raumentwicklung (BBR). http://www.bmvi.de/SharedDocs/DE/Anlage/Digitales/bim-leitfaden-deu.html
El-Mekawy M, Östman A, Hijazi I (2012) A unified building model for 3D urban GIS. ISPRS Int J Geo Inf 1(2):120–145
EN 17412-1:2020, Building Information Modelling - Level of Information Need - Part 1: Concepts and principles
EU (2014) Directive 2014/24/EU of the European Parliament and of the Council of 26 February 2014 on public procurement and repealing Directive 2004/18/EC. https://eur-lex.europa.eu/legal-content/DE/ALL/?uri=CELEX:32014L0024
EU BIM Task Group (2017) Handbook for the introduction of Building Information Modelling by the European Public Sector
Frick TW (2017) Industrie 1.0 bis 4.0 – Industrie im Wandel der Zeit, Erstfassung 2. Oktober 2014, Letztes Update 15. August 2017, https://industrie-wegweiser.de/von-industrie-1-0-bis-4-0-industrie-im-wandel-der-zeit/ (Januar 2019)
Goltermann P, Grube T (2017) How to implement the Bispebjerg Hospital as a BIM-5D project. https://www.dreso.ch/en-CH/learning-from-other-countries-bim-project-danish-style/
GSA (2020) General Service Administration. https://www.gov.uk/guidance/creating-a-digital-built-britain-what-you-need-to-know#what-is-building-information-modelling-and-how-does-it-affect-you
Hijazi I, Donaubauer A (2017) Integration of Building and Urban Information Modeling – Opportunities and Integration Approaches. In: Kolbe / Bill / Donaubauer (Hrsg.): Geoinformationssysteme 2017 - Beiträge zur 4. Münchner GI-Runde
IFC (2020). https://technical.buildingsmart.org/standards/ifc/
ISO 12006-2, Building construction — Organization of information about construction works — Part 2: Framework for classification
ISO 12006-3, Building construction — Organization of information about construction works — Part 3: Framework for object-oriented information
ISO 16739-1:2018, Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries -- Part 1: Data schema
ISO 19650-1, Organization of information about construction works — Information management using building information modelling — Part 1: Concepts and principles

ISO 29481-1, Building information models — Information delivery manual — Part 1: Methodology and format
Khemlani L (2005) CORENET e-PlanCheck: Singapore's Automated Code Checking System, AECbytes http://www.aecbytes.com/feature/2005/CORENETePlanCheck.html
May I (2015) BIM in planen-bauen 4.0. In: Scherer RJ, Opitz F, Bauinformatik – Baupraxis 2015: Leitlinien – Richtlinien – Normen – Ergebnisse aus dem VDI-Koordinierungskreis BIM, Tagungsband 6. Fachkonferenz Bauinformatik Baupraxis, Dresden. ISBN 978-3-8678-0458-5
McAuley B, Hore A, West R (2017) BICP Global BIM Study – Lessons for Ireland's BIM Pro-gramme Published by Construction IT Alliance (CitA) Limited, 2017. https://doi.org/10.21427/D7M049
McGraw Hill Construction (2014) The business value of BIM for construction in global markets. McGraw Hill Construction, Bedford
NATSPEC (2013) BIM and LOD, Building Information Modeling and Level of Development. Australien: Construction Information Systems Limited, 2013, ABN 20 117 574 606
NBIM (2007) National BIM Standard-United States® (NBIM-US™): National Building Information Modeling Standard Version 1.0 - Part1: Overview, Principles, and Methodologies. https://buildinginformationmanagement.files.wordpress.com/2011/06/nbimsv1_p1.pdf
NBIM (2015) National BIM Standard-United States® (NBIM-USTM). [online] URL: https://www.nationalbimstandard.org/files/NBIMS-US_FactSheet_2015.pdf
NBS (2016a). https://www.thenbs.com/knowledge/what-is-building-information-modelling-bim
NBS (2016b) National BIM Report. NBS, RIBA, London
OKSTRA® (2020) Objektkatalog für das Straßen- und Verkehrswesen. http://www.okstra.de/
Reif M (2017) Building Information Modeling (BIM) Grundlagen und Vertragsgestaltung in Pilotprojekten des BBR. 4. Oldenburger BIM-Tag BIM-Baumeister-Akademie am 30./31. März 2017
Remes J, Manyika J, Bughin J, Woetzel J, Mischke J, Krishnan M (2018) Solving the productivity puzzle: the role of demand and the promise of digitization. McKinsey Global Institute (MGI). Copyright © McKinsey & Company 2018
RWS (2020) Ontwikkelingen OTL-RWS. https://otl.rws.nl/publicatieomgeving/#/
Singh I (2017) BIM adoption and implementation around the world: Initiatives by major nations. https://www.geospatialworld.net/blogs/bim-adoption-around-the-world/
Smith P (2014) BIM & the 5D Project Cost Manager. Procedia. In: Social and Behavioral Scienc-es 119 (2014), S. 475 484. 27th IPMA World Congress. https://opus.lib.uts.edu.au/bitstream/10453/35808/3/2013004050_PeterSmith_BIM_Journal_Paper-a.pdf
Tulke J (2017) Bauen digital Stand und Entwicklung der BIM Einführung in Deutschland, http://ikth.de/files/2016/Veranstaltungen/biminthringen.pdf
Tulke J (2018) BIM-basierter Bauantrag. http://geoportal-hamburg.de/xleitstelle_webseite_veranstaltung/downloads/06_BIM-basierter_Bauantrag.pdf (07.2019)
USIBD (2016) Level of Accuracy (LOA) Specification Version 2.0. http://www.usibd.org/products/level-accuracy-loa-specification-version-20
Van Nederveen GA, Tolman FP (1992) Modelling multiple views on buildings. Autom Constr 1(3): 215–224. https://doi.org/10.1016/0926-5805(92)90014-B
VDI (2017) Building Information Modeling – VDI-Richtlinie zur Zielerreichung. Verein deutscher Ingenieure (VDI). https://www.vdi.de/vdi-agendabuildinginformationmodeling/

Water Sector 4.0

Martha Wingen and Holger Schüttrumpf

Contents

The Water Sector strives to protect water as an essential component of nature and habitat for flora and fauna, to enable people to use water in future generations and to protect them from the hazards that can arise from water. As part of the public service, the Water Sector operates critical infrastructures with regard to the supply of drinking water and the purification of wastewater. Further tasks are the maintenance of water bodies as well as coastal and flood protection. Challenges such as the progression of climate change, structural and demographic changes, as well as conflicts of use or the shortage of skilled workers are tasks in which digitalization and the "4.0 approach" can provide support for the Water Sector. Many

M. Wingen (✉) · H. Schüttrumpf
Institute of Hydraulic Engineering and Water Resources Management, RWTH Aachen University, Aachen, Germany
e-mail: wingen@iww.rwth-aachen.de; schuettrumpf@iww.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_35

developments have already taken place; however, comprehensive networking of all systems is still far from being implemented. In this chapter, the path of data through the Water Sector is presented using examples from practice and insights from research.

1 What Is the Water Sector?

Water is the most important natural resource for all homo sapiens. They consist of approximately 65% water (Schaal et al. 2016) and should have access to a minimum of 15 liters per day for drinking and domestic hygiene purposes, according to the SPHERE Standards (Sphere Project 2011). However, water use per capita varies heavily throughout the world (IWA 2021). The production of food and consumer goods also requires large quantities of water. Water supply and wastewater disposal is therefore a basic prerequisite for a functioning society, economy and their development. In addition to humans, flora and fauna also depend on water of sufficient quality and quantity, making its sustainable management essential (Umweltbundesamt 2017).

However, the Water Sector includes a much wider range of tasks than just water supply and wastewater disposal. The tasks of the Water Sector furthermore include:

- Determining basic water management data, i.e., observing the hydrologic balance and studying the condition of streams, rivers, lakes, and groundwater
- Regulation of wastewater disposal
- Securing groundwater resources and water supply
- Flood protection
- Water body development and its maintenance

(MULNV NRW 2018)

This diversity of tasks, involving different groups and interests, results in various conflicts of use, that must be considered (Umweltbundesamt 2017). Due to its direct contact with people, (drinking) water supply and wastewater disposal are more prominent in public awareness, than issues such as the maintenance of water bodies or flood protection. In Germany, the quality of drinking water supply and treated wastewater is above average compared to other European countries (Umweltbundesamt 2018). In contrast to other EU countries, almost 100% of wastewater is treated to the highest EU purification standard (ATT et al. 2015). This high quality benefits resource conservation and sustainability. The Water Sector is the central actor in achieving Goal No. 6 of the United Nations Sustainable Development Goals (SDGs), which aims to ensure the availability and sustainable management of water and sanitation for all people by 2030 (United Nations 2015).

The Water Sector is facing a variety of global and regional challenges. Particularly worthy of mention are megatrends such as climate change, the growth or decline of the population temporally as well as spatially, structural changes on the part of the economy and politics, such as the energy transition, as well as technological progress such as digitalization (Berger et al. 2018). To meet these challenges and develop new solutions, the education and training of professionals is essential. One manifestation of climate change that has increased in recent years is extreme flooding and heavy rainfall events (MULNV NRW and MBWSV NRW 2016). In 2018 however, one of the longest dry periods since the beginning of temperature recording followed (DWD 2018), and with it a drop in water levels in the reservoirs, driving bans and restrictions on federal waterways due to low water, and drying out of flora and fauna (Reisinger 2018). Furthermore, there are new problems such as anthropogenic trace substances, multi-resistant germs, nitrate pollution or microplastics. These challenges complicate the development towards a sustainable use of water as a resource—the goal of a modern and integrated water management.

The global Water Sector is organized differently depending on a country's history of organizational structures. In most cases the organization includes a combination of public administration on the one hand and private-sector companies on the other hand. While the public administration at the federal, state and local levels, with its ministries and authorities, regulates, monitors, advises, invites tenders and awards contracts, the private sector takes over a large part of the planning, implementation and maintenance. In some cases, the private sector is also responsible for (drinking) water supply and wastewater disposal. The most common model for water supply, however, is that of semi-autonomous companies founded by the municipalities. This is also referred to as the "city works" model (The World Bank 1995). In contrast, fully municipally owned and managed companies are common, especially for wastewater disposal. The planning and implementation of water management infrastructures and measures are primarily carried out by engineering firms and construction companies. In addition, the necessary technology and machinery are manufactured by highly specialized companies in the sector. The research institutes of the universities and colleges, which work on questions for the entire Water Sector, are located at the interfaces as public-law institutions. Various associations and societies (e.g., IWA, DWA, WWC, etc.) are active in the political and technical representation of global Water Sector issues and interests. They develop guidelines and regulations in technical-scientific working groups, in which the "state of the art" is defined.

(Drinking) water supply and wastewater disposal are defined as critical infrastructures that require special protection.

> Infrastructures are considered 'critical' if they are of important significance for the functioning of modern societies and their failure or impairment results in lasting disruptions to the overall system. (BMI 2009)

Although (drinking) water supply and wastewater disposal are "basic technical infrastructures," water as a resource is also the basis of health care, nutrition, and disaster prevention (BMI 2009). Significant interruptions in water supply and wastewater disposal occur to different extents, depending on the country. However, with the increase in automated control of electronic monitoring of drinking water and wastewater treatment plants worldwide, the risk of cyber-attacks is growing. As critical infrastructure, physical and digital protection of these facilities is a priority for the Water Sector.

2 Automation, Digitalization, and Water 4.0

While automation has already established itself in most subject areas of the Water Sector, digitization continues to spread further. The current developments in automation and digitalization towards Water 4.0 enable a holistic approach to optimizing processes in terms of efficiency, flexibility and speed. The central element is networking and thus the communication of all machines with each other. Networking via the Internet enables coordination between the often-decentralized locations of the Water Sector. For example, in wastewater disposal, sewer networks can be connected and coordinated with pumping station chains, sewage treatment plants and stormwater retention basins (Grün 2016). The particular relevance for the Water Sector is to map the complexity and interconnectedness of water management and hydrological systems in such a way that the scarcity of resources can be counterbalanced by increasing efficiency (Vestner and Keilholz 2016). Through the networking, visualization and virtualization of information, the digitalization of water management enables this holistic mapping of influences on the water cycle and supports sustainable development. However, to realize this in the everyday work of the Water Sector an increase in data density, data collection, and information density is necessary (Kutschera et al. 2016). This allows for better planning of tasks and more stable operations to ensure water supply (Grün 2016).

Institutions and decision-makers in the Water Sector are dealing with topics and terms such as automation, digitalization, the Internet of Things, and Industry 4.0. The German Water Partnership—Working Group Water 4.0 (2015) defined the term Water 4.0 based on the Industry 4.0 initiative, projecting a future fourth stage of industrial development onto the Water Sector. However, like Industry 4.0, this definition is not scientific and does not refer to a specific technology, but describes the vision of a next development phase of increased automation and digitization of the Water Sector (Vestner and Keilholz 2016). Figure 1 shows the parallel development stages of the industrial and Water Sector revolutions. The first revolution for industry and Water Sector is “mechanization.” Starting from the development of the steam engine, the use of steel made it possible to operate plants with high water pressures. An example at this point is the steam engine-driven piston pump for water pumping (Thamsen et al. 2016). The second revolution was “electrification,” which enabled both energy generation (hydroelectric power plants) and utilization (high-speed centrifugal pumps) (Thamsen et al. 2016). The third and acute revolution of “automation” enabled the first programmable controller and the first hydraulic modeling program (FeFlow). While this stage of development is still ongoing, the influence of digitalization is already partially leading to an intertwining of the real world with the virtual world through so-called cyber-physical systems (CPS) (Vestner and Keilholz 2016).

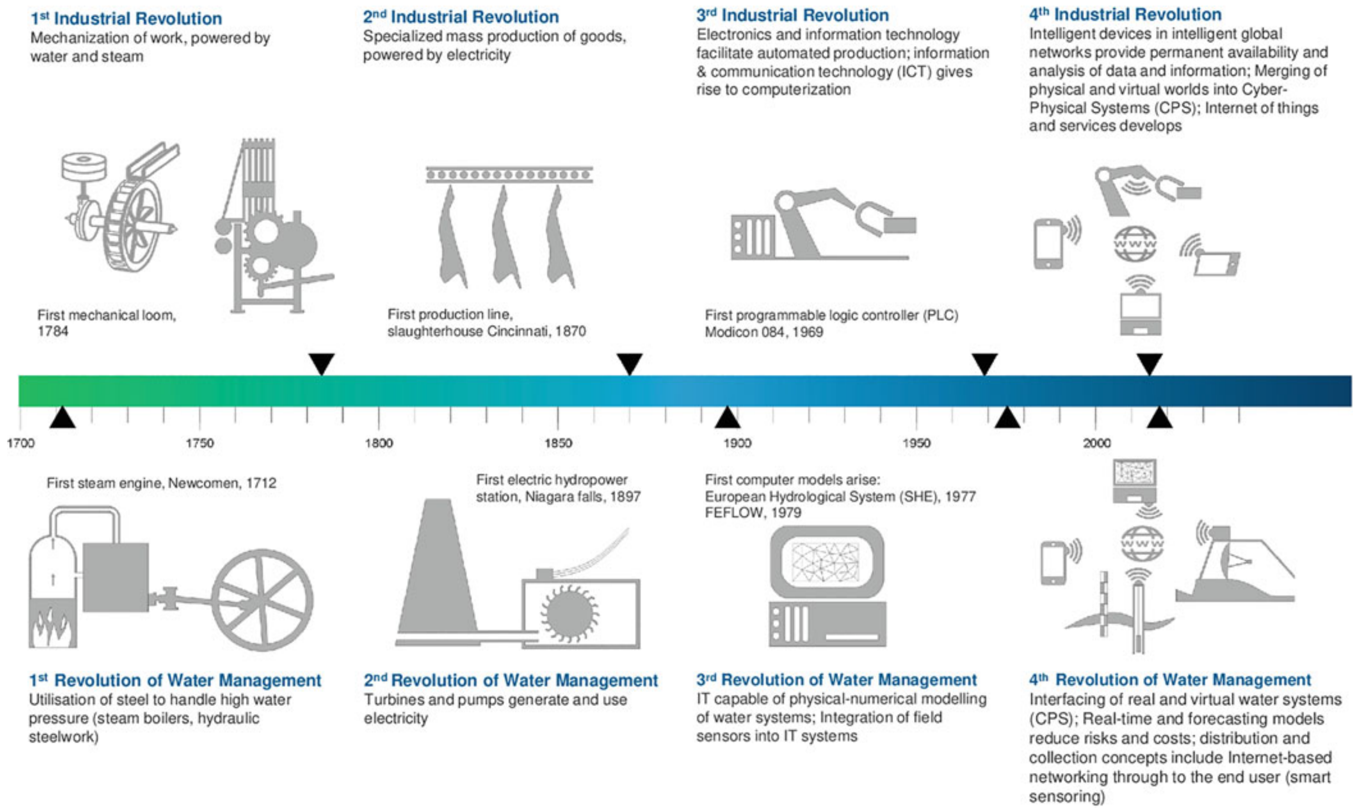


Fig. 1 Comparison of relevant development stages in the industry (in accordance with DFKI (2013)) and the Water Sector (Vestner and Keilholz 2016)

A CPS is a system of

> real (physical) objects and processes linked with information-processing (virtual) objects and processes via open, partly global and at any time interconnected information networks (VDI 2015)

One application in water and wastewater systems today is already the control of very complex networks in large cities such as Berlin with more than a hundred pumping stations and thousands of kilometers of pipe network. What is special here is the possibility of adapting to dynamic demand and changing requirements (Thamsen et al. 2016).

Thamsen et al. (2016) see the pursuit of economic efficiency as the driving force of the Water Sector revolutionary stages. Here, digitalization and Water 4.0 offer the potential, in addition to the secure supply of the population and reduction of costs, also of resource conservation in the form of a reduced energy demand.

3 The Path of Data Through the Water Sector

Data is the central resource of digitization and the 4.0 approach. New technical possibilities and methods enable data generation and collection as well as its processing and visualization (referred to below as processing stages). Modeling and simulation of complex processes play an important role in the 4.0 approach for the Water Sector, enabling a wide range of early warning, monitoring, control and optimization systems. The results are visualized and communicated to decision-makers or the public through digital communication. All processes should take place within a framework characterized by security. The data is available to or processed by various stakeholders in its different processing stages. An overview of all available data in the respective processing stages in a corresponding preparation or visualization for the respective target group is a relevant issue for the Water Sector, especially with regard to transparency. Figure 2 shows the path of data through the Water Sector as a diagram with the respective groups of people involved: The experts on one side and the public and external stakeholders on the other. The individual processing stages are presented in the following chapters.

As in other industries, the quality as well as the spatial and temporal resolution of data in all processing stages is an important aspect. Here, it is necessary to find the right balance between accuracy, required storage volumes and computing speed.

3.1 Data Generation, Collection, and Preparation

To be able to evaluate data, model with it or visualize it, it must first be generated and collected. A distinction can be made between data generated by people (human behavior) and data generated by machines. However, the generation of data does not

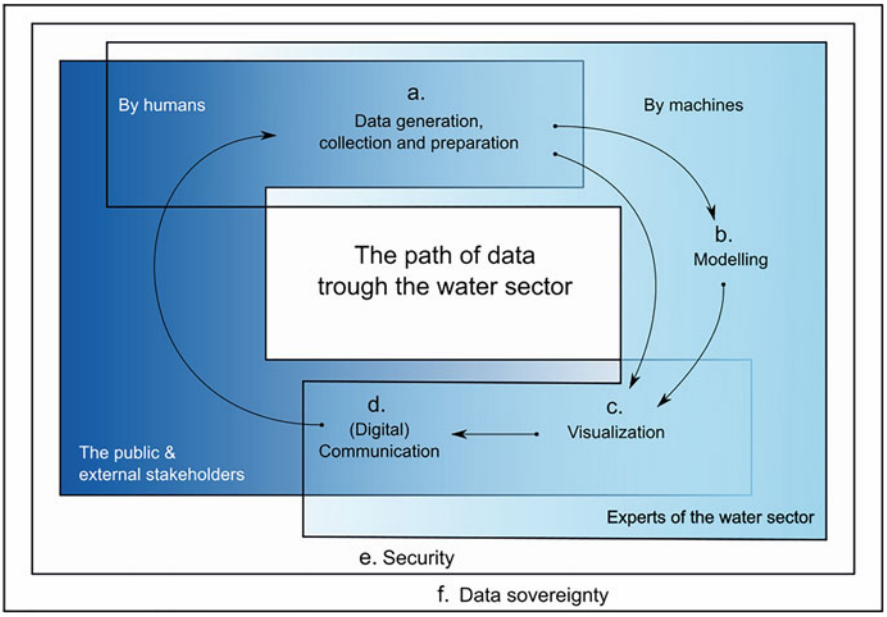


Fig. 2 The path of data through the Water Sector

necessarily include its collection, storage and processing. At the current stage of development, a lot of relevant data for the Water Sector is already generated and collected, however, there is still room for expansion (Vestner and Keilholz 2016). Data that has already been collected from a purely technical perspective includes, for example, precipitation data, water levels at water bodies, flows in canals and pipelines, parameters in basins of wastewater treatment plants, or terrain data obtained through remote sensing. A distinction can be made between continuously collected data and occasion-related data. Continuously collected data is used, for example, in real-time forecasting for flood events or in the control of sewer networks. Occasion-related collected data such as terrain models are used, for example, as the basis for building simulation models. In addition to machine collection, relevant data can also be collected by humans. A classic example of this are censuses, which allow conclusions to be drawn about drinking water consumption or wastewater production. A more recent development is data collected from so-called citizen scientists.

Digital technologies enable a new level of citizen participation in the field of water and nature conservation as well as in the field of environmental education. The term Citizen Science is generally used to describe the process of collaboration between citizens and scientific institutions (Rückert-John et al. 2017). Citizen Science projects are often carried out as monitoring, observation, and crowdsourcing projects and primarily serve to generate new knowledge through the collection, processing, and communication of data. The goal of this collaboration is to complement established science and to involve citizens in science-policy problem definitions and decision-making processes, as well as in the scientific work itself (Rückert-John et al. 2017).

Crowdsourcing projects can collect large amounts of spatially and temporally very complex data on the one hand and save human resources in scientific institutions and companies on the other (Pettibone et al. 2016). Digital devices such as cell phones, tablets, and GPS devices are used for crowd-sourced data collection. Communication of crowd-sourced data often takes place via Internet platforms. These are provided either by the research institutions and companies themselves or by a service provider (Leopold 2015).

When processing data, whether from experts, scientists, citizen scientists, or computers, the goals of further use must be precisely defined so that they are available in sufficient quality and compatible for the further processing stages and work steps of all stakeholders.

3.2 Modeling

Models in the Water Sector represent different situations in reality as detailed as necessary and as simple as possible. Thus, different scenarios from reality can be represented and decisions can be made based on the results. Vestner and Keilholz (2016) summarized different types of modeling used in the Water Sector. To generate realistic results of models, it is state of the art to couple one-dimensional and multidimensional models. Real-time modeling is becoming increasingly relevant in the Water Sector and is used in the planning, design, and operation of water infrastructure. Predictive modeling can be used to represent situations in the future that have a certain probability of occurring. Modeling can be performed using either actual measured data or forecast data as boundary conditions. They serve as a basis for structural measures such as flood protection. To check the prediction quality and accuracy, the model is calibrated and validated on past events.

Figure 3 shows an example of the numerical modeling of a flood event.

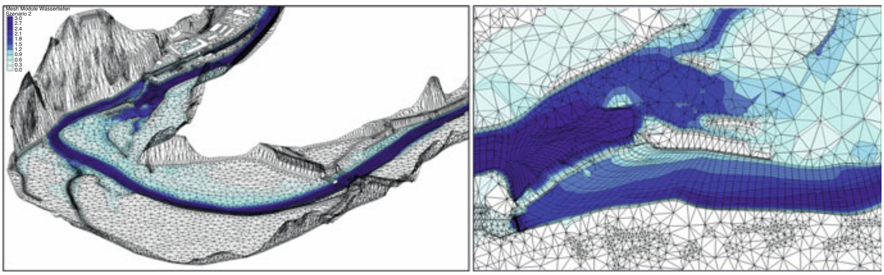


Fig. 3 Numerical model of a flood event

3.3 Visualization

The visualization of data is an important step to present data and model results in a comprehensible, clear, and target group-oriented way. As Fig. 2 shows, the public joins again in this step. Therefore, depending on the target group, it may be necessary to make the data interpretable for laypersons. Especially the public part of the Water Sector fulfills the obligation of transparency towards the public with appropriate visualization.

Visualization is not only used for final results. It also enables an increase in efficiency in the ongoing operation and maintenance of plants and, in particular, in planning. Digital planning, execution and management methods such as Building Information Modeling (BIM) are also used in the Water Sector to visualize future construction projects, to communicate transparently between all stakeholders and to counteract any coordination problems that may arise.

Two kinds of modern technology, augmented reality (AR) and virtual reality (VR), have proven suitable for displaying various data and modeling results. AR involves the computer-based augmentation of reality perception (Marktgraf 2018), while VR describes the representation and simultaneous perception of reality in a real-time computer-generated, interactive, and virtual environment (Bendel 2018). For example, in AR operating and plant data can be displayed on screens through smartphone applications and the device cameras. This facilitates maintenance and, in the event of malfunctions, supports cause identification and rectification, as the intelligent system communicates this directly to the operator at the right location with precise problem identification. AR can also be combined with 3D models and GPS. For example, sewers and pipes can be projected onto the surface at their actual location under the road. Likewise, planned structures can be projected in their original scale to the place where they are planned to be built. Furthermore, simulation and modeling results can be validated in reality. Due to the widespread use of smartphones in society (Koch and Frees 2016), AR technology can also be used in the field of education for sustainable development or in explaining the relevance of precautionary measures against, for example, floods and heavy rain.

For the application of VR, special VR goggles must be procured. For example, they support the maintenance of sewers through virtual inspections (Grün 2018). VR is also frequently used to train employees, as it allows various scenarios to be simulated and practiced.

3.4 Digital Communication

Digitization also has an impact on communication within and outside the Water Sector. The communication of technical topics to the public in particular is in a constant state of flux between communication forms and channels. A core element of digital communication is social media. Many companies and institutions in the

Table 1 Advantages and disadvantages of social media in the Water Sector (Wingen 2018)

Advantages	Disadvantages
Increase of visibility and acceptance	Costs of personnel and resources
Influencing and participating in the public debate on environmental issues	Criticism of the public through participation in public dialogue
Acquisition of skilled workers and trainees	Potential reputational damages
Formation of networks between different institutes	Costly legal violations for non-compliance with laws
Increasing the employees' identification with the institution	Investment in training and development of social media managers
De-bureaucratization and change culture	
Potential for cost savings	

Water Sector are successfully using social media, while others are in the early stages or reject them completely. A study surveyed already established as well as prospective professionals in the Water Sector on the topic of social media and personal usage behavior (Wingen 2018). This revealed a clear difference in the use of and attitude towards social media between different age groups as well as between active, passive as well as non-users. These different groups create a complex initial situation for the acceptance of social media in the Water Sector. The active and passive users formed their knowledge and opinions about social media solely through personal experience, while the non-users formed their opinions without any experience. All interviewees had no training in social media management, but in their respective positions they also had decision-making power about the implementation of social media to be able to use its advantages (see Table 1) and evaluate its disadvantages.

The past has shown that there is a high level of willingness on the part of the civilian population to help during disaster events, such as floods. This help already existed in earlier disaster events, but digital communication enables a new type of networking with the involvement of volunteers. In this context, citizens are already organizing themselves independently via social media, such as Facebook and Twitter. These spontaneous volunteers are not tied to an organization and can take on various activities. This happened, for example, during the 2013 Elbe flood in Germany, where thousands of volunteers organized themselves via the Facebook group "Hochwasser Halle-Saale" (Halle-Saale floods), in which they independently sent out calls for help and provided information about deployment locations (Sackmann et al. 2018). This self-mobilization entails risks but also opportunities that can be reduced or used by the responsible authorities. To benefit from this networking, authorities themselves can be present on social media and directly involve possible spontaneous volunteers in disaster management. Communication with spontaneous volunteers can take place via existing social media or via smartphone applications specifically designed for this purpose (Müller-Tischer 2018).

3.5 Security

Data security at all stages of processing is a major challenge for companies and institutions in the Water Sector. Unlike in other industries, the focus here is also on the security of the critical infrastructures of (drinking) water supply and wastewater treatment and disposal. Since July 2015, the IT Security Act has been in force in Germany, which imposes a special obligation on some of the operators of critical infrastructures to protect their IT systems in accordance with the state of the art, to provide regular evidence of this and to report any malfunctions and security incidents to the Federal Office for Information Security (BSI). This considers not only technical and organizational vulnerabilities, but also human error (BMI 2016). The BSI informs operators about IT security, possible security incidents and vulnerabilities, but also necessitates mandatory reports on corresponding incidents.

3.6 Data Sovereignty

Regardless of the level of data processing, there are different owners, users, processors and responsible parties with different rights, obligations and intentions. Data sovereignty enables those who hold it to bring about the necessary dissemination and thus transparency. Furthermore, they can cause the use of uniform and compatible measuring, simulation and control tools in the course of the different processing stages of the Water Sector's data. It can also support the increase in data density and data collection and information density that will be needed in the future (Kutschera et al. 2016). Data sovereignty is also accompanied by quality assurance.

4 Research and Industry Projects

To further illustrate some of the content from the previous section, selected research projects that reflect current and future developments in the digitization of the Water Sector or Water 4.0 are presented below as examples. Since most of the projects consider all processing stages of the data, projects with specific characteristics for the respective processing stage are selected. Details of the projects are published on the respective websites.

4.1 Data Generation, Collection, and Preparation "Global Flood Monitor"

The Global Flood Monitor is an online tool to detect flood events from social media data instead of traditional early-warning systems based on hydrological models, radar data, gauges and such. The goal is to inform the local flood response organizations as quickly as possible, since the lack of information available causes delays, which cause expensive damages. Researchers at the Institute for Environmental Studies (IVM—VU University Amsterdam) and FloodTags are working on this tool for global, real-time flood detection and monitoring. This new approach was developed using 88 million tweets from which over 10,000 flood events in 176 countries with 11 languages were detected in four years' time. People's need to communicate the appearance of disasters serves as a new and reliable source of data, which was collected and prepared in an online tool (de Bruijn et al. 2019).

4.2 Modeling "Early Dike"

Currently, only water level forecasts are used in early warning systems for flood events. However, other factors such as wind and waves affect the resistance of dikes. With the help of the "Early Dike" project funded by the German Federal Ministry of Education and Research (BMBF), a complex, sensor- and risk-based early warning system is being developed to consider a larger set of parameters critical for dike stability than was previously possible. To this end, smart geotextiles are being integrated into the dike to measure the load on the structure in real time. This will allow for determination of external impacts and the load capacity of the dike to feed software that simulates flooding. All information is collected via a geo-portal and then made available to end users such as the fire department or the technical relief organization in a web-based manner or via app (Krebs et al. 2017).

4.3 Visualization "RiverView"

In the "RiverView" project, which is also funded by the German BMBF, a modular, remote-controlled measuring catamaran is to be used to record visual, hydromorphological, chemical, and physical water data. This data is collected simultaneously and linked with spatial data, so that it's collected for each water body section. Data sets above and below water are collected and linked. Water hydraulics are measured, water monitoring is carried out, the water topography is determined and the surface and surroundings are captured with a 360° camera. This data can be used to create models, which facilitate improvement of operational flood protection and planning processes. The system is also suitable for analyzing the data after a flood event or after pollutant discharges, incidents and accidents. The project also focuses on data visualization. All data can be viewed using VR goggles as part of a virtual waterway experience. In addition, data from several measurement runs can be displayed together on the respective sections to better identify changes and developments (Wöffler et al. 2017).

4.4 Digital Communication "KUBAS"

To illustrate digital communication, the example of spontaneous volunteers during flood events mentioned earlier is brought forward again. For example, the KUBAS project (Koordination ungebundener vor-Ort-Helfer zur Abwendung von Schadenslagen - coordination of unbound on-site volunteers to avert damage) is working on a system in which volunteers can register so that they can be effectively deployed by the authorities in the event of a disaster. The aim of the BMBF-funded project is to create a platform on which volunteers can be coordinated in the event of a disaster. While volunteers currently organize themselves in a decentralized manner, for example via social media, in the future they will be able to register via mobile devices with the help of the KUBAS platform and then be organized centrally from there. The platform is intended to support the crisis team, which can use it to reach volunteers centrally (Rauchecker and Schryen 2016).

4.5 Security "STOP-IT"

Potential physical and cyber threats to the Water Sector have become a problem of general and international concern over the past two decades. Cyber threats include cyber-terrorism, among others. As water utilities continue to invest in digital technology, dependence on digital technology increases. To continue to ensure security, STOP-IT, an alliance of water utilities, industrial technology developers and small and medium-sized high-tech companies in Europe, has been established. The aim is to develop a variety of solutions, including secure communication modules, technologies to protect against physically threatened infrastructures, and implementation of risk assessment and public warning systems (Ugarelli et al. 2018).

4.6 Data Sovereignty "Geoportal Luxembourg"

The national Geoportal of the Grand Duchy of Luxembourg is an Internet platform for the exchange of geodata, maps and related products, services and information. It is operated by Luxembourg's national cadastral and topography administration (ACT) for several years already. It is an example of how an institution with data sovereignty can make data on water and nature conservation available to all interested parties of both professional and private citizens in a comprehensible form. The central element is the map window for the general public, which is free of charge for everyone to use. Moreover, geodata of the cadastral administration itself and many other data offered by the different state authorities and sources are available in the same place. Other comparable institutions in other countries around the world and especially Europe have already developed similar portals, complying with the Infrastructure for Spatial Information in Europe (INSPIrE) Directive (Richardson 2012).

5 Conclusion and Outlook

In summary, it can be said that the Water Sector has already implemented some developments in the direction of Water 4.0 and digitalization and will continue to do so. However, there is still a long way to go before the holistic networking presented here is achieved. The time frame for the implementation of such a project in the Water Sector is characterized by time delays compared to industrial production companies (Wulf and Erbe 2016). Based on the life cycle of the electronics required for Water 4.0, significant changes are forecast for 2025 at the earliest, according to Wulf and Erbe (2016). In addition, not all developments from the field of Industry 4.0 are also relevant for the Water Sector, as it essentially pursues different goals than companies that deal with producing or manufacturing goods.

The 4.0 approach and digitalization offer opportunities to increase efficiency and conserve resources. They support the management of global and local challenges in the areas of water and environmental management. In particular, climate change-driven extreme weather events such as flooding and heavy rainfall can be better managed through real-time forecasting and digital communication in the event of a disaster.

In this context, it is important for the Water Sector to work together on secure and holistic industry standards for the implementation of digitalization and Water 4.0 (Schuster 2018).

References

Arbeitsgemeinschaft Trinkwassertalsperren e.V. (ATT), Bundesverband der Energie- und Wasserwirtschaft e.V. (BDEW), Deutscher Bund der verbandlichen Wasserwirtschaft e.V. (DBVW), Deutscher Verein des Gas- und Wasserfaches e.V. – Technisch-wissenschaftlicher Verein (DVGW), Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V. (DWA), Verband kommunaler Unternehmen e.V. (VKU) (2015) Branchenbild der deutschen Wasserwirtschaft, Bonn

Bendel O (2018) Virtuelle Realität, https://wirtschaftslexikon.gabler.de/definition/virtuelle-realitaet-54243/version-27729, (Accessed 23 Jan 2019)

Berger V, Düwel U, Reuschel B, Merkel W, Schüttrumpf H, Uhl M (2018) Wasserforschung in NRW- ein Überblick, from Energie Wasser-Praxis 10/2018 pp 2–9
Bundesministerium des Innern (BMI) (2009) Nationale Strategie zum Schutz Kritischer Infrastrukturen (KRITIS-Strategie), Berlin
Bundesministerium des Innern (BMI) (2016) Verordnung zur Bestimmung Kritischer Infrastrukturen nach dem BSI-Gesetz (BSI - Kritisverordnung – BSI-KritisV), Bundesgesetzblatt Jahrgang 2016 Teil I Nr. 20, published in Bonn the 2nd of May 2016
de Bruijn JA, de Moel H, Jongman B et al (2019) A global database of historic and real-time flood events based on social media. Sci Data 6:311. https://doi.org/10.1038/s41597-019-0326-9
Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) (2013) Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0 – Final Report
Deutscher Wetterdienst (2018) Deutschlandwetter im Jahr 2018. Offenbach. https://www.dwd.de/DE/presse/pressemitteilungen/DE/2018/20181228_deutschlandwetter_jahr2018_news.html?nn=495078. Accessed 16 Jan 2019
German Water Partnership e.V. – Working Group 4.0 (2015) Wasser 4.0, Berlin
Grün E (2016) Wasserwirtschaft 4.0: die Zukunft gestalten. In: Pinnekamp J (Hrsg) 49. Essener Tagung, Aachen, pp 7/1–7/18
Grün E (2018) Facetten der Digitalisierung: Erste Stationen auf dem Weg zur Wasserwirtschaft 4.0, Mülheimer Tagung, Mülheim
International Water Association (2021) IWA Statistics and Economics: Specific Water Consumption for Households in liters/capita/day, available at: http://www.waterstatistics.org/graph/18 (Accessed 16 Mar 2021)
Krebs V, Herle S, Schwab M, Quadflieg T, Gries T, Blankenbach J, Schüttrumpf H (2017) Implementation of sensor-based dike monitoring by smart geotextiles. SCACR2017, International Short Course on Applied Coastal Research in Santander, Spain from 10. – 06.10.2017
Koch W, Frees B (2016) Dynamische Entwicklung bei mobiler Internetnutzung sowie Audios und Videos. Media Perspektiven 9:418–437
Kutschera G, Barneveld H, Bolle FW, Engels R, Schüttrumpf H, Zethof M (2016) Die Zukunft des Hochwasserschutzes als Teil der Wasserwirtschaft 4.0 - ein erster Blick durch die digitale Brille. In: DWA (Hrsg) Korrespondenz Wasserwirtschaft, pp 674–680
Leopold J (2015) Open innovation und crowdsourcing: Neue Perspektiven des innovations managements. Rainer Hampp Verlag, Augsburg
Marktgraf D (2018) Augmented Reality, https://wirtschaftslexikon.gabler.de/definition/augmented-reality-53628/version-276701, (Accessed 23 Jan 2019)
Ministerium für Bauen, Wohnen, Stadtentwicklung und Verkehr des Landes Nordrhein-Westfalen (MBWSV) Ministerium für Umwelt, Landwirtschaft, Natur- und Verbraucherschutz des Landes Nordrhein-Westfalen (MULNV NRW) (2016) Konzept Starkregen NRW, Düsseldorf
Ministerium für Umwelt, Landwirtschaft, Natur- und Verbraucherschutz des Landes Nordrhein-Westfalen (MULNV NRW) (2018) Lebendige Gewässer in Nordrhein-Westfalen, Düsseldorf
Müller-Tischer (2018) Handbuch Social Media im Einsatz. Grundlagen, Konzepte, Werkzeuge, Amazon Fulfillment.
Pettibone L et al (2016) Citizen Science für alle: Eine Handreichung für Citizen Science Beteiligte, https://www.buergerschaffenwissen.de/sites/default/files/grid/2017/11/20/gewiss_citscifueralle_handreichung_web_0.pdf, (Accessed 8 Jan 2019)
Rauchecker G, Schryen G (2016) Projekt KUBAS: Koordination ungebundener Vor-Ort-Helfer. Im Einsatz 23:252–254
Reisinger K (2018) Eine Dürre, wie es sie noch nie gab. https://www.welt.de/wissenschaft/article181616914/Duerre-in-Deutschland-Was-der-trockene-Sommer-bewirkt-hat.html (2018), (Accessed 16 Jan 2019)
Richardson K (2012) GeoInformatics; Emmeloord Bd. 15, Ausg. 3, pp 18–20
Rückert-John J, Talmon-Gros L, Ziegler D, Bach N, Wiatr M, John R, von Blanckenburg C, Teichler T, Dienel H-L (2017) Konzept zur Anwendbarkeit von Citizen Science in der

Ressortforschung des Umweltbundesamtes. Bundesministerium für Umwelt, Naturschutz, Bau und Reaktorsicherheit (BMUB), Berlin
Sackmann S, Lindner S, Gerstmann S, Betke H (2018) Einbindung ungebundener Helfer in die Bewältigung von Schadensereignissen. In: Reuter C (Hrsg.). Sicherheitskritische Mensch-Computer-Interaktion. Wiesbaden, Springer Fachmedien Wiesbaden, pp 529–549
Schaal S, Kunsch S, Kunsch K, (2016) Der Mensch in Zahlen – Eine Datensammlung in Tabellen mit über 20000 Einzelwerten, Springer Spektrum
Schuster O (2018) Smarte Digitale Transformation in der Wasserwirtschaft. Ein Kompendium der Projektergebnisse, Hof
Sphere Project, Sphere Handbook: Humanitarian Charter and Minimum Standards in Disaster Response, 2011, 2011., Available at: https://www.refworld.org/docid/4ed8ae592.html (Accessed 16 Mar 2021)
Thamsen PU, Mitchell R, Pöhler M (2016) Was versteht man unter Wasser 4.0? In wwt. Wasserwirtschaft Wassertechnik 7–8(2016):16–19
The World Bank (1995) The German water and sewerage sector: how well it works and what this means for the developing countries. International Bank for Reconstruction and Development/ The World Bank, Washington DC
Ugarelli R, Koti J, Bonet E, Makropoulos C, Caubet J, Camarinopoulos S, Bimpas M, Ahmadi M, Zimmermann L, Jaatun MG (2018) STOP-IT - Strategic, Tactical, Operational Protection of water Infrastructure against cyber- physical Threats
Umweltbundesamt (2017) Wasserwirtschaft in Deutschland, Dessau-Roßlau
Umweltbundesamt (2018) Bericht des Bundesministeriums für Gesundheit und des Umweltbundesamtes an die Verbraucherinnen und Verbraucher über die Qualität von Wasser für den menschlichen Gebrauch (Trinkwasser) in Deutschland 2014 – 2016, Dessau-Roßlau
United Nations (2015) Transforming our World: The 2030 Agenda for Substainable Development
Verein Deutscher Ingenieure e.V. (VDI) (2015) Statusreport Industrie 4.0 – Technical Assets: Grundlegende Begriffe, Konzepte, Lebenszyklen und Verwaltung (November 2015)
Vestner RJ, Keilholz P (2016) Was bedeutet der “4.0 - Ansatz” für die Wasserwirtschaft? In Pinnekamp J (Hrsg) 49. Essener Tagung, Aachen, pp 7/1–7/18
Wingen M (2018) Status Quo, Chancen und Risiken von Social Media in der Wasserwirtschaft. Wasser und Abfall; Issue 03; 18. Volume; March 2018
Wöffler T, Engels R, Blankenbach J, Effkemann C, Schüttrumpf H (2017) RiverView – Eine virtuelle Flussfahrt, Internationales Wasserbausymposium Aachen (IWASA), Aachen
Wulf G, Erbe V (2016) Wasserwirtschaft 4.0 - Systematische Weiterentwicklung eines Flussgebietsmanagements. In: Pinnekamp J (Hrsg) 49. Essener Tagung, Aachen, pp 4/1–4/10

Part IV
Electrical and Information Engineering, Mathematics

Artificial Intelligence 4.0

Gerhard Lakemeyer

Contents

1 Introduction

The term Artificial Intelligence (AI) is currently on everyone's minds, especially because of the impressive advances of machine learning (ML) in recent years. This even had the (unfortunate) effect that some people think that AI and ML are one and the same. It is fair to say that we are living in a time of AI hype, since the successes of AI have also created high expectations as to what AI is able to achieve. In fact, influential people such as Elon Musk have warned that AI will develop into a threat for the survival of humanity, with memories of killer robots known from science fiction movies such as the *Terminator* series.

In this chapter, I will first look at AI from a research perspective, which hopefully helps to bring things down to earth as far as the possibilities and limits of AI are concerned. Then I will turn to the chances and challenges which AI offers for Industry 4.0, including some examples from current research.

G. Lakemeyer (✉)
Knowledge-based Systems Group, RWTH Aachen University, Aachen, Germany
e-mail: gerhard@kbsg.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_36

The rest of the chapter is structured as follows. In Sect. 2, I will first try to shed some light on the meaning of AI. Section 3 briefly recalls the ups and downs of AI research and showcases some of the highlights of the past 20 years. In Sect. 4 I will discuss the chances and challenges AI poses for production technology and present several use cases involving machine learning but also other classical AI methods. Section 5 concludes the chapter with further remarks on the challenges for AI in the context of Industry 4.0.

2 What Is Artificial Intelligence?

The term **Artificial Intelligence** goes back to the year 1956 and the by now famous Dartmouth Conference in New Hampshire, USA, where ten scientists met and not only established AI as a scientific discipline but also coined the name. Among those scientists were John McCarthy, Marvin Minsky, Herb Simon, and Allen Newell, who are considered the fathers of AI and who for many decades had a tremendous influence on the direction the field was taking.

But what is AI anyway? Finding a precise answer to this question turns out to be difficult, as one is forced to use analogies from human intelligence, for which we have no clear and precise definition either. It becomes a bit easier when we concentrate on the goals of AI research. In most cases AI research is about building systems, whose behavior we would characterize as intelligent. This is perhaps easiest seen in the case of games such as Chess or Go, where today's programs are much more capable than most human beings. In the case of autonomous robots, which people tend to associate with AI thanks to science fiction movies, researchers are still struggling to equip these machines with goal-directed, rational behavior. Most AI research is actually not concerned with general intelligent behavior. Instead, for the most part, the focus has been on special applications such as translating natural language, detecting objects in images, or just playing games. In other words, by and large, the focus has not been on building artefacts that mimic the thought processes of the human brain, but rather on building systems which behave, when viewed from the outside, in a way that resembles what humans can do or even outperform them, but only for very specific tasks. This distinction is often referred to as **strong** versus **weak AI**. Finally, it is also important to note that AI is different from being **autonomous**. In short, it is perfectly possible to construct machines which act autonomously, that is, without human intervention, but show little or no intelligent behavior.

3 The State of the Art of AI Research

Until the turn of the millennium, AI was mostly an academic discipline, with little notice outside universities or research labs. Beginning in the 50s, the first decades were characterized by alternating phases of euphoria and disillusionment (so-called **AI winters**). When the first game programs were developed in the early 50s, including some which were able to improve by learning, many were convinced that it would only take a decade or so to develop systems which would be at a par with human intelligence or even better. A first disillusionment happened in the late 60s, when it was discovered that research into machine translation and neural networks did not meet the high expectations by far. In the mid-70s, when Edward Shortliffe proposed **MYCIN**, a system which was able to diagnose infectious diseases outperforming medical doctors, the era of **expert systems** was born with much excitement, only to end 15 years later, when it was found that expert system technology often does not scale to larger problem domains and clearly was not a panacea to solve all of AI.[1]

A first media-savvy breakthrough in AI happened in 1997, when IBM developed the chess program ***Deep Blue,*** which was able to beat the then world chess champion Garri Kasparov. While Deep Blue had access to supercomputers, today we find chess programs on ordinary PCs or laptops, and humans have very little chance of winning against them. In this context the recent success of **AlphaGo**, developed by *Google DeepMind,* was truly surprising, when it beat Lee Sedol, one of the best Go players in the world, 4:1 a few years ago (Silver et al. 2016). Most AI experts thought at the time that it would take at least another decade before such performance was possible given the complexity of the game of Go. Today, machines excel even in the game of Poker and can beat the world's best human Poker players (Brown and Sandholm 2019). This is surprising since bluffing, a crucial component of the game, has always been thought to be a very human characteristic.

While the successes of AI concerning games such as chess or Go fascinate lay persons and experts alike, the progress of AI with societal and economic relevance lies elsewhere. One important area is pattern recognition. In particular, object detection on images or videos was revolutionized in recent years through **Deep Learning**, a machine learning methodology based on artificial **neural networks** (LeCun et al. 2017; Krizhevsky et al. 2017). A crucial prerequisite for these breakthroughs was the availability of massive amounts of data in combination with compute power in the guise of GPUs,[2] which make it possible to train large neural networks with millions of data sets. These methods are now also used, with similar success, in machine translation. We already see cases where these systems approach human-level performance.

[1] A concise history of AI can be found in the textbook by Russell and Norvig (2016), which has long served as a reference for teaching AI at universities worldwide.

[2] GPU stands for *Graphics Processing Unit*. As the name suggests, these were originally developed for computer graphics applications and allow highly parallel computations.

After the success of Deep Blue IBM landed another coup with ***Watson***, which in 2011 took part in the quiz show *Jeopardy!* against two human players, who were among the best players, and Watson won decisively (Ferrucci et al. 2010). An interesting aspect of Watson is that, to answer questions, it uses massive amounts of unstructured information available on the Internet. To process this information, it uses many different methods, from natural language processing and machine learning to **knowledge representation** and logical inference. Although Watson was again "only" playing a game, it clearly has the potential for commercial applications in areas where expert knowledge is available mainly in unstructured form. For example, it is envisaged to use such technology to support medical doctors in the diagnosis and treatment of diseases or lawyers searching for cases related to the one at hand. Similar techniques are already being used for speech-controlled **personal assistants** such as Apple's *Siri* or Amazon's *Alexa*.

Ever since the vehicle "Stanley," developed at Stanford University, won the **DARPA Grand Challenge** in 2005 by driving autonomously 200 kilometers through the Mojave desert in California (Thrun 2006), **autonomous driving** has become fashionable. Meanwhile all big car manufacturers are perfecting this technology, an area not only with great market potential but also the potential to drastically influence the way we live. Here, AI methods play an important role as well, in particular machine learning, but also statistical methods for perceiving the environment and **navigation**.

The story behind the development of **robots** is still a bit mixed. Industrial robots have been successfully used in factories for a long time, and they have played an important role in automation and improving productivity. But generally, AI has played little role in developing such systems, as the behavior of industrial robots, such as those found in the car industry, is usually fully deterministic and can be pre-programmed in advance. For autonomous robots, which are supposed to work in human environments such as the home or a machine shop, progress has been much more limited. A major problem lies in the fact that areas occupied by humans or the outdoors are often unstructured and cluttered, which makes it difficult for a robot to orient itself, not to mention perform useful tasks such as setting the table or helping a human to assemble a product. This does not mean that there is no progress in robotics research, to the contrary, but progress has been relatively slow and incremental. Most robots are still largely confined to laboratory environments, with few exceptions such as robot vacuum cleaners or lawn mowers.

When considering the state of the art in AI research as a whole, the picture is rather mixed. On the one hand we see impressive progress in pattern recognition, mainly driven by advances in machine learning in the form of Deep Learning. Other success stories are question-answering systems like a Watson and autonomous driving, which are based on many different AI techniques combined with an amazing amount of engineering. On the other hand, it is not yet well understood how to solve problems where machines need to take rational decisions over a longer period, especially when it comes to unforeseen or rare situations such as driving in a snowstorm. We humans have the ability to deal with such situations using **common sense**. Unfortunately, we do not yet know how to build machines with common sense. Moreover, we are only at the beginning of understanding what it means for a machine to act in an **ethical** way, not to mention having **emotions** or even **consciousness**. In other words, we are far away from developing machines with intelligence that is comparable to human intelligence in a broad sense. On a good note, this also means that we are very far away from creating super-intelligent killer robots like Terminator.

4 AI and Industry 4.0

Despite these caveats, we saw that AI has made significant progress in recent years. Hence it does not seem farfetched when forecasting that AI will play a significant role in the context of **Industry 4.0** and, in particular, **production technology**.

Of course, AI research for production has been around for many years, from applications of classical expert systems (Kusiak and Chen 1988), the use of agent-based techniques (Shen 2002) to machine learning (Monostori et al. 1996), to name but a few. But these results rarely found their way into practice. One reason was that these methods were not mature enough and lacked needed performance. The progress in AI we have seen in recent years, on the other hand, is reason to believe that practical applications will become feasible, even if there are still a number of hurdles in the way, as we will see below.

Machines in production often generate massive amounts of data, so-called **Big Data**, with sensors measuring temperature, pressure, speed, and the like. Up to now this data was rarely used systematically to improve the production process or to make predictions about the quality of a product. Today it seems natural to think that one should use machine learning for this purpose. But doing so is far from trivial. One problem is that the most successful learning methods require **labeled data**. A label could be, for example, an annotation of a data set indicating to what extent the resulting product satisfies given quality criteria. A large number of such annotated data sets could then be used to train a neural network so that it learns which machine **parameters** lead to products of high quality, something currently mainly human machine operators are able to achieve with their expert knowledge and years of experience. While machine learning seems like an obvious choice for automating this process, the crux is that it is often too costly or even impossible to generate enough labeled data. Another issue is that the most successful learning methods like deep neural networks are so-called **blackbox** methods, that is, they are unable to explain to a human why a certain prediction was made. This may be acceptable in certain cases, but to trust the decisions taken by a machine it is desirable, if not necessary, to understand the reasons behind such decisions, especially when mistakes are possible with potentially catastrophic consequences. For this reason, a new

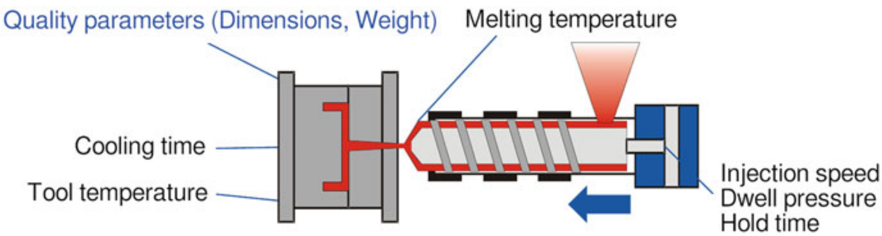


Fig. 1 Schematic depiction of an injection molding machine

research area **Explainable AI (XAI)** was founded a few years ago.[3] While interesting research results on XAI have been obtained, for example Ribeiro et al. (2016), we are still waiting for real breakthroughs.

Despite these hurdles, AI does offer many opportunities for Industry 4.0, making it worthwhile to tackle such challenges. An interesting starting point for investigations is the wealth of **mathematical models** and simulation techniques which have been developed in the context of production technology over many years. Should it not be possible to use such models to improve **data-driven** learning? Likewise, can we improve these models with the help of machine learning? Finding answers to such questions is of central concern for the **Internet of Production (IoP)**, a research cluster of excellence at RWTH Aachen University. In the following, I will present two studies from the areas of plastics and metal forming, which were carried out in preparation for the IoP and which demonstrated that synergies between **model-based** and data-driven views can be achieved. Moreover, one should not forget that classical AI topics such as automated planning can play an important role in the control of production processes. For that we will have a look at an example from production logistics.

4.1 *Machine Learning for Injection Molding*

In a collaborative study between the Knowledge-Based Systems Group (KBSG), the Chair of Informatics in Mechanical Engineering (IMA) and the Institute for Plastics Processing (IKV) at RWTH Aachen University, supervised machine learning using neural networks was applied to the problem of quality prediction for **injection molding** (Hopmann et al. 2017).

Figure 1 schematically depicts an injection molding machine to produce plastic parts. Plastic granulate is inserted into a funnel and then drawn into a cylinder by a spiral while heating it. Molten plastic accumulates at the tip of the spiral and is

[3] As an indication of the importance of this field, the American Defense Advanced Research Project Agency (DARPA) initiated a special funding program (https://www.darpa.mil/program/explainable-artificial-intelligence).

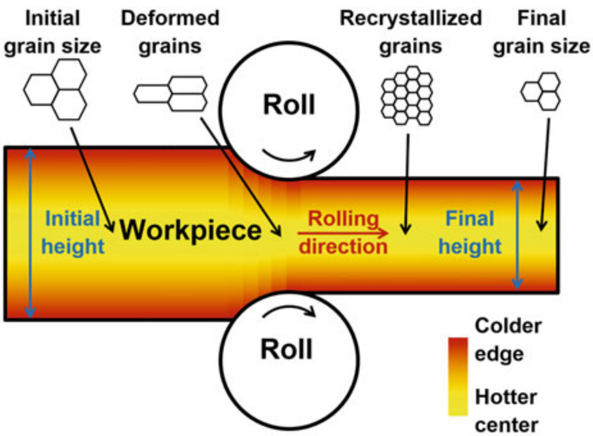


Fig. 2 Principle of reversing hot rolling. (Institute of Metal Forming (IBF), RWTH Aachen University)

pressed into the tool. After a cooling phase the molded part is ejected and the process repeats. Important quality criteria are the weight and dimensions of the molded part. Parameters which influence the quality are temperature, pressure, and injection speed, among other things.

Today these parameters are still often chosen manually by human experts. It would be of great help to have a method which is able to automatically predict the quality of the molded part. Indeed, such methods exist, but they rely on complex **finite-element simulations**, which are very time consuming. The idea behind this study was to automatically learn the prediction of the weight of the part from existing data. A complication was that real data from actual experiments are costly to produce. For that reason, the training of a multi-layered artificial neural network was performed using both real and simulated data. It turned out that this was sufficient to predict the weight of a part for new parameter settings with high accuracy. One of the advantages was that such predictions were possible in real time so that automatic methods for finding parameter settings such as local search were applicable. An important insight of the study was that it is possible to compensate the lack of enough real data by using additional simulation data for supervised learning.

4.2 Machine Learning for Hot Rolling

As another application of machine learning for production technology, we considered the case of reversing hot rolling. This work was conducted in collaboration between the Institute of Metal Forming (IBF) at RWTH Aachen University, IMA and KBSG.

Figure 2 depicts the principle of hot rolling with its most important parameters. Rolling stock, for example, a steel slab, is heated and fed under pressure in several **passes** through two opposing rolls in a reversing fashion. With each pass, the height of the rolling stock is reduced until a desired size and quality is reached. The quality of the work piece in terms of strength and extensibility largely depends on the so-called grain size, which constantly changes during the hot rolling process.

To control the hot rolling process, a **pass schedule**, that is, a sequence of passes, needs to be defined in such a way so that a work piece is transformed into a desired product, for example, a metal sheet of a certain height and quality. Besides information about the material, a pass schedule specifies things such as the height reduction per pass and the waiting time between subsequent passes. The waiting time crucially influences the changes in grain size through recrystallization. In practice pass schedules are mainly designed by human experts, sometimes aided by mathematical models such as ***RoCaT*** (Bambach and Seuren 2015), which, for a given pass schedule, can make very precise predictions about the quality of the end product.

Since, in contrast to the injection molding use case, an excellent model for quality prediction already existed in the form of RoCaT, the goal of this study was not to learn such a model. Instead, **reinforcement learning** was employed to develop a method which was able to automatically generate pass schedules (Meyes et al. 2018). For that purpose, about 300,000 pass schedules with more than 1,000,000 passes in total were generated and evaluated with RoCaT. For a new product new pass schedules were pieced together from these example passes and optimized in an iterative fashion, with desired height and grain size of the end product serving as optimization criteria. It turned out that after 20,000 iterations pass schedules of high quality were indeed found.

This study again demonstrated that learning based on artificially generated data (here: pass schedules) can be successful. An interesting aspect of this work is that the data is evaluated by an existing model, thereby guiding the learning process in important ways. In other words, here we have a case where clear synergies emerged between model-based and data-driven approaches.

4.3 AI Planning Techniques for Production Logistics

It is one of the visions of Industry 4.0 to use autonomous robots in the flexible production of **product variants** with very small lot sizes. Besides the need to solve typical robotic challenges such as navigating safely in semi-structured environments like machine shops or manipulating objects with a gripper, AI techniques are needed so that robots can carry out tasks in a timely manner, either by themselves or as a team. Complications arise because robots usually have only incomplete or uncertain knowledge about their environment, and production should continue even if some of the robots or machines fail.

Before we turn to the challenges of robots in **production logistics**, it is worth noting that robots are already used frequently in **intralogistics**. Here, the robots' task is to carry goods from one location to another in large warehouses, as they are typically found in the online retail business. Here, the robots often travel along fixed and well-marked routes according to pre-defined patterns, that is, largely without the need for AI as far as navigation is concerned. However, research is still needed to coordinate very large fleets of such robots, see for example, Hönig et al. (2016).

To model production logistics scenarios in a realistic fashion and to make it accessible to research, the *Robocup Federation*[4] initiated the **Robocup Logistics League (RCLL)** some years ago in cooperation with the company *Festo* (Niemueller et al. 2016). The idea is to provide a testbed for realistic production logistics scenarios, which allows research groups worldwide not only to develop novel solutions, but to test and evaluate them in competition in a game-like setting. In such a game two teams with three robots each compete against each other. The robots interact with machines, which are randomly placed on a field mimicking a machine shop. Each team has exclusive use of its own machines, but since machines of opposing teams may be next to each other, their robots often need to navigate safely past each other.

The goal in the RCLL is, roughly, to complete as many dynamically generated tasks as possible with the help of robots. Each task consists of an order for a product variant, whose production requires parts and material processed by machines. In addition, a time window is specified, during which a product must be delivered. Depending on the type of product, varying processing steps need to be performed on different machines. The job of the robots is to transport work pieces and sometimes additional material to the respective machines until a finished product can be handed to a delivery station.

The robots are based on the *Robotino* platform built by Festo. Teams are free to add additional sensors such as laser range finders and 2D/3D cameras as well as a gripper. The machines are so-called *Modular Production Systems* (MPS), also built by Festo. Figure 3 shows a robot of the Aachen **Carologistics Team**[5] next to an MPS. The robots can communicate with each other by Wi-Fi. In addition, there is a central server called *Refbox,* which controls the game and handles the communication with machines, among other things.

The products are made up of a cylindrical base, up to three rings and a cap, and all components come in different colors. This way almost 250 different product variants are possible. The main task of the robots is to transport the right pieces or intermediate products to the appropriate machine for further processing. Before a machine can be used, a robot must send a request to the Refbox, which then informs the machine about the planned production step. Sometimes additional steps are necessary. For example, a cap is initially mounted on a base cylinder, which needs to be removed and discarded. Figure 2 schematically shows the production steps for a C3 product consisting of a red base, three rings (yellow, orange, and green) and a gray

[4] At the beginning, the focus of Robocup (www.roborcup.org) was to develop soccer playing robots. In the meantime, many other fields are considered such as rescue or service robots in the home, and more recently, robots in industrial settings such as production logistics.

[5] The Carologistics Team, a collaboration between RWTH Aachen University and The University of Applied Science Aachen, has already won the world title in the RCLL several times.

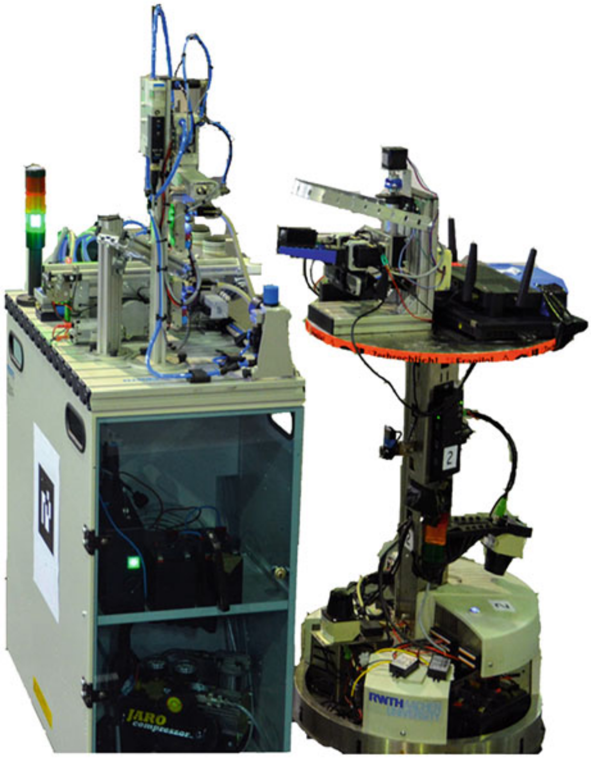


Fig. 3 A robot of the Carologistics Team next to an MPS

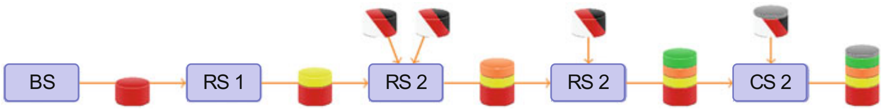


Fig. 4 The production chain for a C3 product

cap. The machines involved in the production are a base station (BS), three ring stations (RS), and a cap station (CS). (The delivery station is not shown.) Overall, the robots have seven different machines at their disposal (Fig. 4).

For navigation, the robots use statistical methods, that is, based on odometry and sensor data the robots estimate their position in a 3D map consisting of x-y coordinates plus an orientation. A by now standard and very robust method for this purpose is **Simultaneous Localization and Mapping (SLAM)** (Thrun et al. 2005). In addition, the robots also need to be equipped with a **collision avoidance** (Fox et al. 1998) system to avoid obstacles, including other robots or human referees moving on the playing field. Another important component is the control software for the gripper, which is needed to hand over work pieces to a machine or to pick them up after processing. Since each team has its own gripper design, the control software varies considerably and is custom made for each individual gripper design.

In most cases, a 3D camera such as the *Kinect* is used for perception and object detection.

While the methods mentioned in the previous paragraph address typical robotics problems, for which a large number of approaches and solutions already exist, the high-level processing of orders in production logistics scenarios poses a number of interesting AI challenges and research questions, for which no satisfactory solutions exist yet. On the surface it may seem that **AI planning** techniques, which have matured considerably in the past 20 years (Bonet u. Geffner 2013),[6] should provide ready-to-use solutions, but this is not the case. The focus of research in AI planning has been to find a plan, usually consisting of a sequence of actions, for a given goal and initial state description, or to show that no plan exists. While many variations of the planning problem have been investigated, such as incomplete state description and nondeterministic or stochastic action effects, the view was static in the sense that all information necessary to find a solution was given at the beginning. In the RCLL, however, this is not the case. For one, new orders (goals) are dynamically generated while existing orders are being processed, that is, it is not known initially how many and what type of orders will arrive during a game. In particular, this may result in situations where a new order arrives and the robots decide that it is better to abort an order which they are currently working on. To tackle problems of this sort, the new research direction of **Goal Reasoning** was recently founded (Roberts et al. 2016). For another, many problems may arise during execution time, making it impossible to successfully complete an order which is currently being processed. In the worst case, this may be due to catastrophic failures of both robots and machines. Consequently, planning and execution cannot be strictly separated from each other, and **execution monitoring**, whose role is to detect unforeseen problems, is of critical importance (Brenner u. Nebel 2009; Hofmann et al. 2016). Nevertheless, AI planning techniques do play a role in figuring out how to satisfy existing orders, but it is just not the whole story. Some of the approaches can be found in Hofmann et al. (2017), Schäpers et al. (2018), and Leofante et al. (2017). To allow researcher in AI planning to address all these problems without having to become experts in robotics, a simulation environment was developed for the RCLL (Zwilling et al. 2014). The simulator provides software for navigation, object manipulation, and sensor data interpretation so that one can focus on planning and execution monitoring without having to worry about the many other facets of robot control.

One aspect, which is not yet considered in the RCLL but which is important for actual production logistics applications, is the interaction with humans. It seems realistic that part of a production chain still needs to be carried out by humans or in cooperation between humans and robots. This is especially true when it comes to assembly tasks that require a high degree of dexterity, where humans are still

[6]The *Int. Conf. on Automated Planning and Scheduling (ICAPS)* not only showcases the latest research results in AI planning but also regularly holds competitions, where state-of-the-art planning systems compete against each other using benchmarks from a wide range of domains.

superior to robots. Nevertheless, the RCLL already poses many research questions, whose solution will take us closer to realizing the vision of Industry 4.0.

5 Conclusions

The examples we considered in the previous section regarding the use of machine learning in production or robots in production logistics are so far little more than case studies. Nevertheless, they already demonstrate, if only in a very selective way, the usefulness of AI for production technology and Industry 4.0 in general. However, much more research is needed to uncover its true potential and then turn this potential into practice. As already mentioned, some of these research questions are currently being address in the excellence cluster "Internet of Production" at RWTH Aachen University. An important theme is to integrate model-based and data-driven methods. Central questions are whether it is possible to improve existing models by machine learning and, vice versa, whether existing models can be used to optimize machine learning methods. The examples from injection molding and metal forming discussed in the previous section can be thought of as first attempts to find answers for these questions.

Moreover, the question of how to explain decisions taken by learning methods will also need to be addressed. Finally, it should be noted that the AI techniques mentioned in this chapter are, of course, not the only ones relevant for Industry 4.0. For example, in the context of human-machine interaction, recent advances in natural language processing[7] seem readily applicable in production technology. And many other areas of AI will undoubtedly follow.

References

Bambach M, Seuren S (2015) On instabilities of force and grain size predictions in the simulation of multi-pass hot rolling processes. J Mater Process Technol 216, Elsevier, pp 95–113
Bonet B, Geffner H (2013) A concise introduction to models and methods for automated planning. Synthesis lectures on artificial intelligence and machine learning. Morgan Claypool Publishers
Brenner M, Nebel B (2009) Continual planning and acting in dynamic multiagent environments. Autonomous Agents Multi-Agent Syst 19(3):297–331
Brown N, Sandholm T (2019) Superhuman AI for multiplayer poker. Science 365(6456):885–890
Ferrucci DA, Brown EW, Chu-Carroll J, Fan J, Gondek D, Kalyanpur A, Lally A, Murdock JW, Nyberg E, Prager JM, Schlaefer N, Welty CA (2010) Building Watson: an overview of the DeepQA Project. AI Magazine 31(3):59–79
Fox D, Burgard W, Thrun S, Cremers AB (1998) A hybrid collision avoidance method for mobile Robots. In: IEEE International Conference on Robotics and Automation (ICRA), pp 1238–1243

[7] An example is the technology behind voice-controlled personal assistants, which have become part of our daily lives already.

Hofmann T, Niemueller T, Claßen J, Lakemeyer G (2016) Continual Planning in Golog. In: Thirtieth AAAI Conference on Artificial Intelligence, pp 3346–3353
Hofmann T, Niemueller T, Lakemeyer G (2017) Initial results on generating macro actions from a plan database for planning on autonomous mobile Robots. Int Conf Automat Plan Sched:498–503
Hönig W, Satish Kumar TK, Cohen L, Ma H, Xu H, Ayanian N, Koenig S (2016) Multi-agent path finding with kinematic constraints. In: International Conference on Automated Planning and Scheduling (ICAPS), pp 477–485
Hopmann C, Jescke S, Meisen T, Thiele TD, Tercan H, Liebenberg M, Heinisch J, Theunissen M (2017) Combined learning processes for injection molding based on simulation and experimental data. In: 33rd International Conference of the Polymer Processing Society (PPS-33)
Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. Commun ACM 60(6):84–90
Kusiak A, Chen M (1988) Invited review: expert systems for planning and scheduling manufacturing systems. Eur J Operat Res 34(2):113–130
LeCun Y, Bengio Y, Hinton G (2017) Deep learning. Nature 521:436–444
Leofante F, Ábrahám E, Niemueller T, Lakemeyer G, Tacchella A (2017) On the synthesis of guaranteed-quality plans for Robot Fleets in logistics scenarios via optimization modulo theories. In: IEEE International Conference on Information Reuse and Integration, pp 403–410
Meyes R, Tewrcan H, Thiele T, Krämer A, Heinisch J, Liebenberg M, Hirt G, Hopmann C, Lakemeyer G, Meisen T, Jeschke S (2018) Interdisciplinary data driven production process analysis for the internet of production. In: 46th SME North American Manufacturing Research Conf. (NAMRC 46), Elsevier
Monostori I, Markus A, Van Brussel H, Westkämpfer E (1996) Machine learning approaches to manufacturing. CIRP Ann Manufact Technol 45(2):675–712
Niemueller T, Ewert D, Reuter S, Ferrein A, Jeschke S, Lakemeyer G (2016) Robocup logistics league sponsored by festo: a competitive factory automation testbed. In: Jeschke S, Isenhardt I, Hees F, Henning K (eds) Automation, communication and cybernetics in science and engineering 2015/2016. Springer, Cham, pp 605–618
Ribeiro MT, Singh S, Guestrin C (2016) Why should I Trust You? Explaining the Predictions of any Classifier. In: Proceedings of the 22nd ACM SIGKDD, pp 1135–1144
Roberts M, Lford R, Shivashankar V, Leece M, Gupta S, Aha DW (2016) ACTORSIM: A toolkit for studying goal reasoning, planning, and acting. In: Workshop on Planning and Robotics (PlanRob) at International Conference on Automated Planning and Scheduling (ICAPS)
Russell S, Norvig P (2016) Artificial intelligence: a modern approach, Global edn. Pearson
Schäpers B, Niemueller T, Lakemeyer G, Gebser M, Schaub T (2018) ASP-based time-bounded planning for logistics robots. In: International Conference on Automated Planning and Scheduling (ICAPS), pp 509–517
Shen W (2002) Distributed manufacturing scheduling using intelligent agents. IEEE Intell Syst 17(1):88–94
Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap TP, Leach M, Kavukcuoglu K, Graepel T, Hassabis D (2016) Mastering the game of Go with deep neural networks and tree search. Nature 529(7587): 484–489
Thrun S (2006) A personal account of the development of Stanley, the Robot that Won the DARPA grand challenge. AI Magazine 27(4):69–82
Thrun S, Burgard W, Fox D (2005) Probabilistic Robotics. MIT Press
Zwilling F, Niemueller T, Lakemeyer G (2014) Simulation for the RoboCup Logistics League with real-world environment agency and multi-level abstraction. RoboCup Symposium, Springer LNCS, pp 220–232

Life Cycle-Spanning Experimentable Digital Twins

Jürgen Roßmann and Michael Schluse

Contents

J. Roßmann · M. Schluse (✉)
Institute for Man-Machine Interaction, RWTH Aachen University, Aachen, Germany
e-mail: rossmann@mmi.rwth-aachen.de; schluse@mmi.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_37

1 Introduction

Digital Twins (DT) revolutionize our view on systems—and this from completely different perspectives. During **development**, the twin type is developed and tested via DT instances, reducing the need for early prototypes of the corresponding Real Twin. In the context of **networking**, e.g., on the Internet of Things (IoT), DTs are networked and thus indirectly the corresponding real machines. In the context of **user interfaces** needed to supervise and command a machine, the worker interacts with the DT, which then performs appropriate actions on the real machine. In the context of the development of Artificial Intelligence (AI) algorithms, DTs generate the necessary **training data** in the virtual world in a short time without endangering people or machines. The same applies to the **training** of drivers or operators, who can safely learn how to operate a machine via the DT. In **Semantic World Modeling** (Sondermann 2018), DTs that represent objects in the environment are created as a basis for subsequent planning activities. Further examples of the use of DT can be found in areas such as predictive maintenance, model-based systems engineering, model-based control, or virtual commissioning (VDI 2016).

It turns out that the DT is both, the result of digitalization (in the sense of transferring real-world artefacts into a digital representation) and the basis of digitalization (in the sense of new approaches using these digital representations, such as networking, automation, big data, AI, new business models, etc.). Despite the wide range of applications, the main strength of the DT concept is its simple and descriptive basic idea, which is the basis for all above-mentioned applications. On the other hand, different perspectives on DTs have evolved, which makes discussions about and development of DT-based applications sometimes quite complicated and, at the end, also hinders to leverage the full potential of this powerful concept. That is why the starting point of this chapter is the widely accepted concept of Cyber-Physical Systems (CPSs) as outlined in Fig. 2 consisting of a Physical System and a cyber part, the Information-Processing System (IPS). Some people now use the term "DT" to refer to a (highly sophisticated) simulation model of the Physical System which has been initialized and calibrated with real-world data received from the real Physical System, others talk about the information-processing part of the CPS, the IPS. In addition to this, the DT is excellent for marketing purposes and is often used unspecifically in this context. Beyond that, the situation is not improved by the fact that the term "Digital Twin" is defined only very vaguely, if at all.

This chapter aims at introducing a common conceptual and technical basis for the different perspectives on DTs. This allows to integrate the different perspectives on DT into an overarching concept and thus enables the life cycle-spanning use of DTs from initial design through commissioning to operation and beyond. A central role is assigned to simulation technology, which makes DTs executable and experimentable—Digital Twins become Experimental Digital Twins (EDT). In the resulting **EDT-Methodology**, a consistent set of tools, methods, and processes for dealing with DTs and their different perspectives and various applications is defined.

The rest of the chapter is structured as follows. In Sect. 2, the terms DT and EDT are defined, and the different facets of an EDT are investigated. Section 3 divides the use of EDTs into different methods, based on the application-specific evaluation of concrete application scenarios, which are modeled as networks of interacting EDTs. By using the same EDT in different applications in different phases, the EDT structures the life cycle of CPSs as shown in Sect. 4. Section 5 then sketches examples for concrete EDT-Applications from different application areas.

2 The Experimentable Digital Twin

Starting with Cyber-Physical Systems, this section defines the notion of EDT in delineation and concretization of existing views on DTs and as a basis for the technical implementation of DTs.

2.1 Cyber-Physical Systems

A Cyber-Physical System links **Physical Systems** with *Information-Processing Systems*. Concrete examples of Physical Systems are machines, vehicles, buildings, houses, or workers (see Fig. 1). From the Industry 4.0 perspective, the Physical System can also be called an **Asset**, an "entity that has a perceived or actual value to an organization and is owned or managed individually by the organization" (Plattform I4.0, 2020). A distinction is made between material Assets, which are the focus of this chapter, and immaterial Assets.

Modern IPSs add smart analytics, information processing and networking capabilities to Physical Systems and provide system models, cognition and decision making (see e.g., Lee et al. 2015). The IPS is created during development and continuously updated with real-time data during operation (see Fig. 2). Following the CPS approach, each IPS corresponds to a specific Physical System. From an information technology perspective, the IPS communicates at the same time with its Physical System and with its environment. This way, the IPS becomes a mediator both between Physical Systems and in particular between humans and Physical Systems (Cichon u. Roßmann 2018).

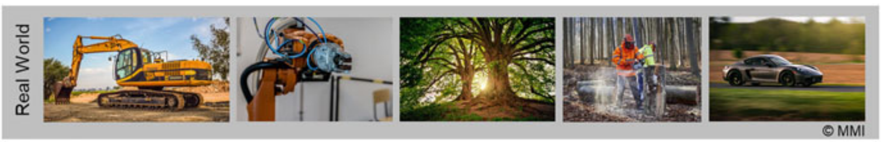


Fig. 1 Examples for Physical Systems

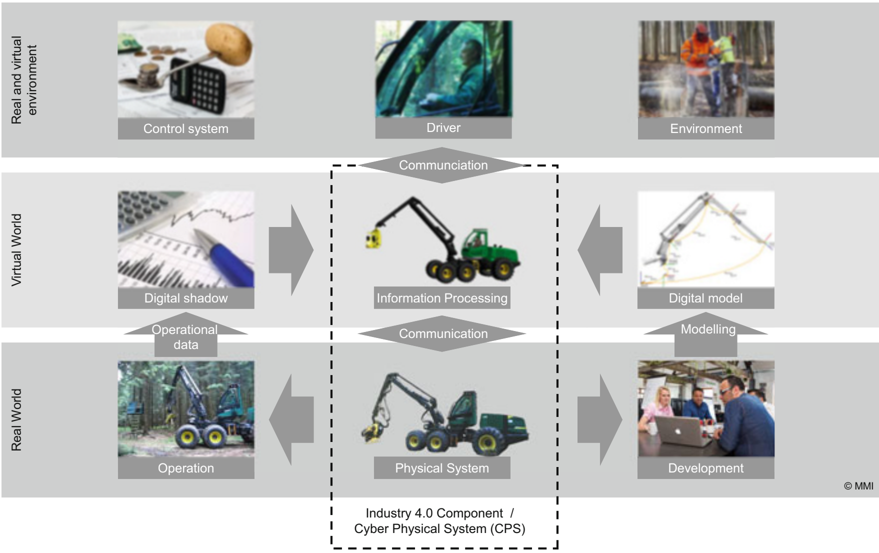


Fig. 2 Cyber-Physical Systems in development and operation

2.2 Interaction of Cyber-Physical Systems

Looking at the interaction of Cyber-Physical Systems, three types can be distinguished. Physical Systems interact with each other on the **physical level**. An example is the interaction between a machine and the workpiece. This interaction leads to energy and material flows corresponding to different disciplines such as mechanics, electrical engineering, hydraulics, or thermodynamics. In the context of **internal communication**, the IPS exchanges data and control instructions with the machine itself (its Physical System) via individual, device- or software-specific IT interfaces (in the forest machine example, for example, via the CAN-bus). With respect to **external communication**, IPSs "talk" to each other via suitable Industry 4.0 communication protocols, e.g., OPC UA (OPC 2017). The interaction possibilities of Physical Systems and IPSs on the physical and IT levels are described in abstract form by interaction points, so-called ports, which in their entirety form the **interface** of a CPS.

2.3 The Digital Twin

Colloquially speaking, a Digital Twin is the virtual equivalent of a Real Twin. But what exactly are a **Digital Twin** and a **Real Twin**? The meaning of the word combination "Digital Twin" is not clearly defined. (Dudenverlag 2019) defines as "twin," among other things, one "of two children born at the same time." In addition, other word combinations such as "twin type," "twin tower," or "twin tire" are used. The term "twin" denotes "the same thing twice," at least according to currently relevant features. The same applies to the DT. Here, too, the same thing is considered twice, namely, the twin once in the real world, the Real Twin, and once in the virtual world, the Digital Twin. Here, the DT denotes the digital half of the twins (or of this pair) and thus provides a descriptive term for the Digital Artefacts created during the digitalization of real-world artifacts.

The use of the DT term in relation to technical systems goes back to Michael Grieves, who mentions the term in 2003 in the context of Product Life Cycle Management (PLM, Grieves 2014). In 2010, NASA revisited the concept for aerospace and used it to refer to "ultra-realistic simulations" (Shafto et al. 2010). Subsequently, the term has been examined from various other perspectives, e.g., from the perspectives of CPSs (e.g., Lee et al. 2015), simulation for product and production engineering (e.g., Boschert and Rosen 2016), IoT (e.g., Bosch 2020), or (model-based) systems engineering (Madni et al. 2019). Grieves himself summarizes diverse use cases in (Grieves 2014). In 2018, Gartner classifies DTs as part of the digitalized ecosystem as one of the five key technological trends of the next decade and predicts the technology to reach the "productivity plateau" in 5–10 years (Panetta 2018). Nevertheless, it is still not finally defined what a DT is or where is the DT focus. The opinions range from simulation models to different combinations of real objects, their physical equivalent, connections between them, services offered by them, and operational data. The lack of definition of this term can be seen in (Plattform I4.0, 2020), which denotes the DT on the one hand a "virtual digital representation of physical assets" and on the other hand as a "simulation model" (see right-hand side of Fig. 3).

The Real Twin denotes the "System of Interest" described by its Digital Twin for the purpose of a special use case. With respect to CPS, the Real Twin can be the Physical System, the IPS or the CPS itself (see Fig. 3). With respect to this Real Twin, the Digital Twin integrates all relevant Digital Artefacts. A Digital Artifact is

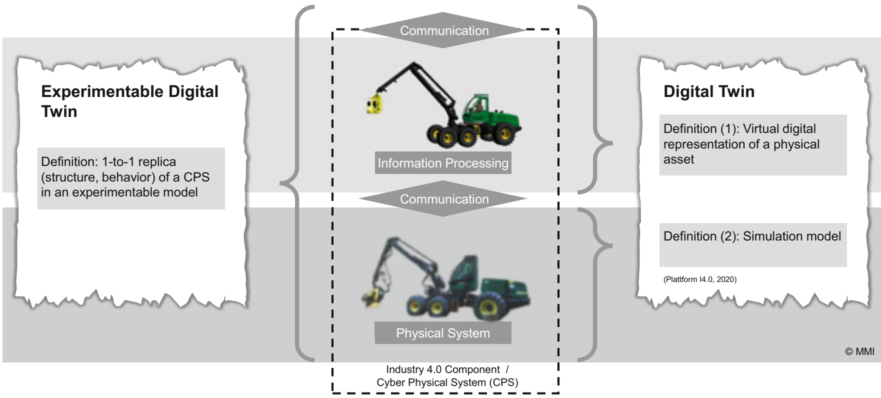


Fig. 3 Towards the definition of Digital Twins (extract from Fig. 2)

an artifact whose design and implementation is accomplished through information technology means (Wimmer 2009). A Digital Artifact can be embedded in a (physical) artifact making up a product. Typical Digital Artefacts in the DT context are static metadata, digital models (geometry, structure, . . .) and data generated during the operation of a Physical System (sometimes called "digital shadow"), functions/services and interaction capabilities provided by the CPS (e.g., interfaces for external communication to others, . . .) as well as simulations used to predict its behavior. Therefore, the following definition is used, which includes the above-mentioned aspects:

Definition *A* ***Digital Twin*** *is a virtual digital 1-to-1 representation of its* ***Real Twin sufficient to meet the requirements of a set of use cases*** *which allows interaction based on up-to-date* ***Digital Artifacts****. Digital Twins can be arranged and connected in hierarchical and heterarchical structures just like their Real Twins.*

2.4 The "Realization Space" of Digital Twins

Starting from the very general definitions of CPS, Digital and Real Twin and their interaction, completely different aspects regarding realization and use of DTs can be developed. Figure 4 provides an example of individual dimensions and related elements. Figuratively speaking, these dimensions span the "realization space" of a DT. As a developer or user of a DT ("user views" dimension), e.g., a businessman has a different view on a machine and considers different interactions than a mechanical engineer ("discipline" dimension). The latter's perspective is in turn different from that of a computer scientist, etc. In simulation technology, a rigid-body dynamics model of the machine looks completely different from a discrete-event model interested in work- and downtimes (dimension "simulation domain"). The DT may be found on the desktop computer of an engineer, in a cloud infrastructure or in an edge approach on a machine ("infrastructure" dimension). The same DT may be used in different applications ("application" dimension). Its life cycle starts with the twin type – a model description of the Digital/Real Twin—which is instantiated by Digital and Real Twins ("life cycle 1" dimension). It must be the goal of a DT to provide a comprehensive "coverage" of this realization space, if possible, using one single but comprehensive DT as "single source of truth."

2.5 The Digital Twin as Node of the Internet of Things

An important perspective on the DT especially in the context of Industry 4.0 is the IT-related networking of Assets. For this purpose, an **Asset Administration Shell** (AAS, called IPS in Fig. 2) is added to the Asset (the Physical System in Fig. 2). The

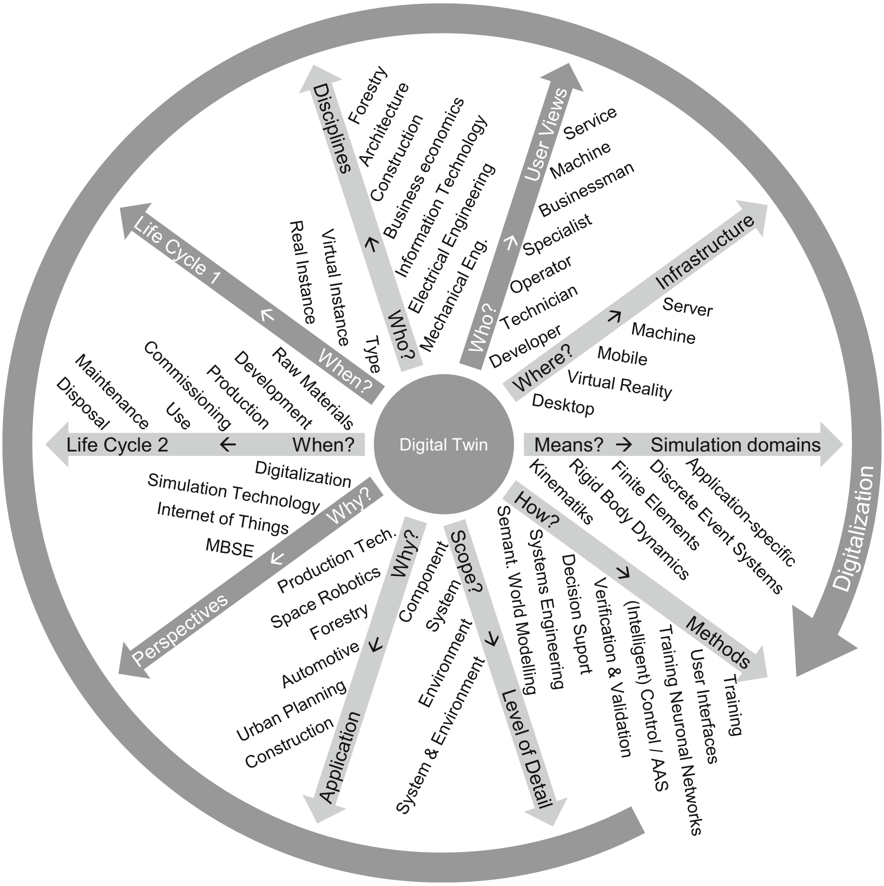


Fig. 4 Selected dimensions of the "realization space" of DTs

AAS is the virtual, digital, and active representation of an **Industry 4.0 component** (I4.0 component, consisting of Asset and AAS, see Fig. 2) in an **Industry 4.0 system** (I4.0 system, Plattform I4.0, 2020). In the technical implementation, the AAS is the run-time environment of a DT (see also Sect. 2.7). This distinction is often neglected and instead, as in the following, the AAS itself is referred to synonymously as the DT (Plattform I4.0, 2020). From the perspective of the DT, a CPS or an I4.0 component can also be described as a Hybrid Twin (see Fig. 15), as it combines parts of the Digital (IPS/AAS) and Real (Physical System/Asset) Twin.

An I4.0 system fulfills its task through appropriate interaction between the individual I4.0 components. Figure 5 outlines this using the example from Fig. 2, the forestry machine application. Here, the forestry machine, the superordinated control system, and the forest worker (via his chainsaw) are networked with each other. The Assets interact with their DTs via internal interfaces (vertical block arrows), and the DTs communicate with each other via suitable IoT communication

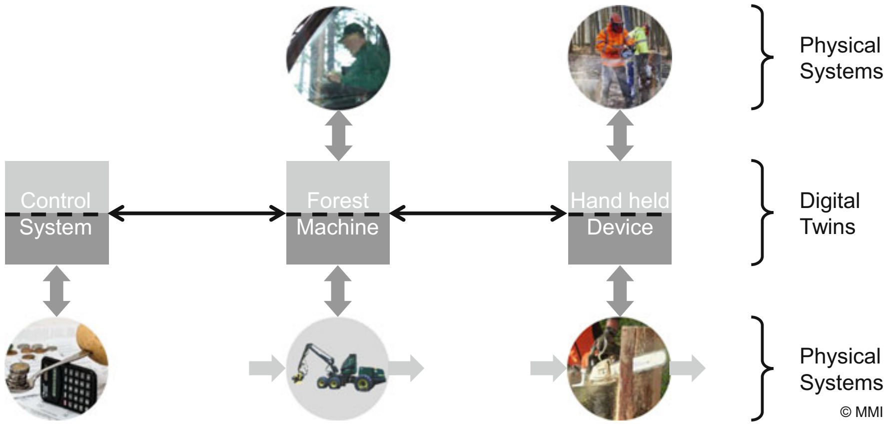


Fig. 5 Digital Twins as nodes in the IoT (the forest machine application) (The figure focuses on the networking aspect and is simplified accordingly. In fact, the driver and the forestry machine also interact on the physical level. The DTs in the middle are shown schematically as bicolored rectangles. This is to visualize that these DTs have an (e.g., company-) internal data area and an external data area (e.g., accessible for cooperation partners, clients, and contractors). A comprehensive overview of the DT in the IoT and this representation in its complete form is provided by (KWH4.0 2019))

means (horizontal arrows). The Assets themselves interact with each other on the physical level (horizontal block arrows). In this way, complex situation-specific value chains/networks can be set up, in which the Assets communicate with each other "at eye level."

2.6 *The Experimentable Digital Twin*

The perspectives mentioned so far each focus on individual elements of the dimensions listed in Fig. 4. For example, the perspective of the DT as a node in the IoT (see Sect. 2.5) focuses on the realization of the AAS, while the perspective of the DT as a simulation model focuses on the accurate simulation of the Asset's behavior. In contrast, an EDT considers the CPS/I4.0 component in its entirety, including all physical interactions as well as internal and external communications (see also Fig. 2). An EDT thus brings together the individual perspectives on a DT. Since an EDT is a 1-to-1 representation of the CPS (see Fig. 3), especially with respect to structure and behavior, the DT in the AAS is created from an EDT by omitting the simulation model for the Physical System and vice versa. In combination with the definition of the experimentable model (VDI 2000), this leads to the following definition of the EDT:

Definition *An* ***Experimental Digital Twin*** *(EDT) is a virtual digital 1-to-1 representation (structure, behavior, interaction) of a CPS in an experimentable model. It is a CPS-level Digital Twin implemented* *for/by the perspective of simulation technology. It integrates the Digital Twins of the Physical System and the IPS, interacts with its Real Twin or is itself part of the IPS.*

2.7 The Technical Realization of a Digital Twin

A closer look at the technical implementation of a DT also leads to different alternatives. It is obviously not clear whether a DT itself is a simulation model or whether it (as an AAS) contains one. To differentiate, Fig. 6 transfers the real CPS/I4.0 component shown schematically (dark gray) at the top, which consists of the Physical System/Asset (yellow), the IPS/AAS (light blue) and corresponding internal, external, and physical interfaces and connections, into different (E)DTs. If we focus on the pure word meaning or **semantics** of the DT, then it is a conceptual description or formal specification of the real CPS as shown in perspective 1 (shaded). If the **simulation model of the Physical System** is referred to as the DT, this leads to perspective 2 (light gray), whereas perspective 3 refers to the **simulation model of the entire CPS**, i.e., the EDT. Perspective 4 focuses on networking, where the terms DT and AAS are used synonymously. From an implementation perspective, a DT is often not implemented directly, but builds on data management, simulation, AI, or other technologies (often referred to as "middleware" or "platform"). If we look at the **use** of the DT in the real CPS against this background, this leads to perspective 5. Here, the DT "lives" in a suitable run-time environment that provides the necessary implementation basis. In perspective 5, the DTs from perspectives 1–4 themselves become a component of the CPS.

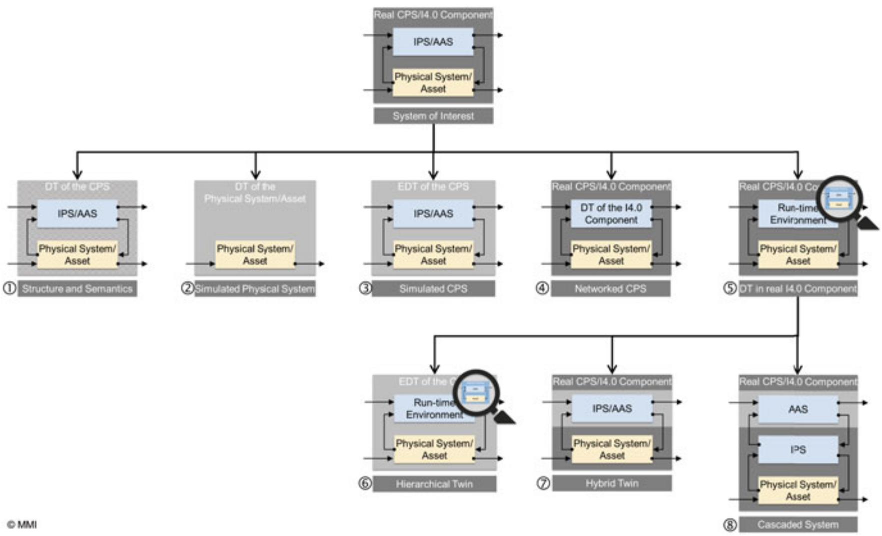


Fig. 6 Different perspectives on a Digital Twin of a real CPS/I4.0 component

The run-time environment provides further functionalities and uses the DT, for example, for simulation, observation, control, planning or operation. In addition, the run-time environment implements the communication interfaces of the DT both internally and externally.

If we take a detailed look at the **technical implementation**, we see that a comprehensive mapping of perspective 5 in the simulation must also consider the run-time environment and the DT(s) living there. This leads to perspective 6, an **hierarchical DT**, in which the (E)DT from the first level (perspectives 1–5) is part of a higher-level DT and here specifically its IPS or the run-time environment located there. The fusion of real and virtual world in perspective 5 can also be called a **Hybrid Twin** (visualized in perspective 7 and correspondingly shown in light gray/ dark gray). In its run-time environment, the DT often extends the IPS already present in the real system, e.g., taking over subordinated control tasks, and adds higher-level functionalities. This leads to the **cascaded system** shown in perspective 8.

Figure 6 provides concrete answers to questions such as “What is the (E)DT with respect to a real system?”, “How to realize an (E)DT?”, “Where does an (E)-DT live?”, or “Where to find the (E)DT?”. None of the representations can be labeled as right or wrong, but reflects different perspectives regarding Real Twin as system of interest and interaction between real systems and DTs. The EDT opens up a way to consistently describe and implement these individual realizations based on the EDT as a 1-to-1 representation. The only prerequisite for the structural realization of the perspectives shown is the interconnection of real and virtual (used in the EDT) communication infrastructures, so that the interconnection can take place across the boundaries of the worlds. This then finally blurs these boundaries between reality and virtuality and results in the convergence of real and virtual worlds.

3 Applications of Experimentable Digital Twins

And what is the added value of an EDT? The consolidation of all information on a Real Twin, be it a CPS/I4.0 component, an IPS/AAS or a Physical System/Asset, in one place taking into account structure and semantics of all Digital Artefacts is certainly a value in itself. This results in a “single source of truth” if it is possible to have exactly one EDT for a Real Twin. Regardless of what the user wants to know about a machine and what he wants to do with this machine, he will find the answer by “talking with his contact person,” the machine’s EDT.

3.1 The EDT-Scenario

On the other hand, the actual added value of EDTs for simulation is the possibility of connecting them with each other in a network-like manner and evaluating the

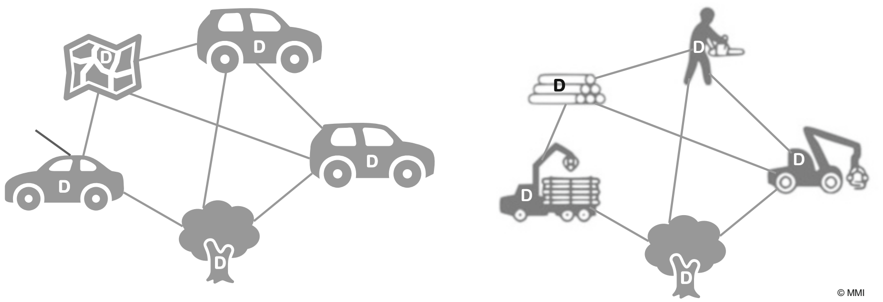


Fig. 7 Modeling of application scenarios using Experimentable Digital Twins

resulting networks appropriately. After relevant CPSs/I4.0 components of an application have been identified and described via EDTs, a concrete application/a specific use case can be modeled in its entirety solely via the networking of all EDTs involved. This networking takes place by connecting their ports on the three levels introduced in Sect. 2.2, internal and external communication as well as physical interaction in different disciplines. The result is an EDT-Scenario. Figure 7 shows an example of an EDT-Scenario from the automotive sector and one from forestry. The vehicles communicate with each other and with a map service and can physically interact with each other, with their environment and with the sensor. Thus, the term can be defined as follows:

Definition: *An **EDT-Scenario** is a model that instantiates multiple EDTs involved in an application according to their type descriptions. It specifies their attributes and describes their explicit interactions.*

3.2 Application-Specific Analysis of EDT-Scenarios

Applications such as the ones mentioned at the beginning are created by the suitable evaluation of these EDT-Scenarios as outlined in Fig. 8. The basis is often the highly detailed interdisciplinary, cross-domain, cross-system, cross-process, and cross-application simulation of an EDT (1). The EDT is created in processes of Semantic World Modelling, e.g., in the context of (model-based) systems engineering (2) or by remote sensing (3). Networked with other EDTs, it is integrated into surrounding systems and the operational environment. On this basis, its behavior can then be tested and analyzed in detail (4) and training data for AI algorithms can be generated (5). Using requirements and test cases defined in the context of systems engineering, virtual validation can take place (6) or system variants and parameters can be optimized in the context of decision support systems (DSSs) (7). The networking of EDTs with the IPSs of real CPSs enables their virtual commissioning (8), the networking of the IPS part of an EDT with the corresponding real Physical System enables the realization of its AAS and its networking in the IoT (9). The integration

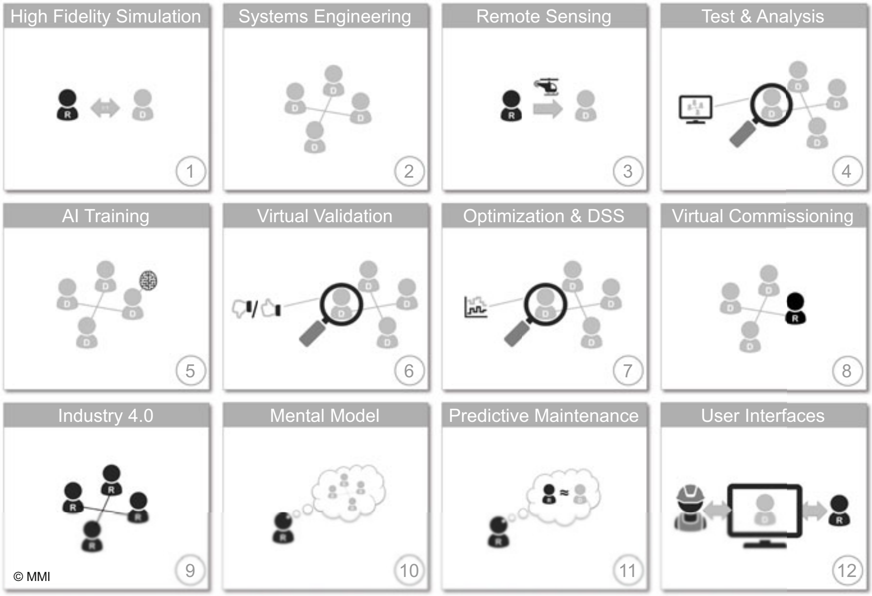


Fig. 8 Examples for EDT-based methods

of an EDT-Scenario updated via sensor data into a CPS realizes mental models, i.e., experimentable representations of the system in its environment for planning, optimization, and execution of actions to be performed (10). The comparison of the simulated with the actual behavior enables the realization of EDT-based predictive maintenance (11), the suitable interactive and intuitive presentation of the EDT for the operator allows for the efficient operation of the Physical System (12). The result are various **EDT-based methods** as outlined in Fig. 8, all realized based on the same set of EDTs.

3.3 Specification of Digital Twins and Their Interaction

It is advisable to initially specify a DT at an abstract level (perspective 1 in Fig. 6). The corresponding CPS is first described in several iterations with increasing level of detail. The domain-specific design and the technical implementation follow afterwards. One possible approach first compiles relevant business cases (step 1). Based on these, relevant CPSs making up the application are identified (step 2) and general information on structure, data, behavior, functions, and networking is compiled for each CPS (step 3). These are then further detailed (step 4) and finally combined in a formal specification (step 5). The first three steps provide an initial overview of the

CPSs involved and their networking, the first four steps consider the CPSs at a conceptual level. Only in the fifth step, the DTs are formally specified. This specification is then the starting point for the concrete technical realization according to the dimensions/perspectives shown in Figs. 4 and 6, where different realizations of the (E)DTs provide concrete technical implementations of the required views of the same (E)DT models created before.

4 Experimentable Digital Twins Cover the Entire Life Cycle

EDTs are ideally suited to bring together all the information arising in the life cycle of a CPS in an experimentable knowledge base. The EDT introduces a level of abstraction to initially describe CPSs in an abstract but at the same time descriptive way (conceptual model) and formally specify and model them afterwards (concrete model). Different application scenarios/business cases (control, command, supervision, networking, maintenance, . . .) then lead to 1) different perspectives (see Figs. 4 and 6) on one and the same EDTs, 2) different contributions to its core Digital Artefacts such as digital models, operational data, prognosis, communication and services, and 3) different methods using these EDTs (see Fig. 8). The descriptions/models of these EDTs are the basis for their subsequent technical implementation (Fig. 6), which leads to domain-specific detailing of the EDT—that at the same time continuously refines the EDT.

4.1 The (Development) Process

Figure 9 shows an overview of a typical development process using modular satellite systems as an example (see Sect. 5.1). Starting with the initial idea, the first step is to model the CPSs involved/required (in this case the satellite modules (building blocks) and their contents). The starting point typically is a systems engineering phase, where the CPSs involved as well as their interfaces are identified and described as EDTs (see also Sect. 3.3). Established development tools from the different disciplines/domains involved are used afterwards to model the EDTs in detail and to create detailed models of the EDT-components. In the next step, EDT-Scenarios for planned applications and operating situations are compiled using the modeled interfaces. These are then executed in so-called **Virtual Testbeds** (VTBs). VTBs are run-time environments/DT platforms for EDTs seen from the simulation perspective, which orchestrate and schedule the work of specialized simulators in co-simulation setups. For this purpose, EDT-Scenarios are automatically transferred into a set of executable models that can be experimented with

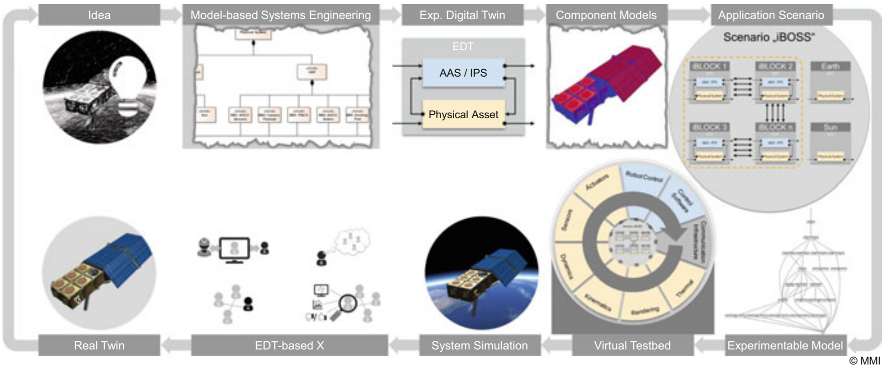


Fig. 9 The EDT in the life cycle of a modular satellite (The EDT-Scenario on the right describes on its left side a modular satellite consisting of four modules (blocks) interconnected by a universal 4-in-1 interface that links these blocks mechanically, electrically, thermally and on an IT level. On the right side, the environment represented by the EDTs of earth and sun is modeled. These two EDTs also interact with each other and the other EDTs on the physical level (e.g., in the dynamics and sensing domains). However, these interactions are automatically resolved by the EDT-Methodology and do not need to be explicitly modeled, which significantly simplifies the EDT-Scenario model (see also Schluse et al. 2018).)

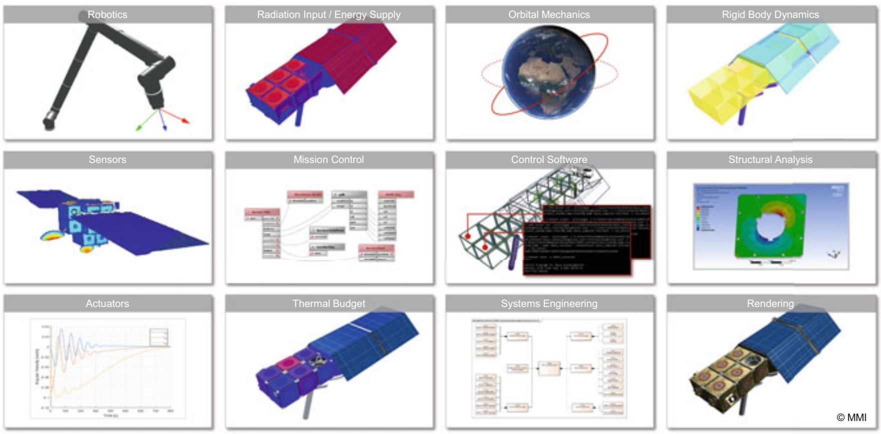


Fig. 10 Relevant simulation domains when developing modular satellite systems

(Schluse et al. 2018). At the same time, the necessary couplings between the selected simulation domains (see Fig. 10) are introduced. In the last step, parts of the EDTs are coupled with parts of real CPSs to implement hybrid, i.e., mixed virtual/real, setups as outlined in Figs. 6 and 8.

4.2 Simulation Domains for Experimentable Digital Twins

In all development steps, EDTs, EDT-Scenarios, and their execution in VTBs are always considered from all relevant perspectives. Only this way, all users, from the development engineer to the project manager, get the necessary overview of the overall system covering all relevant development results from all perspectives involved. Only this way, questions such as “How does the energy consumption of the overall system behave when the robot moves based on sensor data, taking into account orbital mechanics and thermal aspects as well as solar energy generation?” can be answered. For this, no case-specific simulation models have to be generated by the modeler itself. Instead, views on the overall EDT-Scenarios are compiled which in turn lead to the experimentable and coupled models already mentioned in the previous section. Figure 10 shows examples of individual perspectives in extension of Fig. 9.

5 Examples for EDT-Based Applications

Figure 11 outlines some selected examples of the application of the EDT-Methodology in different **EDT-Applications** in different application domains. Obviously, DTs and the EDT-Methodology are not limited to production technology and “classical” Industry 4.0 settings, but can be used in different application areas from aerospace to production technology, construction sites, intralogistics, building (automation) technology and even the environmental sector. Most of the time, only level of detail and scope of an EDT are specific for a concrete application, but not its modeling and implementation. That is why the EDT-Methodology enables an easy

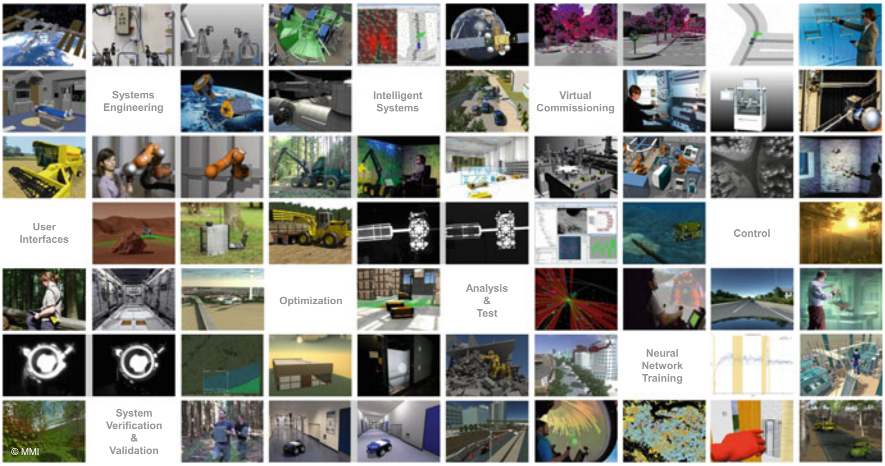


Fig. 11 Selected examples for EDT-based applications

transfer of existing EDTs between the individual applications and application domains. Thus, remote sensing-based tree modeling as well as semi-automated forest inventory, laser scanner-based localization or simulation-based validation of driver assistance systems always address the same EDT of a tree. Only the view on the tree EDT changes between the applications. For further illustration, this section looks at one application each for four of the EDT-based methods introduced in Fig. 8.

5.1 Test and Analysis of Modular Satellite Systems

The first application is test and analysis of modular satellite systems (method 4 in Fig. 8), which were carried out in the context of the development of the iBOSS concept (Weise et al. 2012, Osterloh et al. 2018). This example already formed the basis for the explanations in Sect. 3.3. The basic problem in the development of space systems is that they cannot be finally tested before their launch; the mission is the first overall system test. Therefore, corresponding tests of the overall system and especially the tests for correct interaction of all subsystems have to be carried out in the virtual world. In this context, it is important to systematically investigate all relevant operating situations. The starting point for the work was a detailed specification of iBOSS-based satellite systems in SysML. This defined the components involved as well as their structure and interfaces. In a second step, these components were detailed out by EDTs, i.e., an EDT with corresponding interfaces was created for each module. The EDTs contained detailed models for the simulation domains relevant for development as shown in Fig. 10. The basis were detailed CAD models of the Physical Systems which were extended by models and attributes for rigid body dynamics, orbital mechanics, and robotics. Bodies of particular interest such as the side surfaces of the individual modules were modeled as additional FEM models for the analysis of the structural dynamics. Thermal models of the modules and interfaces describe their thermal behavior. Furthermore, models for sensors and actuators as well as the radiation input were integrated, which particularly relied on the CAD models. The integration of the real control software, a complete robot control system as well as the mission control into the EDT allows for the complete operation of such systems in the VTB. By connecting the ports of the EDT, their physical interactions as well as internal and external communication were modeled. Up to this point, all EDTs have been considered completely independent of each other. Only their interconnection brings them together and is the starting point for converting the EDT-Scenario into an executable model. The figures above outline some selected examples.

5.2 *Control of Reconfigurable Robotic Workcells*

One of the challenges in using a reconfigurable robotic workcell is to configure, program, build and control this workcell in the best possible way for a given manufacturing task using the available cell elements (robots, grippers, tools, mounting frames, positionable fixtures, etc.). These steps were completely implemented using the EDT-Methodology. Again, building on a specification of the overall concept for the individual cell elements, EDTs were modeled as a 1-to-1 mapping of their real CPSs (see also Fig. 12). This was the basis to model concrete application scenarios by a network of interacting EDTs in the form of EDT-Scenarios as shown in the example for headlight assembly (see also (Schluse et al. 2018)). This task is performed by an experienced expert. When programming the assembly sequence, he can use all information about the cell components provided by the EDTs. This also holds for the robot and cell controls, which significantly simplifies the planning/ programming process. To use this information basis to control the virtual as well as the real robots, the required controllers are integrated in the EDTs themselves – which in turn evaluate the EDT-Scenario to do so. For real-world execution, the VTB is also available for the real-time operating system QNX to also provide the controller for the physical robot as outlined in perspective 5 in Fig. 5 and method 10 in Fig. 8. The EDT shown at the bottom right of Fig. 12 thus controls the real Assets shown at the top right.

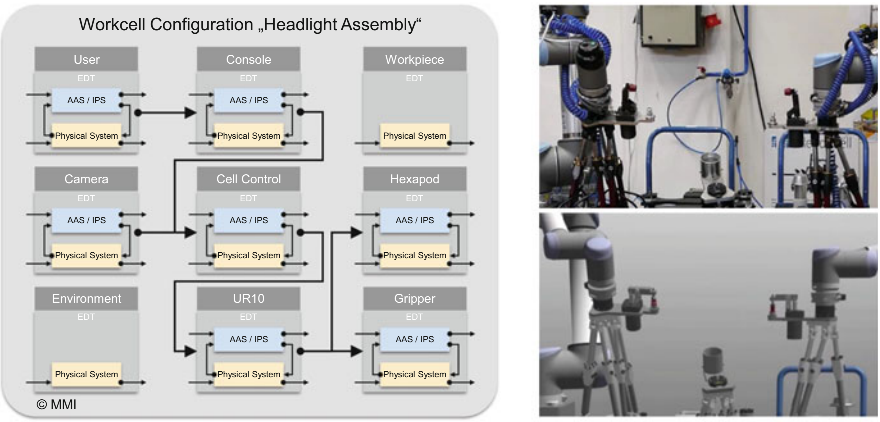


Fig. 12 Development and control of reconfigurable robotic workcells using EDT (see also Schluse et al. 2018)

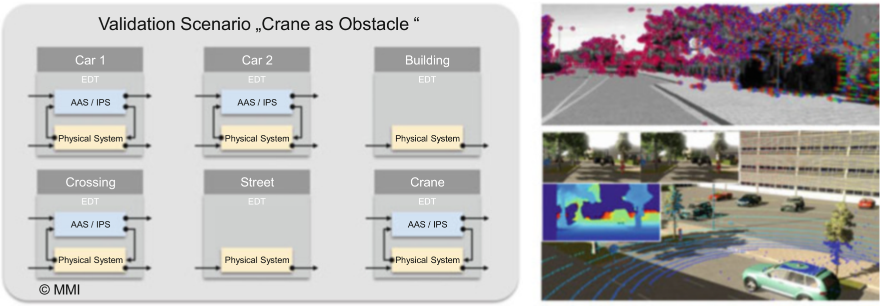


Fig. 13 Exemplary scenario to validate driver assistance systems (l.), examples for simulated sensor data and EDT-Scenarios (r.) (see also Thieling a. Roßmann 2018)

5.3 *Validation of Driver Assistance Systems*

The development of modern driver assistance systems and autonomous vehicles leads to new challenges in the development, testing and functional validation of such systems, since they must perceive their environment more comprehensively and, on this basis, plan actions and intervene more intensively in the driving dynamics than today. To test their final behavior, real test drives reach their limits in terms of time, cost, reproducibility, and test coverage. It is expected that simulations can close this gap, but this requires realistic methods, especially for sensor simulation, which are applied to a large number of (preferably automatically) configured test scenarios in the form of EDT-Scenarios reflecting, e.g., varying environmental conditions (weather, speeds, other cars, pedestrians, etc.). Exactly this flexibility is provided by the EDT-Methodology, where each scenario is modeled as a network of predefined and suitably parameterized EDTs.

Figure 13 outlines an example scenario, where two vehicles and an intersection communicate with each other. On the physical level, they also interact with the road, buildings, and a crane as passive and active CPS in the environment. The right side of the figure shows simulated sensor data, supplemented at the top by the result of a "feature detector" integrated into the EDT for motion estimation. By comparing the results of such algorithms, it could be shown that they provide comparable results for simulated as well as real data.

5.4 *Man-Machine Interaction for Rescue Robots*

Even today, robots can still be used best in structured and well-known environments. This is exactly what is not the case for rescue robots, and which must be compensated for by new "intelligent" robotic systems, but also by suitable user interfaces, which put the robot operator in a position to close the remaining perception and planning gaps in the best possible way. Figure 14 shows a suitable user interface as a

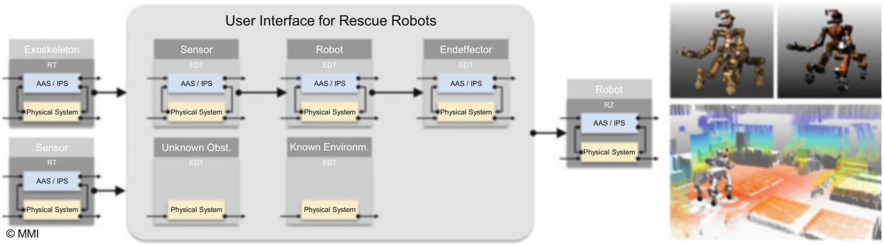


Fig. 14 User Interface for a rescue robot (s. Cichon a. Roßmann 2018)

technical implementation of method 12 from Fig. 8. The operator controls the real robot (on the right in the figure) with an exoskeleton (left). He is supported by the real sensors (left) of the robot, which update the EDT-Scenario to reflect the real environment as close as possible. The operator can directly control the real robot via the resulting EDT-Scenario (see bottom right in the figure). He can also use it to plan his actions or to generate force feedback for the real robot, which is located far away and thus connected with quite large communication delays. In addition, it also serves as a "mental model" for online planning of the robot actions themselves (see method 10 in Fig. 8).

6 Summary

DTs are currently experiencing a hype and are on their way to become a disruptive technology. DTs are both, the result of and the basis for digitalization. They provide semantics and structure, e.g.:

(1) for the formal description and practical use of, e.g., architecture, models, operating data, interfaces, functionalities, and behavior of the corresponding real Twins (CPSs/I4.0 components, Physical Systems/Assets, IPSs/AASs)
(2) for the realization of CPSs/I4.0 components
(3) for the prediction of their behavior
(4) for the situation-specific choreography/orchestration of CPSs/I4.0 components in systems of systems and value-added networks, and
(5) for the automation of complex value-added processes

The concept of DT shows enormous potential, which is only slowly being realized due to the current situation, which is characterized by ambiguous definitions and marketing-oriented use of the DT term. This situation certainly endangers the further development and application of this concept.

Although the concept of DT arose well before the Industry 4.0 approach, it is now often perceived as one of the central implementation components of Industry 4.0. Here, too, the main task of DTs is the synchronization of the real and virtual worlds—or even their convergence. DTs ensure the availability of all data

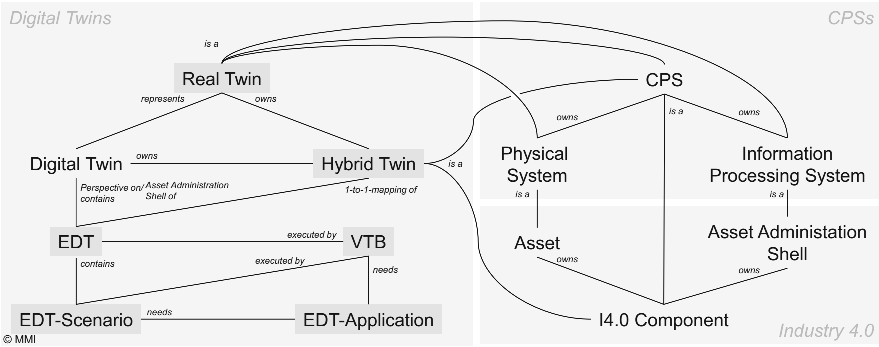


Fig. 15 Semantic net summarizing various terms relevant for EDTs, grouped in terms related to DTs, CPSs and Industry 4.0 terminology. Shaded terms are extensions by the EDT-Methodology

(environment, operation, process, . . .) at any place at any time and make (production) processes more reliable, flexible, and foreseeable. As nodes in the IoT, they enable the merging of data and functions and the exchange of information between machines as well as people and machines. They are used in development, production, maintenance, and many other application areas.

This chapter brings together the enormous range of applications of DTs on the one hand and the associated perspectives on DTs and methods using them, which are often perceived as competing and not matching each other, on the other hand. The EDT defines a common basis to consistently derive application/perspective-specific views for EDT-based methods and applications. The starting point is the perspective of simulation technology, which has proven itself for years as a sustainable basis for the integration of systems in the virtual world but also for the implementation of hybrid (i.e., partly virtual and partly real) systems. Figure 15 concludes by illustrating the established and newly introduced terms used in a semantic network whose center is formed by the triangle of Digital, Hybrid and Real Twin.

Acknowledgements This work was supported by the research projects INVIRTES, iBOSS-3 and ViTOS, funded by the German Aerospace Center (DLR) with funds from the Federal Ministry of Economics and Technology (BMWi), support codes 50RA1306, 50RA1203 and 50RA1304. The projects ReconCell and Centauro received funding from the European Union's Horizon 2020 research and innovation program under grant agreement Nos. 680431 and 644839.

References

Bibliographisches Institut GmbH Dudenverlag (2019): Duden Online, www.duden.de

Bosch (2020) Bosch IoT Things. Retrieved April 10, 2020, from https://docs.bosch-iot-suite.com/things/

Boschert S, Rosen R (2016) Digital twin-the simulation aspect. In: Hehenberger P, Bradley D (eds) Mechatronic futures: challenges and solutions for mechatronic systems and their designers.

Springer International Publishing, Cham, pp 59–74. https://doi.org/10.1007/978-3-319-32156-1_5
Center of Excellence for Forestry 4.0 (2019) Der Digitale Zwilling in Wald und Holz 4.0 – Semantik, Struktur, Perspektiven und Umsetzung. www.kwh40.de
Cichon T, Roßmann J (2018) Digital twins: assisting and supporting cooperation in human-robot teams. ICARCV, Singapore
Grieves M (2014) Digital Twin: Manufacturing Excellence through Virtual Factory Replication - A Whitepaper, www.apriso.com/library/Whitepaper_Dr_Grieves_DigitalTwin_%0AManufacturingExcellence.php
Lee J, Bagheri B, Kao H A (2015) A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. Manufact Lett 3(October 2017):18–23. https://doi.org/10.1016/j.mfglet.2014.12.001
Madni A, Madni C, Lucero S (2019) Leveraging digital twin technology in model-based systems engineering. Systems 7(1):7. https://doi.org/10.3390/systems7010007
OPC (2017) OPC Unified Architecture (OPC UA), opcfoundation.org
Osterloh T, Dahmen U, Roßmann J (2018) Full lifecycle support for modular satellite systems provided by comprehensive Virtual Testbeds. i-SAIRAS, Madrid
Panetta K (2018) 5 Trends Emerge in the Gartner Hype Cycle for Emerging Technologies. www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018
Plattform Industrie 4.0 (2020) Glossary. https://www.plattform-i40.de/PI40/Navigation/EN/Industrie40/Glossary/glossary.html
Schluse M, Priggemeyer M, Atorf L (2018) Experimentable digital twins – streamlining simulation-based systems engineering for Industry 4.0. IEEE Trans Ind Inform 14(4)
Shafto M, Conroy M, Doyle R et al (2010) DRAFT modeling, simulation, information - technology and processing roadmap. Technol Area 11
Sondermann B (2018) Simulationsgestützte Landmarkendetektion, Lokalisierung und Modellgenerierung für mobile Systeme. Dissertation, RWTH Aachen University
Thieling J, Roßmann J (2018) Highly-scalable and generalized sensor structures for efficient physically-based simulation of multi-modal sensor networks. ICST, Limerick
VDI/VDE-Richtlinie 3693 (2016) Virtuelle Inbetriebnahme, Modellarten und Glossar
VDI-Richtlinie 3633 (2000) Simulation von Logistik-, Materialfluß- und Produktionssystemen. Düsseldorf
Weise J, Briess K, Adomeit A, Reimerdes H-G, Göller M, Dillmann R (2012) An intelligent building blocks concept for on-orbit-satellite servicing. i-SAIRAS, Turin
Wimmer L (2009) Die Gestaltung digitaler Artefakte - Designtheoretische Ansätze in der Human-Computer Interaction. Universität St, Gallen

Towards the Digital University

Hans-Joachim Bungartz

Contents

1 Introduction

Digitization is on everyone's lips today. There is hardly any research agenda without digitization components, no trade association that does not refer to the impending impoverishment without it, hardly any social debate in the course of which someone does not talk about the opportunities and risks of digitization. Gone are the days when high school graduates ignored informatics due to vague ideas about the field and preferred to opt for something more down-to-earth or even more solid, such as mechanical engineering. **Informatics** is "in." Its devices and processes have a decisive influence on the zeitgeist. They not only stand for technical or economic progress, but also for lifestyle and, thus, they make classic icons such as cars look rather old increasingly often. In politics, too, digitization has established itself as a large, independent topic. In Germany, the federal and state governments are launching large programs for infrastructures (**National High-Performance Computing—NHR, National Research Data Infrastructure—NFDI, Competence Centers for Big Data and Machine Learning,** etc.) as well as for relevant research, the latter currently under the umbrella term of the latest digital hype,

H.-J. Bungartz (✉)
Department of Informatics, Technical University of Munich, München, Germany
e-mail: bungartz@tum.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_38

Artificial Intelligence. But what is it about, beyond these buzzwords? And what does digitization mean for universities? What does, to add another term of its own, “University 4.0” look like?

Basically, it is about comprehensive, profound changes that have now reached almost all areas of our private and professional life and are unparalleled. Informatics naturally plays a key role here as the primary supplier of methodology and technology. It is no coincidence that the Department of Informatics is now clearly the largest department in terms of numbers of students at my university, the Technical University of Munich, and we observe similar developments elsewhere. But it is not a claim to sole representation, but rather a leading role as primus inter pares, because all disciplines are affected by digitization. On the one hand, they urgently need access to the appropriate infrastructure, services, and competence; on the other hand, they also contribute in one way or another to its progress. In the following, we want to deal with a few facets of this transformation—one quickly speaks of revolution—and their effects on science.

2 Science 4.0 Is “Computational”: Computer- and Computation-Aided

People of my generation have experienced several transitions from the analog, continuous to the digital, step-line very closely, and often somewhat bizarre fan communities have made the parting cult in the process. This was for example the case when the CD and later the DVD replaced long-playing records, compact cassettes, and video cassettes. The same holds true for the introduction of digital cameras. Only with the telephone, the dial plate was not really mourned. The extent of the whole development was not really foreseeable—I remember all too well statements of the kind “Now let’s take it easy – who knows how long this will go on with these computers.” Even heroes of the digital cosmos repeatedly fell victim to the pace of development and in retrospect made themselves a laughing stock with statements such as “640 kB ought to be enough for anybody” or, from today’s perspective, completely absurd prognoses about the world market for workstations. “Kilobyte” is now the Stone Age. Mega and Giga are long behind us, Tera and Peta are there. Exa is at the door, and you can already read about Zetta and Yotta. One can only hope that our Greek friends do not run out of prefixes, because the international system of units, SI, has thrown in the towel at Yotta. If you consider that all of this only began 70 or 80 years ago, then this development is truly breathtaking. But you can also see that although the digital transformation of these days may have made its mark on business, politics, and society, a lot of it is technologically not new.

With all the focus on “Industry 4.0,” it sometimes takes a back seat to the fact that the transformation processes in science are no less fundamental than in business. The “digital university”—as unfortunate as the term may be—will consequently be completely different from the Alma Mater we are familiar with. Disruptive processes everywhere. Research works differently today than it did 20 years ago. This applies to experiments as well as to theories, and now for almost all subjects. Thus, in 20 years everything will be different again. Specialist literature is used less and less in printed form and libraries are mutating into information centers. The **open science** debate has already led to new business models (by the way, not necessarily cheaper ones), and more will happen. Even confidential things are done electronically—one click, a hidden authentication and authorization, and all documents are accessible. Another click and your own statement is uploaded. Teaching and learning also take place via online channels—a development that has seen a tremendous acceleration due to the COVID-19 pandemic, and anyone who, like me, is convinced of the added value of face-to-face teaching will find a wealth of supporting formats. The legendary CIP pools—not so long ago the means par excellence for the broad penetration of university education with computers—are increasingly disappearing. In contrast, today's WLANs have to master thousands of mobile devices at the same time. My office is wherever I am right now—thanks to Wi-Fi and Eduroam. On the one hand, I could no longer function otherwise. On the other hand, every work-life balance threatens to fall apart, which in view of the high intrinsic motivation in the academic field is likely to be an even greater problem than in the private sector. Increasingly, administrative processes are leaving the world of paper—or, let us say, could leave it. **Campus management systems** provide an example of this, including promise and horror. **Research data management** has become a national task to a **national research infrastructure**—not only for legal or documentation reasons, but also out of a fundamental research interest. And this list can be supplemented almost at will.

Of course, this also entails a changed, expanded range of tasks for a **data center**—which actually can no longer be a mere data center. A change only in the service units or their offerings is not enough. Much more has become different, so much more has to change. But more on that later.

Let us first look at the changed subject landscape. Of course, Humboldt is still omnipresent today, but something is happening. A particularly visible example of the penetration of science by computer technology and—much more importantly—by the methodologies, ways of thinking and previously hidden access to research and teaching associated with it, has been given a name of its own. Today one speaks of "***computer-aided*** *science*" or "***computational science and engineering***," whereby *science and engineering* now extends far beyond the classical natural and engineering sciences.

The digitization of the sciences started quite harmlessly: "***Computational*** *1.0*" was little more than theses in Word, scientific publications in LaTeX, laboratory books in Excel, process computers for controlling experiments, scanners, and software for digitizing an old handwriting, and literature research at cozy terminals. All of this is unquestionably extremely helpful, but still more of the classic pocket calculator type: nice tools that you definitely do not want to miss anymore, but which do not yet shake the foundations of scientific work.

The big breakthrough came with "***Computational*** *2.0*"—numerical simulation and computing on a large scale. **High- and high-performance computing** (the fine, also political, difference between "hoch" and "höchst" in German will not be discussed further at this point). It felt like a long time ago when computers were still primarily used for arithmetics. Today, **simulations**—computer-aided experiments—have become an indispensable part of scientific practice in many disciplines.

Such a simulation is a complex thing. You need powerful physical and mathematical models, efficient algorithms, powerful software, and meaningful visualizations as well as data analysis to interpret the calculated results. And that is exactly what makes the simulation so wonderfully interdisciplinary as well as orthogonal to all established subjects. As soon as things get really complicated, no mathematician, computer scientist, engineer, or natural scientist has a real chance to go the long way successfully on their own. Here, the digital transformation is also challenging our academic structures: Silo-like departments not only no longer reflect the modern research landscape, they can even hinder it. Instead of noticing new trends quickly, or even setting them up, people prefer to get bogged down in never-ending discussions and quarrels about to whom "belongs" a new topic, a new subject, a new professorship, or a new course of study. As a result, we must also question the structure of universities in the course of digitization.

Most universities respond by introducing a sort of cross-department structure. On the one hand, of course, this is by no means a wrong approach. On the other hand, it is also a dangerous crutch, because you do not dare to completely question the existing structures as a hole. However, many decades after Humboldt, exactly the latter might be advisable. Where is it written that the existing **subject structure** must remain the same forever? Let us be honest. If we had to map the disciplinary landscape from scratch today, without the centuries or at least decades of tradition of what we now call "disciplines"—would then "physics" or "chemistry" or "electrical engineering" really come out again, or maybe a completely different constellation? The professional profiles "outside" have long since shed the narrow disciplinary corsets and developed further—even if a number of professional associations sometimes, extremely annoyingly, act as dinosaur-like Lord Privy Seals and try to exert a very backward-looking influence on the universities.

The computer, as its central enabler of course, plays a prominent role in **digitization**. There is, for example, **Moore's** famous and much strained **law**, which—formulated a good 50 years ago and still somehow valid despite signs of age—predicts an increase in performance of around a factor of 10 every 5 years. Originally, this related to the integration density of transistors on computer chips. Later, it concerned more and more to the computing speed. A factor of 10 every 5 years—that is remarkable. Even if such comparisons always lag in a certain way: Where would we be in terms of the fuel consumption of cars or the speed of airplanes, if something similar had happened since the Model T and the Super Constellation?

So, we are now approaching the exa-flop with 10^{18} floating point operations per second. If you take human performance as benchmark, that would mean going around the equator in approximately 40 fs. The important fact here is we not only have these computer capabilities, we also need them for digital science. For this reason, from 2013 to 2015 the German **Science Council** dealt with models for a sustainable supply of the German science landscape with computing capacity and computational competence in the field of **high-performance computing**. Actually, for the second time, it must be pointed out that the first toss in the 90s had brought about the service and supply pyramid and in particular the three national HPC centers in Stuttgart, Jülich, and Munich at the top, which later became the troika in the Gauss Center for Supercomputing—tier-1 in the national hierarchy, in Europe as tier-0. And since no pyramid can only stand on the top, the Gauss Alliance was also founded—as a platform for tier-2.

Now, on the second attempt, a few birth defects from back then should be cleared up. The result—the draft of an infrastructure for the **National High-Performance Computing, NHR**—contained some pioneering proposals that had been called for a long time: institutional funding, overall consideration of investments, operating costs, and personnel—no racks without brains. After intensive consultations in the GWK, the joint science conference of the federal and state governments, the **NHR** was launched in autumn 2018. The key points from the paper of the **Science Council** have been implemented—if not completely, then more than ever before. This is not the case everywhere—at the European level, EuroHPC is for example moving, unfortunately, much more in the direction of the racks, the hardware. Regarding the NHR, the following can be said: The German Research and Educational Network (DFN) was commissioned to set up the NHR office. A strategy committee was installed to transfer the concept to implementation. A central challenge is the transition from the "old" system—the **Gauss Alliance**—to the new—**NHR**. Now that the first NHR centers have been selected, a successful one will not just be a provider of cycles—this could also be operated centrally from Buxtehude or Iceland. It is also about the competencies associated with the computer cabinets, which are much more important in the area.

Of course, we also come up against the limits of previous technology. When there are just a few atoms left on the chip between the components; when the clock frequency requires higher signal speeds than the speed of light; when power plants have to cluster around mainframes; when jobs have to be dynamically distributed among hundreds of millions of compute cores; then, it becomes clear that fundamental algorithmic innovations are required here, which no system manufacturer can deliver offhand. Thus, it does not happen by itself—we have to do something for it.

An example: How would a giant such as the Airbus A380 aircraft be designed, if it had to be lifted into the air not by four mighty engines, but by a few hundred thousand standard hair dryers with the same total output? Our engineers would manage the challenge, but the colossus would probably look quite different (the fact that the A380 has withdrawn in the meantime is another topic that popped up while writing this article—but far less digitization-driven…). With computers, we have exactly the same situation: ant processors, no elephants—which one has to coordinate somehow. The measurement criteria also tumble. Perhaps we no longer need the numerical solver for linear systems of equations that gets by with the smallest number of arithmetic operations, but the "coolest" one—that is, the one that solves a given system of equations with the least of heat production. Whether it will look so different than before is another matter. It is therefore also and above all to make algorithms and software fit for the "Exa-scale Age," with *evo*lutionary and *revo*lutionary approaches. This is also digital transformation.

But back to the **simulation**. The next step has long been taken. If the experiment on the computer was initially used to supplement or replace the familiar experiment, the numerical simulation itself is now questioned. For example, instead of a new crash simulation, one learns from the data already available today—be it measurements from real crash tests or the results of previous simulations. If this is done with sufficient skill, simulation results for model variants or for an intermediate generation with only a slight "facelift" can be predicted surprisingly precisely—and with considerably more efficient use of large calculations. We "learn" from all available data and thus save ourselves time-consuming crash tests *and* otherwise required simulation processes. In contrast to many widespread application scenarios of **machine learning**, the starting position here is not the legendary "Da stell'n wer uns mal janz dumm" (we are going to pretend we are very stupid now) from the German movie "Feuerzangenbowle," but a lot of prior knowledge from physics. That is why one speaks in this context of "physics-informed **machine learning**" or "scientific **machine learning**"—a term that not without good reason evokes memories of the genesis of "scientific computing." For the first time, a real bridge between the classic approaches in science—theory and experiment—opens up. Mathematical models allow for an improved design of experiments through simulations, and the wealth of data from experiments in turn allows for the calibration and optimization of models. It is this bridge that made ***computational*** the "third paradigm," the third way to gain knowledge alongside theory and experiment, models, and data.

This brings us to one of the more recent digital hypes that make up "4.0" in the title of this handbook: Big Data—or "***Computational*** *3.0.*" Indeed, the data-intensive process, the focus **on data management and analysis** is just another form of "**computational**." And this focus continues the transfiguration of scientific work. Still, much more is different than you might think because data-centering is accompanied by a further step in dissemination, if you will, a step towards the democratization of the digital revolution in science. While the STEM subjects were still largely among themselves in the club of "computing disciplines," everyone is now involved. With business analytics, business administration has discovered, among other things, how amazing insights can be gained from and with enterprise data. Sociologists pounce on Twitter data. The life sciences are taking on their treasure trove of sequencing data. Finally, the humanities also got involved—the digital humanities are on everyone's lips today.

Suddenly, you do not have to want or be able to solve a differential equation to play along in this concert—everyone has access to data. One's own data usually also appears big, and everyone can imagine doing something with it. But anyone who believes that **machine learning** can do with less mathematics will be amazed. The variety of data sources knows hardly any boundaries. Just think of the large-scale experiments in physics such as the Large Hadron Collider (LHC) or Copernicus, medical image data or digitized books as well as works of art. And thus, disciplines that 10 years ago did not really know what data was, today talk about **data science**, **big data**, and **machine learning**—and, mind you, that is a good thing. This also contributes to the transformation and requires a corresponding rethinking of the actors in science. Such *data science*, *engineering, and analytics* are, of course, also to a large extent interdisciplinary. Consequently, it is pretty clear. The possibilities of **data analytics**, **machine learning**, **deep learning**, or **Artificial Intelligence** will change science forever. Even if a lot of the methodology has been on the move for a long time, and even if above all it is the increased computing power that makes complex analysis of even the largest amounts of data feasible today—it is happening now.

In general, the ties between **HPC** and **AI** are closer than one might think. Classic **AI** and classic **machine learning** are strongly heuristic. You can do amazing things with amazing accuracy, but you usually do not really know why this or that comes out. In any case, it is enough to analyse customer profiles. But it is not enough to drive autonomously or to support heart transplants. To do this, we need the whole thing to be reliable and much more mathematical. One speaks of reliable and comprehensible **machine learning**. This is only possible if you work numerically. Eigenvalues, singular values, kernel-based methods, dimension reduction and discretization of the feature space are a few keywords in this regard—keywords that are well known in numerical mathematics and HPC. **The machine learning** codes must also become more efficient, if big problems are to be tackled—this is also a topic that has been present in the HPC for decades. So, if **AI** can do something today, it can do it because of **HPC** and "**computational**." Thus, if it is to be able to do even more, it urgently needs more HPC.

All of this makes it clear that the digital revolution in research, **simulation,** and **big data**, did not invent any lucky bag that allows us to pipette faster or to find what we are looking for more quickly in libraries. **Simulation** and **machine learning** are not "hip" today and will be replaced by a new "bandwagon" tomorrow. Not at all—the way in which research is carried out is currently being completely redefined. Experiments are designed in an automated experimental design. Models are derived based on data. Uncertainty quantification allows us to increase the reliability of forecasts in the sense of a "**predictive science**." Experiments, models, and simulations are integrated and, thus, create their next generation. As a result, completely new cross-references can be made, but more importantly analogies, of which one was not even remotely aware, can be discovered. This in turn means that, on the one hand, **research data** of the communities must be systematically organized, but on the other hand, it must not only remain a matter for the individual communities, which usually adhere to the principle "my data is completely different from yours." This, too, has now been identified as a national task in Germany and thus materializes itself as the **National Research Data Infrastructure (NFDI)**. If this instrument will be used correctly, hardly anything will remind you of the classic monodimensional and monodisciplinary laboratory.

3 Implications for Education

In view of all this, the fundamental question of the type of education also arises at universities and colleges—how much digitization does a subject need or tolerate? Do we need digitized biologists and Romanists, or biologized or Romanized computer scientists or, in the end, even completely new "**Computational Scientists**," as investigated by another working group set up by the **Science Council**, or "**Data Scientists**," which many are already calling for? This often happens without having a clue where they come from and what they should be able to do… Many have not yet understood, perhaps also do not want to understand, that neither an informatics course plus some Baudelaire or black hole nor a Romance studies course with partial focus on digital literacy really works. No, it is not that simple. There are cases in which new programs help—the now numerous "**Computational X programs**" are good examples. However, it is just as important that the established courses continue to develop. A degree in physics or archeology must look completely different in the digital age—without losing its foundation. Nothing else than physics or archeology, but in a different way. This means not to get lost in details à la "Where should we get the weekly hours or credits, respectively, from?", which works so wonderfully as a killer against all change and is used repeatedly with the greatest enthusiasm, even dedication. There must also be more collaborative formats, beyond department borders—modules, courses, or something completely different.

Furthermore, we need a new, or should we say, an expanded **scientific ethic** that is reflected in the processes and must be communicated early on in the training. After all, it is no longer just a question of whether someone kills mice in animal experiments, but also of whether virtual cloning actually has the same ethical dimension as the physical endeavor. Again, we must deal with the interplay between the real and the virtual world. Everyone knows Dolly the sheep—but who of us knows the extent to which virtual cloning is taking place around the world right now? Rarely since World War II has there been more research into nuclear and similar weapons than at the present time. But when we do not have an atoll in the South Pacific blowing up in our faces, in contrast to just a few bytes, nobody seems to care. Why should somebody? Apparently, nothing breaks. The topic of data is also highly ethically relevant—**privacy** and **open science** clearly show that. Statutory regulations must not be taboo neither. They often find their origin in a time with completely different possibilities—for the cops as for the robbers. Sometimes an additional protective mechanism is required quickly, sometimes a dusty one should be disposed of with verve. There is no question that clinical data, for example, requires special protection, but the fact that it is often not allowed to leave the hospital premises is technologically obsolete and hinders biomedical research. Namely, this leads to "secondary IT worlds" at university hospitals, which must then continuously being expanded. Much is scientifically and economically nonsensical and is not even necessary for the undisputed safety requirements.

The question above about adequate education in the digital age brings us to teaching. Here, too, the digital change arrived in various phases. Laptop and beamer replaced the good old overhead projector, which actually disappeared quite quickly and silently, whereas the slate, especially in subjects such as mathematics, has been a great analogue stronghold to this day. Rightly so! Then came the time of "**virtual universities**" and **e-learning**—with a mixture of motivations that is difficult to see through: new technical possibilities, new didactic possibilities, easier access to university education, but also hope for possible savings.

Teaching and learning platforms emerged. Lecture recordings became popular. These should enable the optimally supported follow-up of the lecture hour or, in the event of a hindrance, a comfortable re-learning of the missed material. But these examples also contain the problem that you must be able to deal sensibly with those new possibilities. A typical consequence of offering lecture recordings: First, the rows in the lecture halls are thinned, then in the exercise groups and finally the success rate in the exams drops. We recorded participation in the lecture and exercises as well as the grades in the exams in various lectures for 2 years—simply on the fly, without any claim of statistical relevance or empirical research. The result: The average grade of the "present" was 1.1 to 1.4 grade levels better than that of the "absent." Now one must be careful with the conclusions from such shots from the hip since, after all, the better students are typically more likely to be among those present. But still—the digital shot can obviously backfire.

The next trend in the wonderful world of digital teaching was the often-mentioned MOOCs (Massive Open Online Courses) or their little brother SPOCs (Small Private Online Courses). As so often, it started in the US—an AI lecture by the Stanford computer scientist Sebastian Thrun from Germany found hundreds of thousands of enthusiastic listeners. At the end, a test was held all over the world—several hundred examinees also turned up in Munich, many of them flown in, especially from Eastern European countries. Well, something else makes sense, even without a CO_2 debate. Nevertheless, distance learning offers new possibilities, can facilitate access, and provide effective support, particularly in the transition from traditional continuous study to life-long learning.

Needless to say that the pandemic dramatically accelerated the process of digitization of teaching in an unrivalled way. Once everyone suddenly "had to," it did work . . . somehow.

In addition to the noble goals of further democratizing **university education** or the individualization of study programs, there are of course others as well. Imagine, for example, that each candidate pays an examination fee of just 10 dollars. The dollar signs on faces are immediately recognizable in the view of this great new business opportunity. Thus, many MOOC protagonists immediately went for a sabbatical—"has taken a sabbatical to found a company" was the standard response. This can of course be praised as true entrepreneurship and some relevant players such as Alison, Coursera, Udacity, and edX are now present. But one can also have doubts as to whether this is really the dawn of the beautiful new era of teaching or whether this is what this movement is really about. But that is the way it is—opportunities and risks are closely related.

In addition to the question "How does **digitization** change **teaching**?" there is of course another issue. How—in which formats, in what depth and to what extent—

must a **training** in digitization technologies and techniques flow into all courses, especially into these "further away from digitization?" While for engineering and natural science subjects, a well-founded education in "higher mathematics" is self-evident today and depending on the importing degree program, is well established, it has largely been different with information technology topics up to now.

Here a one-semester "Introduction to Computer Science," there a programming course, perhaps combined with a dash of numerics—that is it usually. In the future, these may be supplemented by "Foundations of Data Science" or "Machine Learning for everyone." This should by no means be misunderstood as a plea for a four-semester cycle "Higher digitization 1–4" or the like—that would probably be doomed to failure, because it seems far too much "old school." Nevertheless, something has to happen here—the mere "They will manage that, after all, the young generation is information technology savvy" is certainly no longer enough. We cannot avoid one or the other conflict. New topics have become relevant and have to be integrated. Hence, something else must be sacrificed, for better or for worse. That may be painful, but that is how it is with change.

4 Administrative Implications

Talking about science and universities, you naturally think of research and teaching first. We have highlighted the effects of digitization on these areas in the previous two sections. The consequences for structures and overall processes are discussed in the following section. The first thing to discuss here is the **administration** of the university. To this end, information technology systems have been introduced at a large scale in recent years, for example corporate software for the entire accounting area or **campus management systems** for the study area. There are commercial solutions from established software providers, which, however, often suffer from insufficient fit for specific university requirements (for example, missing or not offered functionalities) as well as from high procurement costs. In addition, other providers have emerged from a university context or from joint initiatives such as HIS (Hochschul-Informations-System eG, formerly GmbH), especially for **campus management software** and **university information systems**. Other universities rely on self-built systems that have evolved over time.

Without digitally subservient spirits, the diverse project activities as well as the diversified and Bologna-based study and doctoral activities of today's universities can no longer be managed effectively and efficiently. During study operations, for example, isolated solutions have emerged in many places over the years—sometimes as innovative pioneers, however mostly in response to necessity, as there were no central tools, but an immediate need for action. Lecture hall occupancy, allocation of seminar and internship places to students (with the greatest possible matching of the individual wishes of both sides), short-term booking of learning rooms, organization of exam inspections, or online exam systems are a few of the topics to be mentioned here. To take just one example, the exam inspection: If almost 2000 students are registered in a beginners' course and only 20% of them want to subsequently use their right to inspect their corrected exams, how can this be logistically accomplished without an appropriate online procedure? As a result, you have topsy-turvy tools, but no clean interfaces, but at least you have something.

Today, there are still whole areas in which the archaic is more the rule than the exception: colorful cardboard covers on which the vacation times are to be entered, forms in which deviations from the standard flextime are to be noted; recruitment processes that trigger fear and horror in the forests of the Amazon region in terms of their paper consumption and are often so protracted that those to be recruited have meanwhile started elsewhere; business travel processes that—from the application to the accounting—take exclusively place on paper; reports on expenditure of funds, for which the numbers are repeatedly transferred manually from one system to the other and then finalized on paper—including transfer errors; rebooking—which is part of everyday university business especially nowadays; and in particular the resulting changes in vacation entitlement or teaching load, which are calculated manually and entered somewhere (hopefully in the right place), or... The list could be continued almost infinitely. It is pointless to argue about the reasons, especially since these are actually well known. Particularly popular and fatal is certainly the reference to force majeure: "The ministry wants the data in paper form anyway, so why should we..." This, of course, is also a knockout argument that should finally be knocked out by someone. Thus, the entry into the change must be a quick one, and it must not stop at entry level.

5 Processes and Governance

In addition to research and teaching, the digital age has also captured all processes at universities. **Information infrastructure** is the key word here. *"Information technology has become an indispensable part of any research infrastructure, thus information has become a central location factor for science and business."*—this is how the foreword by the former DFG President Kleiner on the recommendations 2010–2015 of the KfR, the Commission for **IT Infrastructure** of the **DFG**, begins. The current edition (2016–2020), unfortunately, was downgraded to a mere "statement," but contains similar messages. These concern the organization, structures, and **governance** of universities. At this point, briefly summarized and to the point, a few of the theses as they can be found in the places mentioned and elsewhere.

First: The change that we are currently experiencing in science (autonomy of universities, excellence initiative or excellence strategy, Bologna process, internationalization, alliances) is, on the one hand, continuously changing the profile of requirements for information technology at universities, but on the other hand can only be adequately controlled through their use.

Second: IT infrastructure has become the backbone of modern science. Science is not only using information technology to an even greater extent, it is changing fundamentally because of IT.

Third: The “I” in **IT infrastructure** stands for “information”—which shows the close proximity of **information infrastructures** to IT infrastructures. It is undisputed that the former are far more than a wonderful world for techies. Nevertheless, it does not work without the technological dimension. One cannot ponder or plan such infrastructures for information or data, if one only starts with “one’s” data and ignores computers, storage and networks—the necessary basic technologies.

Fourth: IT support needs process modeling. For example, anyone who tries to implement campus management software before all processes have been modelled, has forfeited any right to be annoyed by coffee machines that, in the best slapstick fashion, first dispense the coffee and then the mug. . .

Fifth: To clear up a big and unfortunately very widespread misunderstanding. The big goal of digital technology has never been, is not, and will never be “cheaper,” but “better,” even if some university chancellors may wish otherwise. This will also not only apply to a transition phase (always associated with additional costs), but permanently.

Sixth: Outsourcing must be carried out with caution, but offers further efficiency potential in the sense of an “economy-of-scale.” Activities from science and its stakeholders (examples include the cloud offerings of individual universities or data centers, local or regional university associations at the state level, at the level of research organizations such as the Helmholtz Association or also throughout Germany, for example within the framework of the federated services of the German Research and Educational Network, the “**DFN**-Cloud”) certainly will gain more relevance. Even if there are good reasons for reservations about some commercial providers when it comes to security, reliability and sustainability, initiatives such as the recently successful Europe-wide tender for cloud services within the framework of GÉANT, in the result of which all **DFN** users and thus practically all research institutions in Germany can participate (an interesting model that can also access the offers of commercial providers and enables to do so without having sleepless nights), for example, offer great possibilities.

Seventh: IT infrastructure needs to be anchored in the **governance** structure of the university. A Chief Information Officer (CIO) with a strategic mandate—no matter in which of the various forms (strategic, operational, as a body, . . .)—is not a panacea, but can be an accelerator. The entirety of all computer-aided processes in research, teaching, and administration, especially regarding digitization and the IT infrastructure, are the crown jewels of a university that is increasingly becoming sensitive to CRCs, ERC grants, clusters of excellence, and Leibniz prizes (and whatever universities are looking for today). This is clear as daylight and increasingly perceived internationally as such. In this country, however, the word has not yet got around everywhere.

6 Risks

Let us briefly turn to the risks of digital transformation because there are not too scarce. As always, they are especially scary when they are new, and their consequences are still difficult to assess. As always, it is also true that the individual is initially challenged to deal responsibly with the new possibilities: six hours a day in front of the TV is not healthy and six hours of chatting on social networks is probably neither. Anyone who willingly reveals their private data on the Internet should not be outraged when others take note of it. Moreover, anyone who sends unencrypted emails is acting grossly negligent (even if I would rather feel sorry for someone who has to, may, or even wants to read my emails)—therefore you do not have to brand every email traffic as generally unsafe. If we come to a mature approach to the digital world in every respect, the opportunities will clearly outweigh the risks.

Primarily, it is about the omnipresent security question—also in science. In this case, the German word "Sicherheit" has at least three different facets: First, we need ***safety***—nothing should be damaged or lost. This applies permanently wherever the legislature or even just a self-binding code prescribes long-term storage—for example for medical data or as the circumstances require for research data. And that is technically not easily possible. Will our storage devices, reading or evaluation software still work in 10 or 20 years? Is the archiving redundant enough to enable adequate restoring procedures in a worst-case scenario? Second, we demand ***security***—no one should be able to read along without authorization, write manipulatively or do anything illegal. And third, we long for ***privacy***—a piece of real or virtual existence without an audience.

While **safety** is generally considered to be the aspect that can still best be mastered technically, **privacy** typically concerns the individual, and **security** is more of a concern for the state. It is not too seldom that one thing is harmed by what promotes the other. This is simply the usual game of cops and robbers. Spam, viruses, trojans, malware, denial of service attacks, but also simple eavesdropping—at ever shorter intervals we read what terrible things are happening. Most of it takes place in the digital world—attacks on the physical infrastructure, such as fiber optics, do not seem to be such a big issue yet. But it does not always have to be external attackers. There will also be a new undersea cable for scientific data exchange between Brazil and Portugal, because Latin America wants its research landscape to be connected to the European one, without having to use a connection node in Florida on US territory.

In the meantime, malware versioning works so frighteningly fast that countermeasures often miss their target, even if they are introduced immediately, because the next generation of rogue software is already scurrying through the network. The attackers are as diverse as the attacks—pranksters, the challenged, the frustrated, criminal individuals and organizations, terrorists, industrial spies, secret services, and the military. The revelations of Edward Snowden have confirmed that it is not only those who are, in the diction of the West, commonly classified as rogue states that do evil.

All of this creates nervousness, especially in view of the openness that science naturally needs for its existence. However, also in view of the often-negligent approach to security in science, let us be honest—who encrypts their emails by default? And that, too, is typical of the digital age: threats and rescue are carved from very similar cloth and spur each other on to top performance. Informatics finally must decide whether it wants to be part of the problem or part of the solution, it was said recently. You might see it that way, but there is nothing to decide. We are of course both.

7 Conclusion

The good old Alma Mater is going digital and changing its face in the process. This contribution wanted to shed light on a few of the many aspects that are relevant here. As far as the technical side is concerned, incentives have been and continue to be created—the **DFG** program "Centers of Excellence for Research Information" a good 15 years ago was an early example—and accents were set. But we are still a long way from a "digital university." Hence, the way to get there is not a purely technical one. Instead, a fundamental change is required. The university should also not change its essence and values, despite all the enthusiasm for renewal. That will be one of the great challenges of digital change.

Part V
Energy, Georesources and Material Technology

Economy 4.0: Aspects of a Future Energy System

Frank-Michael Baumann, Eckehard Büscher, Stefan Rabe, and Georg Unger

Contents

F.-M. Baumann (✉) · E. Büscher · S. Rabe · G. Unger
EnergieAgentur.NRW, Düsseldorf, Germany
e-mail: baumann@energieagentur.nrw; buescher@energieagentur.nrw; rabe@cef.nrw; unger@cef.nrw

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_39

1 Global Challenges of the Energy System

The current debates on the transformation of our global electricity generation system and on climate protection in science, politics, business, and society clearly show that the global energy system is facing a profound transformation. Not only global initiatives such as “Fridays for Future” demand compliance with the Paris climate protection agreement of the United Nations, but also the former American vice president and environmental activist Al Gore, e.g., in his films “An Inconvenient Truth” (2006) and “An Inconvenient Sequel: Truth to Power” (2017) for combating climate change. The breadth of the discussion is also evident in the encyclical “Laudato Si” by Pope Francis from 2015. “The climate is a common good of all for everyone. (. . .) Humanity is called to be aware of the need to make changes in life, production and consumption in order to combat warming or at least the human causes and climate change is one of the most important current challenges for humanity” (Der Heilige Stuhl 2015, p. 23). Further evidence of this is the societal discussions on the exit from nuclear power and coal, the expansion of renewable energies and a CO_2 tax in Germany.

In 2011, the German Advisory Council on Global Change (WBGU) presented three main areas of action for a global change in the direction of sustainability in its main report “World in Change - Social Contract for a Great Transformation”: energy systems, urban spaces, and land use systems (WBGU 2011). Furthermore, the authors conclude: “Normatively, the carbon-based economic model is also an unsustainable situation, as it endangers the climate system’s stability, and therefore the natural life support system for future generations” (WBGU 2011, p. 1).

1.1 Historical Review

First, a look at the development of today’s fossil-fueled global energy system. The nineteenth and twentieth centuries are considered the epoch of expansive modernism. Large parts of the world followed the industrial and growth path. A large part of the population, particularly in the industrialized countries, experienced material and, above all, immaterial progress (Loske 2016, p. 11). The basis for this progress was the apparently unlimited use of “cheap” fossil fuels with their welfare-enhancing effects. New means of transportation and communication set in motion a globalization of economy and society, the consequences of which humanity is only just beginning to understand. The sociologist Werner Sombart accordingly described the fossil fuels coal, oil, and gas as lottery winnings for mankind (Edenhofer 2017, p. 30).

Until the beginning of the Industrial Revolution, society’s energy supply was based on a few energy sources, especially biomass. In the course of industrialization, biomass and animal and human labor were gradually replaced by fossil fuels in combination with new technologies (steam engine, train, car, tractor). First coal was

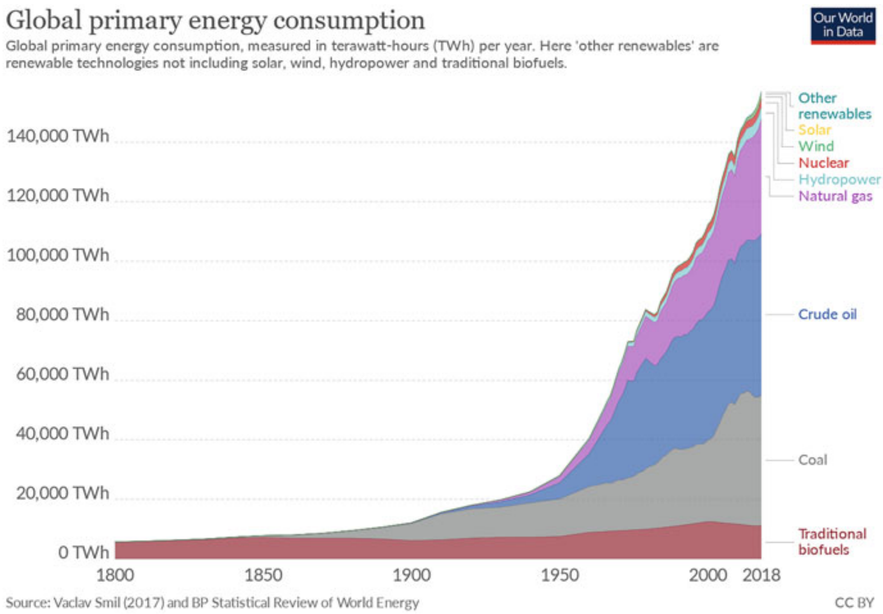


Fig. 1 Global primary energy consumption (Source: Our World in Data 2020a; Vaclav Smil 2017; BP Statistical Review of World Energy)

used, later oil and gas. Industrialization was therefore not only a complex process of economic and social restructuring of pre-industrial societies, but above all a change in the energy regime (WBGU 2011, pp. 92 ff.). This conversion of the energy supply system has led to corresponding welfare effects, especially in the industrialized nations (Fig. 1).

In the last four decades alone, global primary energy consumption has increased from more than 6 million tons of oil equivalents to almost 14 million, more than doubling. Overall, fossil fuels such as oil, coal, and gas accounted for 81% of primary energy supply (Bundeszentrale für politische Bildung 2019). This trend continued in 2018. Global energy demand increased in 2018 by 2.9%, while CO_2 emissions grew by 2.0%, faster than ever since 2010/2011 (BP 2019, p. 2).

The International Energy Agency (IEA) describes the situation in its World Energy Outlook 2019 as follows:

> The energy world is marked by a series of deep disparities. The gap between the promise of energy for all and the fact that almost one billion people still do not have access to electricity. The gap between the latest scientific evidence highlighting the need for evermore-rapid cuts in global greenhouse gas emissions and the data showing that energy related emissions hit another historic high in 2018. The gap between expectations of fast, renewables-driven energy transitions and the reality of today's energy systems in which reliance on fossil fuels remains stubbornly high. And the gap between the calm in well supplied oil markets and the lingering unease over geopolitical tensions and uncertainties (International Energy Agency 2019, S. 1).

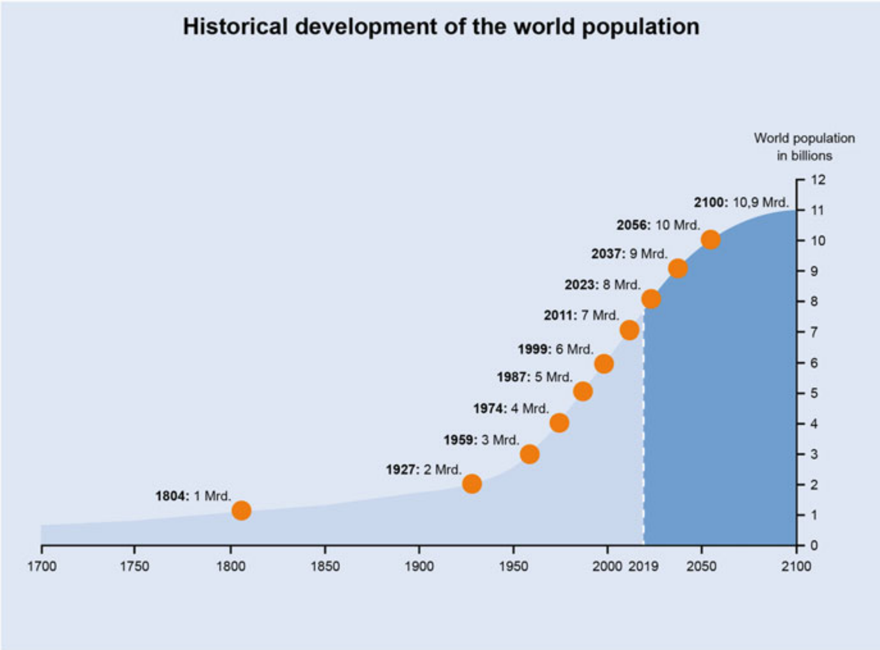


Fig. 2 Historical development of the world population (Data taken from: Deutsche Stiftung für Weltbevölkerung 2019)

However, this rapid increase in global energy consumption is not only based on the increasing industrialization of the world, but also has at least two other components. On the one hand, the world population grew from under 1 billion people around 1800 to almost 8 billion people in 2020 (Fig. 2).

On the other hand, the global gross domestic product multiplied from 1980 to 2018. The world's four largest economies alone, the United States, China, Japan, and Germany, account for more than half of global economic output with a gross domestic product of around \$42.9 trillion (Urmersbach 2020). As Fig. 3 shows, further growth in global gross domestic product can also be expected.

According to BP's scenarios, GDP will more than double again in the next 20 years, driven by increasing prosperity in the rapidly growing developing countries (BP Energy Outlook 2019, p. 7). However, historically speaking, economic growth has been very different in the industrialized, emerging, and developing countries over the past 200 years (Fig. 4).

These different economic growth trends including their external effects, such as CO_2 emissions are also the subject of controversial discussions at international level. In this context, it is often argued that the world's major emitters, such as China and the United States, would first have to reduce their greenhouse gas emissions. Germany had a share of around 2% of the global greenhouse gas emissions in

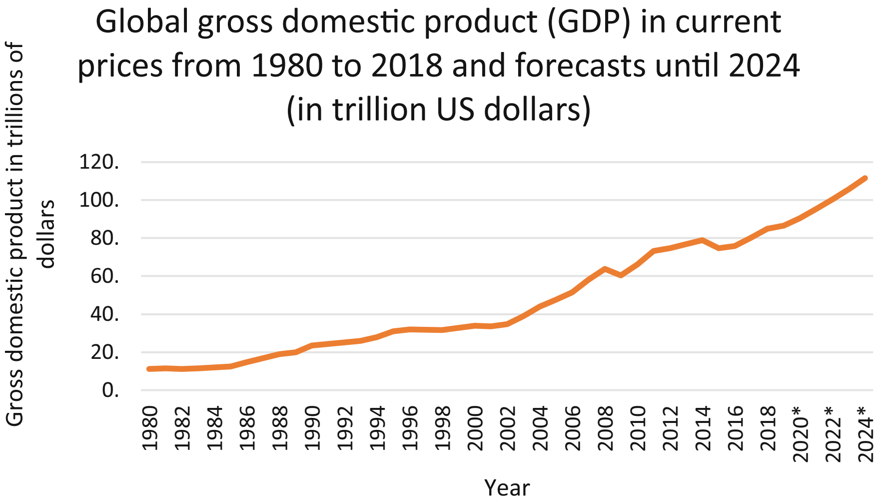


Fig. 3 Global gross domestic product (Data taken from: Statista 2020a, b)

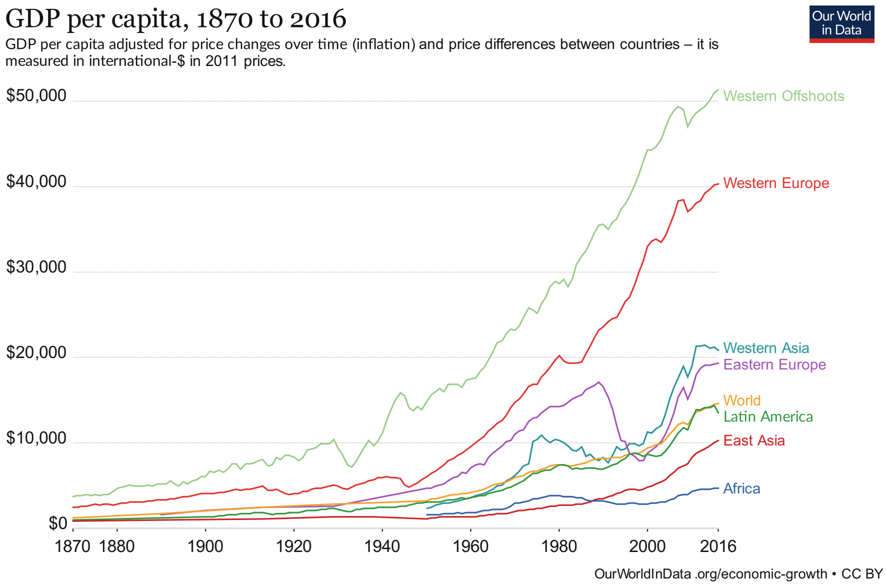


Fig. 4 GDP per capita. Note: These series are adjusted for price differences between countries using multiple benchmark years, and are therefore suitable for cross-country comparisons of income levels at different points in time (Source: Our World in Data 2020b; Maddison Project Database 2018)

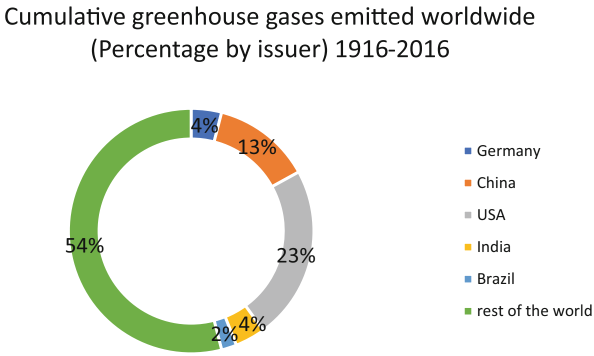


Fig. 5 Cumulative greenhouse gases emitted worldwide (Data taken from: Friedrich Ebert Stiftung 2019)

2016, which totaled 47.23 gigatons of CO_2 worldwide. China had a share of 26.9%, the USA 13.9%, India 6.1%, and Brazil 2.2%. Due to the longevity of carbon dioxide in the atmosphere, the cumulative values of the past decades are of crucial importance. If you look at the cumulative greenhouse gas emissions for the period 1916–2016, the situation changes as follows.

The USA had 23.2%, China 12.6%, Germany 4.3%, India 3.9%, and Brazil 1.7% (Friedrich Ebert Stiftung 2019, p. 9). These figures illustrate once again the responsibility of Germany and other industrialized nations in implementing the goals of the Paris decisions (Fig. 5).

At the same time as the industrialization of the world, the increasing global gross domestic product and the increase in the global population, the CO_2 concentration in the atmosphere has risen sharply. Before the Industrial Revolution, the CO_2 concentration was around 280 ppm, but today it has already exceeded 400 ppm, which is an increase of almost 50%. The so-called "Keeling Curve" also shows this long-term trend. At the beginning of the measurements in 1958, the Mauna Loa Observatory still determined an atmospheric CO_2 concentration of 317 ppm to increase steadily since (Schellnhuber 2015, p. 73). According to the NOAA Earth System Research Laboratories (ESRL), the value was already 414.02 ppm in December 2020 (NOAA Global Monitoring Laboratory 2021) (Fig. 6).

If one looks at these three determinants in parallel, the link between these developments becomes clear. With an increasing world population and growing global gross domestic product, global energy consumption and the resulting CO_2 emissions have increased at the same time (Fig. 7).

This also reveals the dilemma of our modern world society, i.e., the close connection of the growth paths has on the one hand led to corresponding prosperity, especially in the industrialized nations. On the other hand, this development has numerous negative effects on our environmental and climate system. With an increasing tendency, mankind now emits more than 36 billion tons of carbon dioxide into the atmosphere each year by burning fossil fuels. Because CO_2 is a long-lasting greenhouse gas and the earth's population produces far more of it than any other greenhouse gas, it is responsible for climate change like no other substance (Plöger and Böttcher 2016, p. 44). In this context, it also becomes clear that the planetary

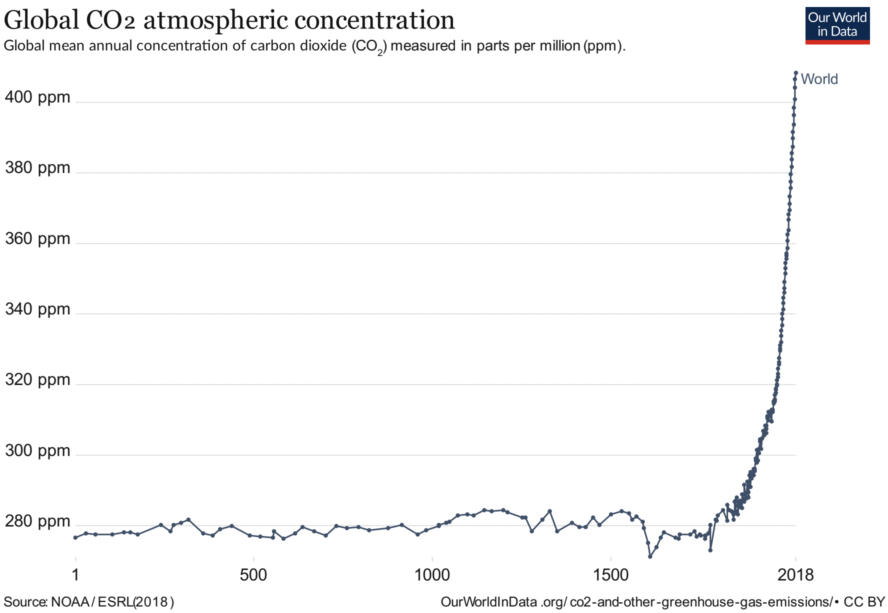


Fig. 6 Global CO_2 atmospheric concentration (Source: Our World in Data 2021; NOAA/ESRL 2018)

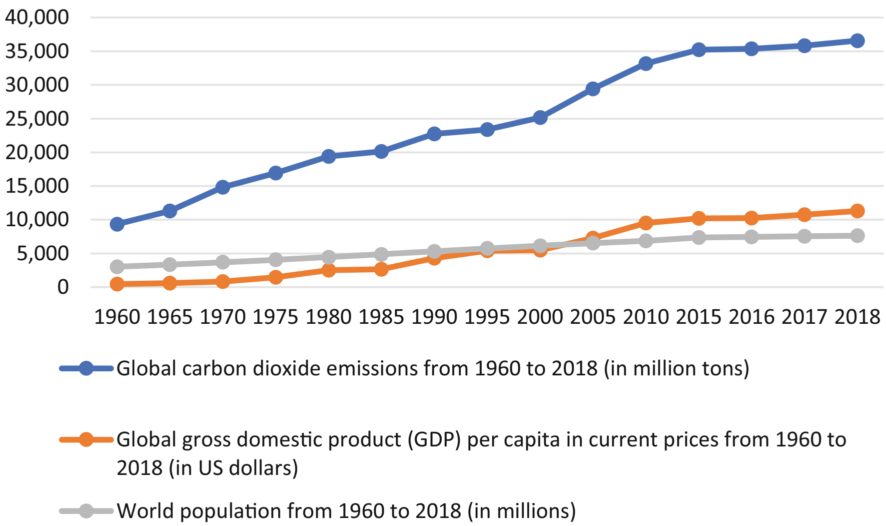


Fig. 7 Global development: CO_2 emissions, GDP, and global population (Data taken from Statista 2020a, b and United Nations 2019)

side rails would be breached with over 10 billion people at a high level of prosperity. Accordingly, the anthropogenic drivers of global environmental changes, such as climate change must be brought to a standstill (WBGU 2014, p. 4).

It is therefore crucial for the design of a global future energy system to decouple the three growth paths already described and the resulting CO_2 emissions, without foregoing socially relevant prosperity effects. Ahead of the Berlin Conference on Climate and Security 2020, Prof. Dr. Edenhofer, sums it up as follows: “Security and prosperity in the 21st century will depend to a large extent on how we manage global common goods, including climate stability” (Potsdam-Institut für Klimafolgenforschung 2020).

For example, the increasing incineration of fossil fuels, deforestation, land use and industrial processes contribute to the anthropogenic greenhouse gas effect and are already responsible for a temperature rise of approximately 1.0 °C global warming compared to the pre-industrial level, with a probable range from 0.8 °C to 1.2 °C (IPCC 2018). Despite the knowledge of mandatory action requirements (IPCC reports) to avoid risks of global warming by 1.5 °C compared to the pre-industrial level, such as sea level rise, extreme weather events, ocean acidification, health effects, migration and conflicts, costs of climate change etc. greenhouse gas emissions continue to increase worldwide.

In this context there is another determinant that is also essential, the development of energy efficiency and energy productivity. An energy transition cannot be achieved without increasing efficiency in all sectors. Energy and resource efficiency and a strengthening of the circular economy can counteract the global trends towards increasing energy consumption through economic growth and increased consumption. According to the dena Leitstudie “Integrierte Energiewende 2018,” e.g., the technical potential for energy efficiency through technical progress and digitalization in Germany is often well over 50% of energy consumption.

The so-called rebound effects should also come into focus. Otherwise, it can quickly happen that desired reduction targets in energy consumption are not achieved despite energy efficiency measures (Dena Leitstudie 2018, pp. 21 f.). For example, economic growth and energy consumption in Germany have been decoupled in the past. The German economy has grown by 50% since 1990, but greenhouse gas emissions have fallen by 30% in the same period. This trend is also evident in energy productivity, which has increased by over 70% since 1990 (BMU 2020, p. 50).

Again, a look back at the history of energy and climate discussions. Since the beginning of the 1970s, recognizable environmental problems such as the greenhouse effect and the ozone hole have been publicly discussed for the first time and the so-called neoclassical environmental and resource economy emerged, as well as ecological economy in the 1980s. In addition, the Club of Rome published its report “The Limits to Growth” in 1972 and the World Commission for Environment and Development published the “Brundlandt Report” in 1987. The international world conferences such as the Earth Summit in Rio de Janeiro also took place in these decades 1992, three “WMO World Climate Conferences” in Geneva in 1979, 1990, 2009 and the first climate conference in 1995 in Berlin or other annual climate conferences (Karlsruher Institut für Technologie 2017, pp. 4 ff.).

It can be stated paradoxically that on the one hand the international negotiations steadily increased and on the other hand the worldwide energy consumption and the energy-related emissions rose in parallel.

1.2 Germany's Energy Mix

A comparable development to the global use of fossil fuels outlined is also reflected in the development of the German energy supply system, which is still largely based on the use of conventional energies (Fig. 8).

Despite the steadily increasing expansion of renewable energies in Germany in recent years, primary energy consumption in 2019 was based on the use of 6.4% nuclear energy, 9.1% lignite, 8.8% hard coal, 14.8% renewables, 35.3% Mineral oil and 24.9% natural gas (AG Energiebilanzen 2020).

Nevertheless, it can also be stated that the share of renewable energies in primary energy generation in Germany has increased from 3.2% to 46.2% since 1990. In the first quarter of 2020, 72.3 billion kwh of electricity were generated in Germany for

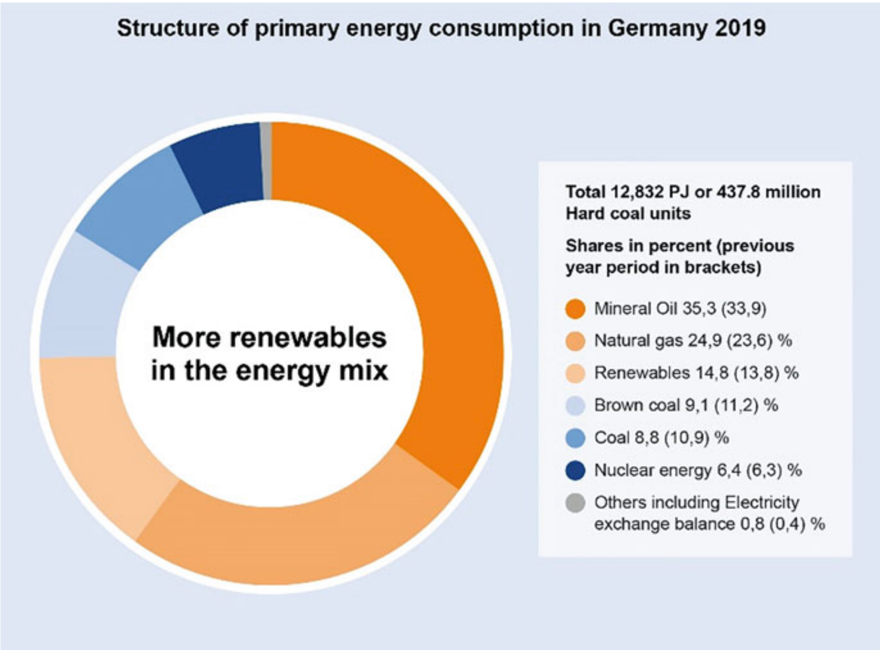


Fig. 8 Structure of primary energy consumption in Germany 2019 (Data taken from: AG Energiebilanzen 2020)

the first time from renewable energies than from conventional energy sources. According to the preliminary results of the Federal Statistical Office (Destatis), the amount of electricity from renewable energies rose to 51.2% of the total amount of electricity fed into the grid compared to the first quarter of 2019. The highest increase was recorded by electricity from wind power with +21.4% (Statistisches Bundesamt 2020).

1.3 Climate Protection Goals in Germany and the EU

Both the European Union and Germany have set themselves ambitious climate protection goals to shape the energy system of the future. In December 2019, the heads of state and government of the European Union committed themselves to the goal of climate neutrality by 2050. By 2050, all greenhouse gas emissions in the EU are to be avoided as far as possible. With the so-called Green Deal, the European Union is showing that it continues to play an international pioneering role in climate protection. At the end of 2020, the EU heads of state and government agreed to raise the EU climate target for 2030 to at least 55% compared to 1990. This means that internal EU greenhouse gas emissions will decrease by at least 55% by 2030 compared to 1990 levels. By 2050, Europe wants to become the first “climate-neutral continent” through an ambitious climate policy (Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit 2021).

There are also comparable developments in Germany. In June 2021, the German Bundestag passed a new Federal Climate Protection Act (KSG). With the revised law, the German greenhouse gas reduction target for 2030 will be increased to minus 65% compared to 1990. Greenhouse gases must be reduced by 88% by 2040 and greenhouse gas neutrality must be achieved by 2045. With the much more ambitious goals, Germany is the first EU country to implement the new European climate goals that were decided last year under the German Council Presidency (Bundesministerium für Wirtschaft und Energie 2021).

To achieve these goals, the so-called strategy paths of sustainability must be applied consistently, also to the energy sector:

- efficiency
- sufficiency
- consistency

To implement these strategic paths, first, consequent increases in efficiency in energy generation, second, changes in behavior and consumption, and third, the rapid expansion and integration of renewable energies into an energy system of the future. This transformation process must be flanked by the digitalization of the energy supply system and by accelerated sector coupling.

An essential prerequisite for a successful transformation is the development of the Levelized Cost of Electricity. The 2020 report by the International Renewable Energy Agency (IRENA) shows that technologies for generating renewable energy have become the most cost-effective option for new capacities in almost all parts of the world and have triggered a corresponding irreversible transformation process (IASS 2019b, p. 12).

> Electricity costs from renewables have fallen sharply over the past decade, driven by improving technologies, economies of scale, increasingly competitive supply chains and growing developer experience. As a result, renewable power generation technologies have become the least-cost option for new capacity in almost all parts of the world. This new reality has been increasingly reflected in deployment, with 2019 seeing renewables account for 72% of all new capacity additions worldwide. According to the latest cost data from IRENA, the global weighted-average levelized cost of electricity of utility-scale solar photovoltaics (PV) fell 82% between 2010 and 2019, while that of concentrating solar power (CSP) fell 47%, onshore wind by 39% and offshore wind at 29% (IRENA 2020, p. 12).

If one assumes that the Paris climate protection agreement, the EU goals, and Germany's energy policy goals have been reached, this means nothing less than that a complete global system transformation of the energy supply sector or an energy regime change must take place in the coming decades. There are at least three important reasons that make this international generation task seem achievable. First, the increasingly urgent voices from science, such as The IPCC reports show, secondly, the increasing global ecological pressure, especially in the emerging countries, such as air pollution and water scarcity, and third, the availability of real alternatives to substitute fossil fuels (Fücks 2016, p. 43).

2 Future Development of Electricity Generation and Demand in Germany

2.1 Development of Electricity Generation and Consumption in Germany

In 2020, more than half of the electricity in Germany is produced by wind, sun, or biomass. The emissions of climate-damaging greenhouse gases were reduced by around 36% by 2019 compared to 1990 (Agora Energiewende 2020a, b). The last three German nuclear power plants will go offline in 2022. This will be followed by an exit from coal-based electricity generation by 2038. Much has been achieved in the past two decades. But for the step into a new energy era to succeed the climate targets agreed at the climate conference in Paris to be achieved by 2050, many steps still have to follow. One measure alone will not be enough.

This includes, among other things, a significant increase in energy efficiency in all areas as well as completely CO_2-free electricity production. The study "Electricity System 2035+" by the Öko-Institut e.V. (2019) shows that the accelerated phase-out of electricity generation from coal is a key element in meeting climate protection goals. The German lignite and hard coal power plants currently cause a good 80% of the CO_2 emissions of the electricity sector. The exit from fossil power generation therefore makes a significant contribution to reducing CO_2. Of course, this cannot be done without structural change in the affected areas.

The experiences from energy transition of the last two decades have been enabled by feed in tariffs. So the costs of wind and solar systems could enormously be reduced in recent years. Current results of the tender in Germany show that the costs of onshore and offshore wind power and photovoltaic systems on open spaces have converged and are now only 3.5 to 7€ cents per kilowatt hour (BNetzA 2020). This means that the full costs of renewable energies are already lower than the costs for the construction and operation of conventional power plants. In the USA, Australia, China, Chile or Morocco, solar and wind power are now cheaper than nuclear and coal-fired power plants.

The basic feasibility of a power system based on more than 90% renewable energies is no longer in question. However, the challenge of the costs of this new power system has not yet been fully resolved (Fig. 9).

The expansion and conversion of the power infrastructure is the main task for long-term energy supply. The replacement of oil and gas in industry, trade, building heating, and mobility with renewable electricity or, in the future, also with "green" gases, shifts the energy consumption from fossil fuels to electricity. In its network development plan (NEP), the Federal Government assumes that the total electricity consumption remains constant at around 590 TWh/a as the consumption plus for traffic and heating is reduced by energy efficiency. The Federal Association for Renewable Energies (BEE) assumes, for example, that gross electricity consumption will be visible in 2030, relationships 740 TWh/a.

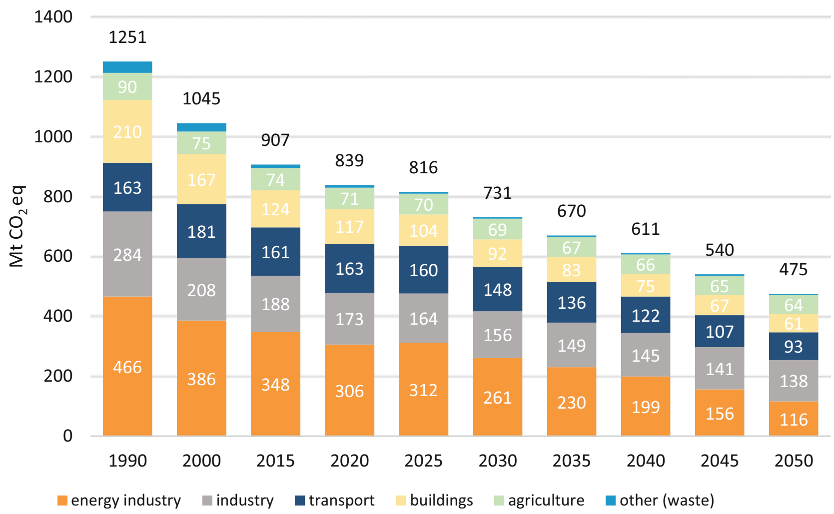


Fig. 9 Greenhouse gas emissions regarding each sector of the Climate Protection Plan (Data taken from Öko-Institut e.V. 2019) [Source: Data up to 2015 are based on German Environment Agency (UBA 2019); data from 2015 from Prognos 2020]

The more electricity that comes from fluctuating sources such as wind power and photovoltaics, the network infrastructure must be flexible. Power grids and storage systems need to be intelligently expanded as required to bring direct wind power from north to south and to compensate for the weaknesses of power generation between day and night or windy and windless times. This also plays an important role in the storage of electric energy when building up an intelligent electricity infrastructure. Digitization is the most important tool for intelligence in the infrastructure.

In the long term up to the year 2050, the pan-European supply situation, according to the Öko-Institut e.V. (2019), is characterized by the fact that Germany generates significantly more electricity than it consumes due to the strong addition of renewables. Therefore, Germany can become a net exporter. In the case of less ambitious European climate policy, the flexibility of the power plant park abroad is used to export excess energy abroad, especially in times of high wind supply. When the level of ambition increases abroad, net exports decrease significantly. To meet the need for flexibility, the use of storage (e.g., power-to-gas with recuperation or battery storage) is becoming increasingly important within Germany.

2.2 Perspectives from the Network Development Plan

The scenario framework of the network development plan of the German Federal Network Agency uses various assumptions regarding generation, load and consumption to describe how Germany's energy supply could develop in the future. It shows different possible development paths ("scenarios"). The Federal Network Agency approved the 2019–2030 scenario framework relevant for the 2019–2030 network development plan on June 2018 and confirmed the figures in Table 1 in June 2020. It contains a total of five scenarios. Three scenarios describe development paths until 2030, one scenario until 2035 and one scenario until 2025.

The individual scenarios contain different assumptions about how strongly and how quickly the energy landscape changes. The different assumptions are described in detail in the NEP. In all 2030 scenarios a percentage of the same but different expansion of the relevant renewable energy sources is assumed. The efficiency savings are also the same with—30 TWh/a in all 2030 scenarios.

The dimensions of the conventional power plant fleet decrease from scenario A2030 to scenario C2030 due to ever lower assumptions about the technical-economic operating life of the power plants. In contrast, the drivers of sector coupling and thus the net electricity consumption as well as the annual peak load continue to increase from scenario A2030 to scenario C2030.

Critics of the current network development plans point out that a part of the network expansion could be avoided by a decentralized energy transition. There are indications that a strongly decentralized energy turnaround could avoid or at least postpone individual elements of the planned network expansion. However, this would only be possible if we could agree socially to accept a very strong expansion of wind turbines near the centers of electricity consumption. So far, there has hardly

Table 1 2030 Network Development Plan (Data taken from BNetzA June 2020)

	Installed capacity [GW]				
Energy sources	Reference 2019	Scenario A 2035	Scenario B 2035	Scenario C 2035	Scenario B 2040
Nuclear energy	8.1	0.0	0.0	0.0	0.0
Lignite	20.9	7.8	0.0	0.0	0.0
Coal	22.6	0.0	0.0	0.0	0.0
Natural gas	30.0	38.1	42.4	46.7	42.4
Oil	4.4	1.3	1.3	1.3	1.1
Pump storage	9.8	10.2	10.2	10.2	10.2
Other conventional generation	4.3	3.8	3.8	3.8	3.7
Sum conventional generation	**100.1**	**61.2**	**57.7**	**62.0**	**57.4**
Onshore wind	53.3	81.5	86.8	90.9	88.8
Offshore wind	7.5	28.0	30.0	34.0	40.0
Photovoltaics	49.0	110.2	117.8	120.1	125.8
Biomass	8.3	6.8	7.5	8.7	8.2
Hydropower	4.8	5.6	5.6	5.6	5.6
Other renewable generation	1.3	1.3	1.3	1.3	1.3
Sum renewable generation	**124.2**	**233.4**	**249.0**	**260.6**	**269.7**
Generation in total	**224.3**	**294.6**	**306.7**	**322.6**	**327.1**
	Electricity consumption [TWh]				
Net electricity consumption	524.3[3]	603.4	621.5	651.5	653.2
Total electricity consumption	570.9[3]	639.8	656.9	686.9	688.6
	Driver for sector coupling				
Domestic heat pumps [million]	1.0	3.0	5.0	7.0	6.5
Electric mobility [million]	0.2	9.1	12.1	15.1	14.1
Power-to-heat (district heat/industry) [GW]	0.8[3]	4.0	6.0	8.0	7.0
Power-to-gas [GW]	<0.1[3]	3.5	5.5	8.5	10.5
	New storage and demand-side flexibility [GW]				
PV battery storage	0.6	11.0	14.1	16.8	14.9
Large-scale battery storage	0.4	3.6	3.8	3.8	3.8
DSM (industry, trade, commerce and services)	1.5[3]	4.0	5.0	8.0	7.0

been an honest and well-founded debate about whether there are realistic alternatives to parts of the network expansion.

2.3 *Regionalization of Production*

The performance of an existing generation plant can easily be assigned to a network node in the regionalization model. It is more difficult, for example, to predict the location and performance of wind farms or new conventional power plants that will be built by 2030 or 2035. Depending on the generation technologies used, different approaches are used for these forecasts. For the renewables, their regional potential and the regional use of land are considered. In addition, the grid connection applications submitted to the distribution system operators are evaluated and reliable information about the location and performance of the new renewable energy plants planned for the next few years is obtained.

For example, if the operator of a wind farm that is being planned requests a connection line with a capacity of several hundred megawatts, you can draw conclusions about how much electricity and where this wind farm will feed into the grid in the future. Based on the expected wind speeds and the solar radiation levels, the volatile generation time series for wind and solar energy can be calculated with the help of the system data (mast heights, simultaneity factor due to unfavourable alignment angles, shading effects). Also, plants whose production is less volatile, such as biomass or run-of-river power plants are considered.

For the regionalization of offshore wind energy, the information and the regional breakdown from the scenario framework are considered in the sense of power plant connection planning. The scenario framework considers the requirements of the area development plan for the invitation to tender for future areas in the North and Baltic Seas. It is therefore analyzed which connection line and how much power is transported to where on land. The grid connection points between the offshore connection line and the transition to the shore-side high-voltage grid are synchronized so that the feed-in of offshore wind energy is correctly transferred to the transmission grid on land in terms of space and quantity.

2.4 Regionalization of the Load

The load time series for the network development plan 2030 are determined based on a model-based approach. The study encompasses both the determination of demand for the whole of Germany and the assignment of demand to various inner German regions. First, the national demand for electricity in the sectors of households, industry, transport as well as trade, and services is modelled. The result is a change in the historical electricity demand curves of the individual sectors.

The total results in the overall national demand curve, which must meet the specified annual energy volume and considers the maximum annual load. This is distributed to the counties using sectoral distribution keys. In this step, numerous district-specific parameters such as the number of households, the regional population development, the number of employees per region and sector or climatic factors are considered.

For example, the domestic demand curve of the household sector will have a relatively strong impact on the regional demand curve of a region with high disposable income and high influx. To enable a calculation of the power flows in the transmission network, it is necessary in the last step to assign the rural districts to certain network nodes of the transmission network. The procedure creates regional load time series that are not based on historical profiles, but rather represent model-based forecasts. The resulting time series show corresponding differences to the load time series of previous network development plans. In the south of Germany in particular, the new modeling leads to increasing loads, while the loads in the eastern counties decrease in some cases.

Compared to today's electricity demand, the model predicts an increase in large parts of Bavaria and in parts of Baden-Württemberg in the household and manufacturing sectors. This is particularly due to the assumption of population growth in the corresponding regions, which is based on the spatial planning forecast 2035 of the Federal Institute for Building, Urban and Spatial Research. In addition, there is the assumption of a comparatively large increase in heat pumps in southern Germany. Their regionalization is mainly based on the distribution of single-family houses, which goes back to the regional database of the Federal Statistical Office. Since there are already many single-family houses in Bavaria and population growth is also assumed there, a large part of the forecast load is due to heat pumps. Especially in scenario C 2030, but also in scenarios B 2030 and B 2035, the heat pumps have a significant share of the total load. But without ambitious increases in efficiency, especially in the heating sector, the demand for electricity increases significantly.

For the first time, the 2019–2030 network development plan provides for the application of FBMC (Flow-Based-Market-Coupling) at all borders of Germany to neighbouring countries. This new approach is based both on the fact that FBMC is already being used at most borders or that it will be used in the next few years. The main difference to the use of NTCs (Net Transfer Capacities) is that available capacities are not specified between market areas, but on so-called "critical branches." The lines that are particularly heavily burdened by trade are defined as critical branches. The resulting power flows on these lines due to the amount of electricity exchanged by the trade must not exceed the specified capacity values. It is therefore a line-based specification of capacities for the trade.

2.5 Fundamental Considerations by the Federal Network Agency on Integrated Power Grid Planning

For some time now, various players in the energy market have been calling for so-called integrated network planning of the electricity and gas networks, sometimes coupled with the heat supply. The topic of sector coupling therefore requires joint planning to advance the integration of electricity and gas as a "hybrid technology-open" energy system. Because gas is increasingly no longer seen as a problem, but as part of the solution to long-term decarbonization of the energy industry.

In the past months and years, several reports have shown that "green" gas as biomethane, but above all as synthetic gas or as hydrogen, can contribute to an economically efficient decarbonization strategy. In most cases, however, the practical implementation of this integrated network planning is not explained in detail. It remains open whether an integrated calculation of the gas and electricity networks is meant in a common simulation process, whether it is a matter of carrying out a holistic energy system planning or whether the scenarios and thus only the input data of the electricity and gas network calculations should be standardized as far as possible.

An integrated network calculation is not carried out in the current system of the two processes, the scenario framework and the network development plan for electricity and gas, since the two simulation processes are fundamentally different. There is no trade restriction within the German electricity capacity for the electricity network calculation within the German electricity price zone. Assumptions are made about the installed services of the producers and consumers as well as the exchange capacities with the neighbouring countries and, based on electricity market simulations, grid usage cases are determined for all 8760 h of the target year to determine what the electricity grid has to do (in the future).

In all areas of the energy industry value chain, we believe that the opportunities of digitalization will be placed even more strongly than today:

- The digitalization of the energy industry ensures the efficient interaction of producers, consumers, and the grid.
- Customers benefit from new business models that can make Germany a pioneer in the fields of smart grids and smart homes.
- Data protection and data security are becoming increasingly important with the increasing digitalization of the energy industry.

3 Sector Coupling: Electrification of the Heat Supply and Transportation

3.1 Decarbonization of the Electrical System

The global climate change is a huge challenge for all countries in the world. The need for reducing the emission of greenhouse gases means that the energy economic sector as well as the industrial sector must reduce the use of fossil energy sources. The increasing usage of renewable energy sources (RES) is an important option.

Germany has subsequently increased the share of renewable energy sources. In 2019, the net electrical energy production from RES (i.e., wind, PV, biomass, and hydro power) was about 239 TWh corresponding to 46% of the total net energy production (Fraunhofer Institut für solare Energiesysteme 2020).

The increasing share of fluctuating renewable energy sources such as wind and solar power in the energy supply system is a challenge for the system stability. Energy supply and demand have to be balanced at every time with respect to system stability. This is much more complex in a decentralized energy system and requires more flexibility options especially on the demand side.

Storage Systems are important elements to achieve a sustainable and flexible energy supply system. A wide range of different storage and flexibility options are available which are able to support the energy market as well as the stability of energy grids.

Thermal storage and stationary power storage systems can provide efficient technical solutions with respect to the integration of renewable energy sources, the reduction of greenhouse gas emissions and sector coupling. They are therefore important technologies for the transformation of the energy supply system.

Other measures of flexibility are:

1. the extension of the electricity grids to balance the power supply and demand between regions
2. application of smart grids and smart markets to minimize the power transport
3. demand side management and demand respond applications in the economic sector trade, commerce and industry which allow a time shift of the power demand

During the last years, technologies which enable the connection of various energy sectors (i.e., power, gas, heat, mobility) have gained increased interest. These so-called "sector-coupling" concepts (i.e., power-to-X technologies, X = i.e., heat, gas, chemicals) offer several advantages in the framework of the energy transition. They have a large flexibility potential and can help to increase the reliability of the power system and they can help to avoid greenhouse gas emission by replacing conventional fuel by renewable energy sources.

Definition of Sector Coupling

A definition of sector coupling was given by the German Association of Energy and Water industries (BDEW 2017). According to this definition, sector coupling

> is the energy engineering and energy economy of the connection of electricity, heat, mobility and industrial processes, as well as their infrastructures, with the aim of decarbonization, while simultaneously increasing the flexibility of energy use in the sectors of industry and commercial/trade, households and transport under the premises of profitability, sustainability and security of supply (Robinius et al. 2017)

The European Commission understands sector coupling "as a strategy to provide greater flexibility to the energy system so that decarbonization can be achieved in a more cost-effective way" (Van Nuffel et al. 2018). The authors of the study also define two strategies for sector coupling:

- End-use sector coupling: this term is linked to the direct use of power, i.e., for heating purposes (heat pumps) or battery electric vehicles.
- Cross-vector coupling: this term refers to the integration of the electricity supply sector (mainly electricity) into the gas and industry sector (i.e., Power-to-X-technologies).

Sector coupling increases the flexibility of the energy system and can help to optimize the use of renewable energy sources. Furthermore, it has been pointed out that sector coupling could also reduce the costs of the transition into an energy system with high shares of renewable energy sources. A study for Germany revealed that the transition to an energy system with 80% GHG-reduction in 2050 could save

Table 2 Summary of the framework of sector coupling and the various technologies which can be applied

Energy supply	Energy conversion and infrastructure	Energy end-use
Conventional power plants Renewable energy sources (wind, solar, biomass, hydropower)	*Infrastructure* • Electricity grids • Gas grids (H_2 and CH_4) • CO_2—sources • Heating and cooling distribution networks *Technologies:* • Power-to-gas (H_2 and CH_4) • Gas-to-power • Power-to-chemicals and fuels (via Syngas: H_2 and CO) • Power-to-heat (heat pumps)	Transport Industry Commercial/ trade Households

Table 3 Overview mobility in Germany (2018) (Data taken from: Kraftfahrtbundesamt 2019 and Bundesministerium für Verkehr und Digitale Infrastruktur 2019)

Number of passenger cars in Germany	ICE		BEV	PHEV
47,095,784	Gasoline: 31,031,021 Diesel: 15,153,346		83,175	66,997
Fuel consumption				
Diesel (L)	20,633,000,000 (224 TWh)			
Gasoline (L)	26,083,000,000 (204 TWh)			

ICE internal combustion engine, BEV battery electric vehicle, PHEV plug-in hybrid electric vehicle

about 600 billion euro if a broad technology mix is applied compared to a system with a strong focus on electrification only (dena Leitstudie 2018) (Table 2).

Although sector coupling offers a great potential for flexibility, there are several barriers to overcome. An important techno-economical obstacle is that several sector coupling technologies (i.e., power-to-gas) can actually not compete with conventional technologies (i.e., use of natural gas). The policy and regulatory framework can also act as a barrier. For Germany, it has been pointed out that the actual energy market design including the tax system is not suitable for the implementation of sector coupling instruments. A level-playing field including carbon dioxide price signals is needed to foster the development and implementation of new technologies in an efficient manner (BDEW 2017).

3.2 E-Mobility

The development of the future mobility will strongly affect the energy supply system during the next decades. An overview of the actual mobility status is shown in Table 3. About 47 Mio passenger cars were registered at the end of 2018 in Germany corresponding to an accumulated driving distance of 630.8 billion km (Kraftfahrtbundesamt 2018). The mobility is almost completely based on passenger cars equipped with combustion engines. The share of electric vehicles in the market (battery and plug-in hybrids) is currently small (ca. 0.3%) but fast increasing. Consequently, the consumption of fossil-based fuels is quite high: in 2018, 46.7 billion litres of fuel were burned in passenger cars. Thus, the mobility and transport sector contributes significantly to the total GHG-Emissions of Germany (2018: 163.6 Mio t, 19%).

Perspectives
In 2019, about 7.9 Mio electric vehicles were sold on the worldwide market (Zentrum für Sonnenenergie- und Wasserstoff-Forschung Baden-Württemberg 2020). The largest market is China (3.8 Mio). It is expected that the number of electric vehicles will strongly increase: one million electric vehicles will drive on German roads by 2022 as forecasted by the National Platform for Electromobility at the end of 2018.

Battery electric vehicles are very energy efficient and can contribute to the transformation of the mobility sector if renewable energy sources are used for power generation. This can be illustrated by an example: assuming an electricity consumption of 15–20 kWh/100 km, a total electricity amount of 102–136 TWh/a would be required to cover the cumulated driving distance in Germany 2018 (630.8 billion km) with battery electric cars. This is much less than the imported fossil energy amount of gasoline and diesel (428 TWh/a, Table 3).

Conclusions
The ramping-up of the E-Mobility impacts the energy system on several levels:

- The introduction of many battery electric vehicle into the mobility sector will change the energy supply. The capacity of the power generation by renewable energy sources must be largely increased to meet the electrical demand. In contrast, the demand for fossil fuels will decrease.
- The electric infrastructure must be developed including a suitable charging infrastructure.
- Battery electric vehicles will increase the load in the distribution grids which must therefore be extended. In this context, the introduction of intelligent charging technologies offers a great potential for cost reduction and an interesting market for public utilities.
- Electric vehicles offer flexibility options for grid and energy services.

3.3 Decarbonization of Industrial Processes: Hydrogen Economy

A considerable amount of GHG-emissions origins from industrial processes. In the European Union in 2018, petroleum refinery accounted for 111 Mio t (2.8%) CO_2-

equivalents, iron and steel industry for 98 Mio t (2.3%), and chemicals for 70 Mio t (1.7%) (European Environment Agency 2020). To meet the climate goals, industrial processes need to be modified.

Hydrogen is a very flexible energy carrier. It can be used for electricity and heat generation as well as a feedstock for synthetic fuels and chemicals. Hydrogen is one option to store energy in large amounts and for a longer timescale. Thus, hydrogen plays a key role in sector coupling technologies. Furthermore, hydrogen offers opportunities for the decarbonization of industrial processes, i.e., in the chemical or in the steel industry (Van Nuffel et al. 2018; Dechema 2017). Hydrogen can therefore be regarded as a key vector in future energy systems with high shares of renewable energy sources. Consequently, several European countries set up national hydrogen strategies and roadmaps, i.e., the Netherlands (Government of the Netherlands 2020) and Germany (BMWi Die Bundesregierung 2020). Some European regions are working on hydrogen strategies and roadmaps as well (Ministerium für Wirtschaft, Innovation, Digitalisierung und Energie des Landes Nordrhein-Westfalen 2019).

The main goals of the German hydrogen strategy can be summarized as follows:

- Decreasing the costs for hydrogen technologies (i.e., electrolysers) to make them competitive to other technologies. To this end, a hydrogen market shall be developed. Market incentives can help to enable the application of hydrogen technologies on a larger scale.
- Development of a suitable hydrogen transport and distribution infrastructure.
- Establish hydrogen as a base material in industry (i.e., chemistry, steel production) to enable the transition to a GHG-neutral industrial production.
- Introduction of hydrogen as an energy carrier for a clean mobility with a focus on sectors which are difficult to electrify (i.e., shipping, aviation, heavy duty transport, etc.).
- Promotion of research and development of hydrogen technologies comprising the generation, transport and use of hydrogen. This includes the education of skilled personnel.

The proposed large-scale hydrogen production would have a strong impact on the energy system of the future. The Federal Government of Germany expects a hydrogen consumption of 90–110 TWh/a in 2030. The goal is to cover a part of this demand via electrolysis. About 5 GW electrolyser capacity is targeted. This corresponds to a total amount of ca. 20 TWh/a renewable generated power. The decarbonization of the steel industry until 2050 would require about 80 TWh/a hydrogen. These numbers illustrate that Germany's power consumption will be significantly increased even if only a part of the needed green hydrogen is produced from renewable energy sources. A hydrogen economy will have a large impact on the power transport and distribution grids as well.

Fig. 10 Drivers of change

4 Digitization and Cyber Security

4.1 Introduction

In an energy system that intelligently combines electricity, heat and mobility, the digitization megatrend is already of outstanding importance. Digitization will further expand its key role in the transformation of the energy supply system. It is considered as an irreversible development process with social and economic opportunities. In an increasingly decentralized system, new complexity arises at various points that can only be mastered with a high degree of automation. Accordingly, digitization offers great market and change potential along the entire value chain up to the potential of disruptive innovations that can revolutionize entire markets in the short term. Technologies such as artificial intelligence, blockchain or the Internet of Things, already recognizable today, will have a major impact on the design not only of the energy industry. However, many of these technologies are still in the making and will only become effective in the future.

Figure 10 shows that digitization is an important but not the only necessary step for the energy transition. But digitization is more than just installing smart meters. In all four subject areas, digital processes play the most important role in the change process as a tool

- climate protection
- profitableness
- service security
- acceptance

Digitization is designed to make the system more efficient overall and to open up new technologies and business models to the energy market. This applies particularly to the areas of virtual power plants, smart grids, digital electricity trading, intelligent load management, sector coupling, and the control of consumption for the end customer.

4.2 Digitization for the Energy Transition

The transformation of the energy industry into a system with integrated, renewable generation units is characterized by inherent structural changes. To do justice to the growing complexity of this system, various solutions have to be established. Among other things, advancing digitization offers an opportunity to make tomorrow's complex systems controllable through automation. As a result, industries are increasingly complementing each other, and information and communication technology (ICT) is becoming the key technology for implementing the energy transition.

The need to make tomorrow's complex systems manageable, controllable and efficient through automation requires the digitization of the entire energy industry value chain. This is illustrated by the rapid increase of generation plants in the system—from 500 conventional large power plants in the end of the century to 1.8 million largely decentralized plants. The system transformation can only be achieved through a structural transformation among energy suppliers and other market players. However, digitization as the core element of this structural change is often still perceived as a threat. Many of the previously known challenges concern the collection, processing, analysis, use and security of data.

It remains to be seen whether new data can really be developed from the amount of data to be expected or whether the market players will eventually become a big data cemetery. In general, however, there are much more opportunities than risks. There are numerous synergy effects and possibilities for intelligent networking, for controlling fluctuating feed-in and for services with end customers as well as for ensuring security of supply. In an increasingly complex system, information and communication technology is becoming a key technology for implementing the energy transition. Overall, digitization offers many opportunities for new business models and better energy efficiency:

- use of new information and communication media
- automation of previously analog processes along the value chain
- measurement, analysis, and control of complex systems in the energy sector
- acquisition of all consumption and generation data in real time

- use of the data as a raw material source to derive regularities and irregularities on the customer side
- creation of market transparency
- efficient coordination of all market players and units

4.3 Digitization and Virtual Power Plants

Due to the expansion of renewable energies, there is a large number of small power generation plants. They are spread across the country and only generate electricity when the sun is shining, or the wind is blowing. The energy transition therefore poses major challenges for electricity suppliers and network operators. Intelligent software makes it possible to efficiently market renewable energies, make optimal use of the grid or bundle many small plants. It can also be used to optimize the own consumption of locally generated green electricity. Another option is to control the energy requirements of flexible consumers depending on the load on the networks (Fig. 11).

The simplest variant of a virtual power plant can only consist of generating plants. According to a general definition approach, virtual power plants are often characterized by the fact that they combine spatially separate generation and consumption systems and storage systems and market them together. The core element is the interconnection of a large number of mostly decentralized systems via a central communication infrastructure.

The core of every virtual power plant is the central system control using optimization software, which activates or regulates certain systems in the portfolio based on market signals. Market signals ideally represent how the situation in the networks

Fig. 11 Virtual power plants compensate fluctuations in electricity generation (© Next Kraftwerke GmbH)

is presented. If many renewable energies feed in, the prices drop and there is an incentive to draw electricity from the grid and to stabilize it. If renewable energies are unable to meet the demand for electricity, prices rise and there is an incentive to feed in additional electricity or reduce consumption. These measures are intended to prevent the grid from being under-fed. The exchange of the necessary signals and the use of the generation, consumption or storage systems takes place in the virtual power plant.

4.4 Smart Grids: Optimized Network Operation Through Digitization

The term smart grid stands for an intelligent energy supply system in which modern information and communication technologies (ICT) are used to coordinate the networking and control of all consumers. In addition to the electrical cables themselves, these modern network structures incorporate suitable storage systems, intelligent switching and control elements, controllable transformers, and intelligent measuring devices (smart meters) for consumers to provide them with information and incentives for efficient energy consumption.

Knowledge of the current network status is required to implement automated power control. To do this, the currently unmonitored distribution network must first be equipped with appropriate sensors. However, to not lose the economic advantage of network automation, the sensor equipment should be such that sufficient calculation accuracy can be achieved with the smallest possible number of sensors. If an impermissible operating state is determined, countermeasures are taken automatically by sending control commands counteracting the state violation to the actuators. For this purpose, a three-stage control algorithm was developed at the University of Wuppertal, the aim of which is to remedy the state violation that has occurred with the least possible effects thanks to digitized processes.

4.5 Digital Energy Trading and Blockchain

The changes in the European energy markets in recent years have also placed energy trading companies under pressure to innovate. Low wholesale margins, the continuous expansion of renewable energies and increased cross-market liquidity have increased the complexity and competitive pressure and require corresponding rationalization processes from energy trading that only work with digitized processes. Efficiency increases and the flexibility and speed of energy trading are particularly important for successfully dealing with the new market situation. The market for automated algorithm-based energy trading systems is therefore growing correspondingly quickly. These enable the fully automatic conclusion of trading transactions based on previously defined strategies. Digitization simplifies and optimizes the

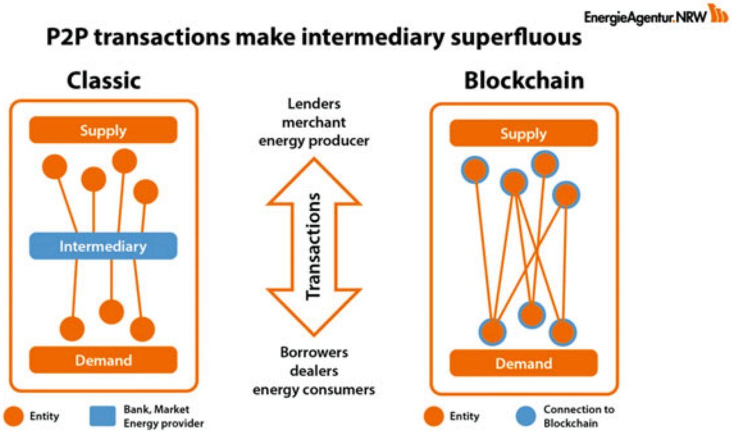


Fig. 12 The potential needlessness of the intermediary (Data taken from Verbraucherzentrale NRW 2016)

daily trading business for individual employees from decision-making to execution, including logistics.

Blockchain technology is a shared database technology in which the consumer and supplier of a transaction are linked directly to one another. The "block chain" is a chain of transaction blocks. It creates a digital register that records transactions between a consumer and a supplier. The resulting online network is managed by several computers—the participants in the transaction. Before a transaction can take place, it must be confirmed and encrypted from every computer to guarantee the security of the transaction (Fig. 12).

The potential of blockchain must be very differentiated and considered in relation to the possible application. Public blockchains, e.g., to produce Bitcoins have the disadvantage of a very high-power consumption. In the case of the energy industry application, data is not stored centrally, but for every participant in the consortium—but not decentrally for all those involved in a product. Here, the problem of excessive electricity consumption does not arise.

The Wuppertaler Stadtwerke (WSW) have developed an application example with its green electricity product "Tal.Markt." "Tal.Markt," the world's first blockchain-based trading center, has been activated for green electricity since the beginning of 2018 and was expanded in 2019 with several partners from Germany and Austria due to the good results.

4.6 Cyber Security

However, the many opportunities and possibilities of digitization also face risks, such as Cyberattacks and data misuse. Not only the example of the successful attack on the Ukrainian power grid at the end of 2015, but also an increasing number of cases reported from Germany show that critical infrastructures can also be attacked successfully.

A basis of our coexistence—even in health crises and Corona-times—is a secure electricity and water supply that citizens, hospitals, trade, and industry can rely on. In addition to the smooth functioning of the hardware, this also includes secure software that is protected against external interference. In digitization, there are great opportunities as well as risks in internal company processes as well as in external digitization for the development of new business models. Increasing networking and permanent communication options within all stages of the value chain in the energy industry, from power generation to distribution and smart meters in every household, create millions of potential gates of entry for uninvited guests.

Cyber Security's job is to guard these gateways and close them if necessary. In the meantime, a remarkable service market has developed here, which is booming with sales of over € 6 billion and double-digit growth rates. This is not least due to the rapidly growing number and scope of the attacks. An impressive visualization of the seriousness of the cyber threat can be found on the website "*Sicherheitstacho.eu.*" Here, the cyberattacks on Telekom's infrastructure and its regional origin are shown in real time on a world map—thousands per minute, over 10 million per day.

The number and intensity of events for the energy industry have also increased in recent years: in 2010 the computer worm "Stuxnet" made a name for itself, primarily aimed at power plants in Iran. Large-scale sabotage attacks became known in 2014, particularly on gas power plants, pipelines, and wind turbines, particularly in Spain. In 2015, there was a hacker attack on the power supply in Ukraine, resulting in a power failure. Around 700,000 Ukrainians were without electricity over Christmas. In 2017, only an accidental safety shutdown of a power plant in Saudi Arabia saved physical destruction from an explosion triggered by hackers. EnBW speaks of 1000 attacks on its infrastructure every day and, according to its own statements, employs 200 people to counter these attacks.

In NRW, too, there are increasing cases where customer data from municipal utilities and city administrations are blocked by digital extortionists and only released again after the ransom has been paid. In May 2020, cybercriminals managed to collect sensitive data from a southern German municipal utility—including addresses and customer account details. These appeared a little later in the darknet. The energy supplier had previously refused to pay ransom in the tens of millions. So far, the attacks have been limited to customer data and not network operation. At the end of 2019, however, a case was reported from Bavaria where the BKA (Federal Criminal Police Office) informed the managing director of a municipal utility that Russian-language worms were hidden in its network operation, which could become active at any time. At this point it becomes clear that the question of cyber security does not only concern customer data and thus the liquidity and credibility of a local municipal utility, but also national security via the power supply.

As part of the federal government's cyber defense strategy, the BKA, the BSI, the BND (Federal Intelligence Service), the BfV (Federal Office for the Protection of the Constitution) and the police authorities of the federal states have been working together in the National Cyber Defense Center in Bonn since 2011. All messages are registered and evaluated here. The observations of the BSI (Federal Office for

Information Security) relativize the picture of tens of thousands of attacks on the entire infrastructure. The 2019 BSI management report describes a total of 38 cases of reportable attacks on energy supply (29) and water supply (9). The BSI report continues to speak of a high level with regard to the risk situation in the area of critical infrastructures. For 2019, however, no threats can be identified that exist only for critical infrastructures.

The most important tools against cybercrime appear:

- good preparation
- clear distribution of competences
- training of all employees

In addition to training and tests for dealing with mail attachments, etc., there are also training offers for network operators. In the Cyber Security Center in Essen, Germany, for example, the operating teams of power grids are attacked digitally under real conditions and their defense behavior is trained and analyzed.

In conclusion—despite certain risks of new digital territory—it should be pointed out that the Federal Ministry for Economic Affairs and Energy already stated in 2015 in its fourth monitoring report on the energy transition that the electricity market will be one of the first fully digitized sectors of the German economy.

5 The Importance of Socio-Economic Aspects for the Transformation of the Energy Supply System

Global megatrends such as digitization, demographic change, the overuse of natural resources and global climate protection are fundamentally changing our society and business practices. The transformation of the global energy supply system can be understood as a complex interplay of social, technological, economic, and political innovation processes (FVEE 2018, p. 51).

A look at the German energy supply system shows that more and more new energy generation technologies such as wind turbines, solar and biomass plants, and new power lines are spreading and shaping new energy landscapes, as they can change them considerably at times. The spread of these systems does not only stand for a few additional artifacts of industrial technology, but also for a comprehensive transformation of the energy supply system including follow-up effects. This made Germany a leader of the energy transition for several years and thus become a buzzword for a technical as well as socially-economic change at the beginning of the twenty-first century (Holstenkamp and Radke 2018). This development illustrates the departure into the so-called “post-fossil age” and initiates the path towards a greenhouse gas-neutral economy. It covers all technological, economic, cultural, and ecological levels fundamentally and changes them in the long term (Schneidewind 2018) (Fig. 13).

This transformation process is currently shaping the global discussion on climate protection and presents political, economic, and social challenges. This is especially

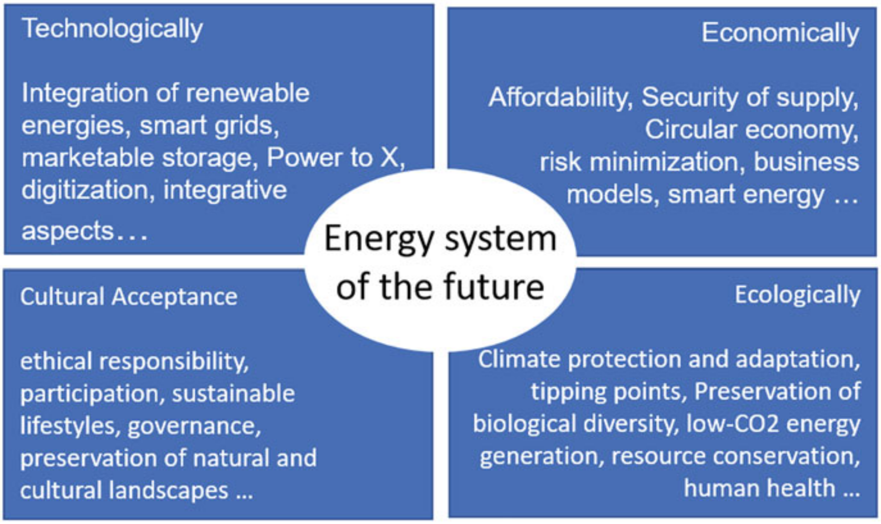


Fig. 13 Energy system of the future

true for Germany as a strong industrial location. In addition to the political measures to implement the Paris climate protection agreement, science, business, and civil society play a key role in shaping the energy system of the future. After all, every actor has to make his or her specific contribution towards sustainable social/industrial metabolism. However, this wide variety of actors with different interests also leads to corresponding conflicts of interest and goals.

With the energy transition, an irreversible paradigm or energy regime change has been initiated in the energy supply, which intervenes deeply in existing structures and ways of life in our society and requires new approaches. Only if the energy transition is understood as a project for society as a whole, which includes technology development, organizational change, effective control measures and targeted behavior, a successful implementation can take place, since the energy transition is a complex system transformation (Renn 2015, p. 76).

In this respect, system changes, if they are to be designed successfully, require broad social discourses on the design options and limits, that is, an open and fair discussion about the pros and cons of different development lines. A social acceptance of the transformation process can only be achieved and sustained if all actors are involved in the design process, combined with active participation in the restructuring of the energy system (Fischedick et al. 2014b). According to the "Social Sustainability Barometer of the Energy Transition 2018," on the one hand, more than 90% of people living in Germany are in favor of the "Community Work Energy Transition." On the other hand, the population's judgment of their implementation in terms of costs, political management, proximity to citizens and justice, as well as in the overall balance sheet, is significantly more negative. For example, a majority consider the energy transition to be too expensive and chaotic (IASS 2019a, b, pp. 2 ff.).

The approval of the population is, however, of significant importance for the rapid expansion of renewable energies, since the further expansion of wind farms and large solar parks in particular will lead to drastic changes in the landscape of entire regions. The energy transition can only be implemented with and not against civil society. This requires a constructive, honest dialogue culture and priority for fact-based discussions and decisions. All in all, a broad socio-political discourse is needed on the right goals and suitable implementation paths for the energy transition. In addition, future-oriented forms of governance are required as well as diverse concepts for participation in the implementation process (Lietzmann 2014).

Distributive fairness (costs and benefits) as well as procedural fairness (participation and empowerment) are relevant for the German energy transition. Topics such as electricity costs, distribution effects of subsidies and participation formats such as real laboratories, planning cells or energy cooperatives play a decisive role here (Virtuelles Institut "Transformation – Energiewende NRW" 2019, p. 6). A profound, innovation-driven transformation of the energy system cannot take place without acceptance and early involvement of the citizens (ENavi 2018, p. 6).

It should also be noted that the technical energy transition was preceded by an international change in values that rejects the harmful environmental effects of a fossil-atomic energy system and prefers a sustainable system. This means that the energy transition is not only based on political decisions but is being driven by large sections of the population, e.g., the worldwide "Fridays for Future" movement or the high approval rate of the population for the implementation of the energy transition are impressive evidence.

The scientific monitoring of the energy transition requires a systematic approach that requires increased interdisciplinary cooperation between scientists from various disciplines and an increasing coupling of scientific processes with political and social decision-making (Fig. 14). In addition, the non-scientific actors must also be involved to develop feasible and promising recommendations for action (FVEE 2018, p. 51).

As described in the Sect. 4 digitization also plays a special role in the transformation. While it has already shaped many areas of everyday and professional life, the energy transition is still at the beginning of a digital transformation. On the one hand, digitization offers the potential to manage the feed-in from a large number of renewable energy systems and to couple different sectors with the energy system. On the other hand, it presents the actors with the challenge of developing new skills and ways of thinking. This applies equally to production and consumption (Virtuelles Institut "Transformation – Energiewende NRW" 2019, p. 4).

In this context, the link between technological and socio-economic research plays a key role. As the evaluation of numerous studies shows, e.g., the German climate protection goal is technically and economically achievable, largely based on known technologies. At the same time, climate-friendly technologies still have to be developed or brought to market maturity. In addition, through intensive research efforts and new approaches to research and innovation, it is important to tackle the technical and social aspects of the energy transition, especially in acceptance and transformation research, e.g., to further investigate the understanding between socio-economic and sociological-technical relationships (BMU 2016, p. 27; FVEE 2018, p. 52).

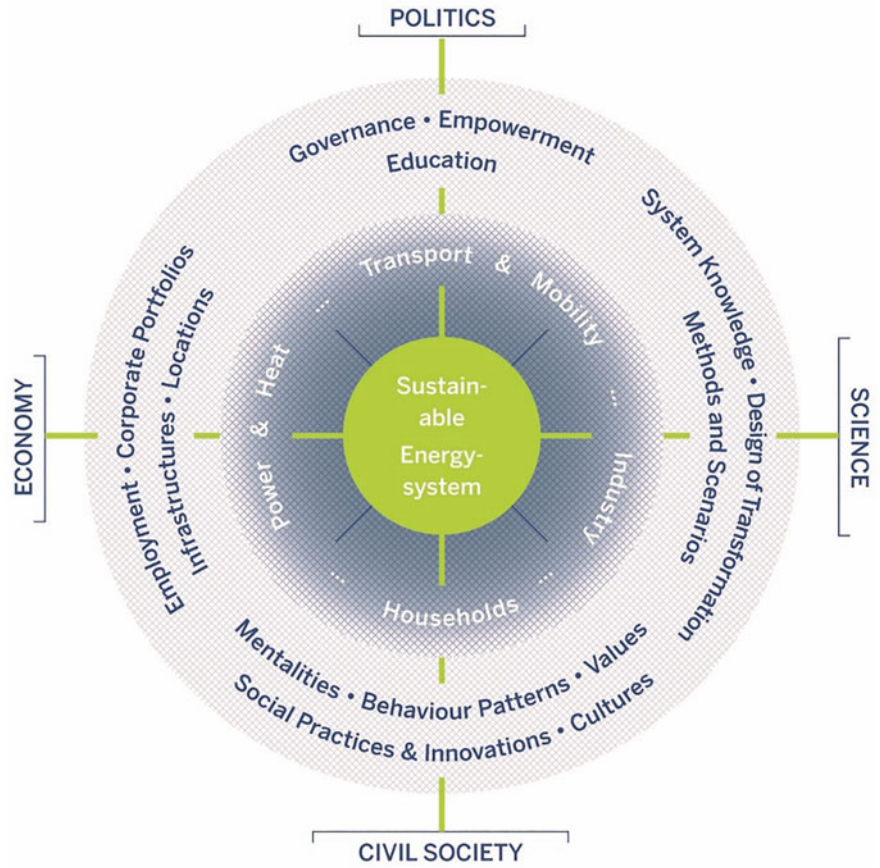


Fig. 14 Essential perspectives and requirements of society on the common path of the energy transition (Source: Fischedick et al. 2014a)

In this sense, social or cultural sciences are not a decorative accessory, responsible for the after-dinner speech, but an integral part of a systemically integrated research landscape that provides politicians and society with the necessary knowledge bases to make them economically justifiable and ecologically compatible in the sense of sustainability and to develop socially adapted energy strategies for an energy system of the future (acatech 2011, p. 37).

6 Conclusion

In conclusion, it can be stated that the design of a future energy system presents politics, business, science, and society with a great challenge, but this appears to be solvable in an international and ambitious cooperation. Accordingly, the international community must account for future generations to what extent it can and wants to limit future climate change (Edenhofer 2017, pp. 7 ff.). The decarbonization of our economic system is therefore one of the major global challenges of the twenty-first century.

References

acatech (2011) Die Bedeutung der Gesellschafts- und Kulturwissenschaften für eine integrierte und systemisch ausgerichtete Energieforschung. Berlin

AG Energiebilanzen (2020) Arbeitsgemeinschaft Energiebilanzen legt Bericht zum Energieverbrauch 2019 vor. Pressedienst Nr. 02/20

Agora Energiewende (2020a) Auswirkungen der Corona-Krise auf die Klimabilanz Deutschlands – Eine Abschätzung der Emissionen 2020. https://www.agora-energiewende.de/veroeffentlichungen/auswirkungen-der-corona-krise-auf-die-klimabilanz-deutschlands/. Zuggegriffen am 06.07.2020

Agora Energiewende (2020b) Die Energiewende im Stromsektor: Stand der Dinge 2019. Berlin. https://www.agora-energiewende.de/veroeffentlichungen/die-energiewende-im-stromsektor-stand-der-dinge-2019/. Zugegriffen am02.07.2020

BDEW (2017) Positionspapier—10 Thesen zur Sektorkopplung. Berlin

BMWi Die Bundesregierung (2020) Die nationale Wasserstoffstrategie. 2020

BNetzA, Bundesnetzagentur, Network Development Plan 2030 (2020)

BP (2019) Statistical review of world energy, 68th edn. https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2019-full-report.pdf. Zugegriffen am 08.06.2020

Bundesministerium für Umwelt, Naturschutz, Bau und Reaktorsicherheit (2016) Klimaschutzplan 2050. Klimaschutzpolitische Grundsätze und Ziele der Bundesregierung. Berlin

Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit (2020) Klimaschutz in Zahlen. Fakten, Trends und Impulse deutscher Klimapolitik. Berlin

Bundesministerium für Verkehr und Digitale Infrastruktur (2019) Report "Verkehr in Zahlen 2019/2020". Berlin

Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit (2021) EU Klimapolitik. Berlin. https://www.bmu.de/themen/klimaschutz-anpassung/klimaschutz/eu-klimapolitik. Zugegriffen am 06.10.2021

Bundesministerium für Wirtschaft und Energie (2021) Deutsche Klimaschutzpolitik. Verbindlicher Klimaschutz durch das Bundes-Klimaschutzgesetz. Berlin. https://www.bmwi.de/Redaktion/DE/Artikel/Industrie/klimaschutz-deutsche-klimaschutzpolitik.html. Zugegriffen am 06.10.2021

Bundeszentrale für politische Bildung (2019) Primärenergie-Versorgung. https://www.bpb.de/nachschlagen/zahlen-und-fakten/globalisierung/52741/primaerenergie-versorgung. Zugegriffen am 07.07.2020

Dechema (2017) Low carbon energy and feedstock for the European chemical industry. Frankfurt

Dena Leitstudie (2018) "Integrierte Energiewende 2018". Berlin

Der Heilige Stuhl (2015) ENZYKLIKA LAUDATO SI VON PAPST FRANZIKUS ÜBER DIE SORGE FÜR DAS GEMEINSAME HAUS. Rom. http://www.vatican.va/content/francesco/de/encyclicals/documents/papa-francesco_20150524_enciclica-laudato-si.html. Zugegriffen am 04.06.2020

Deutsche Stiftung für Weltbevölkerung (2019) Historische Entwicklung der Weltbevölkerung. https://www.dsw.org/infografiken. Zugegriffen am 07.07.2020

Edenhofer O (2017) Klimapolitik. Ziele, Konflikte, Lösungen. München

ENavi (2018) Energiewende braucht den Dialog mit der Gesellschaft. Potsdam (Geschäftsstelle des Kopernikus-Projektes Energiewende-Navigationssystem)

European Environment Agency (2020) Annual European Union greenhouse gas inventory 1990–2018 and inventory report 2020. https://www.eea.europa.eu/publications/european-union-greenhouse-gas-inventory-2020

Fischedick M et al (2014a) Essential perspectives and requirements of society on the common path of the energy transition. https://www.vi-transformation.de/begleitprojekt. Zugegriffen am 29.01.2021

Fischedick M et al (2014b) Herausforderung der Transformation des Energieversorgungssystems aus nationaler Sicht – Phasen der deutschen Energiesystemtransformation. In: Cluster EnergieForschung.NRW (Hrsg) Transformationsforschung NRW. Wege in ein nachhaltiges Energieversorgungssystem, Düsseldorf. www.cef.nrw.de

Fraunhofer Institut für Solare Energiesysteme (2020) Energiecharts. www.energy-charts.de. Zugegriffen am 10.06.2020

Friedrich Ebert Stiftung (Hrsg) (2019) Die Debatte um den Klimaschutz Mythen, Fakten, Argumente. Bonn

Fücks R (2016) Freiwillige grüne Revolution, in: IP Zeitschrift für internationale Politik, März/ April 2016, Seite 42-45

FVEE (2018) Gemeinsam forschen für die Energie der Zukunft. Berlin

Government of the Netherlands (2020) Government strategy on hydrogen. https://www.government.nl/documents/publications/2020/04/06/government-strategy-on-hydrogen. Zugegriffen am 24.06.2020

Holstenkamp L, Radke J (Hrsg) (2018) Handbuch Energiewende und Partizipation. Wiesbaden

IASS (2019a) Soziales Nachhaltigkeitsbarometer der Energiewende 2018. Potsdam

IASS (2019b) Stärkung der internationalen Zusammenarbeit für eine globale Energiewende. Potsdam. IASS Policy Brief 3/2019

International Energy Agency (2019) World Energy Outlook 2019. Executive summary. https://webstore.iea.org/download/summary/2467?fileName=1.English-Summary-WEO2019.pdf. Zugegriffen am 08.06.2020

IPCC (2018) Summary for Policymakers. In: Global Warming of 1.5 °C. An IPCC Special Report on the impacts of global warming of 1.5 °C above preindustrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. Geneva. https://www.ipcc.ch/site/assets/uploads/2019/03/SR1.5-SPM_de_barrierefrei-2.pdf. Zuggriffen am 09.07.2020

IRENA (2020) Renewable Power Generation Costs in 2019. International Renewable Energy Agency. Abu Dhabi

Karlsruher Institut für Technologie (2017) Die Weltklimakonferenzen. Karlsruhe

Kraftfahrtbundesamt (2018) Entwicklung der Fahrleistungen nach Fahrzeugarten seit 2014. Flensburg. (short report). https://www.kba.de/DE/Statistik/Kraftverkehr/VerkehrKilometer/vk_inlaenderfahrleistung/vk_archiv/2014/verkehr_in_kilometern_kurzbericht_pdf.pdf?__blob=publicationFile&v=1. Zugegriffen am 01.07.2020

Kraftfahrtbundesamt (2019) Fahrzeugzulassungen (FZ): Bestand an Kraftfahrzeugen nach Umwelt-Merkmalen. Flensburg

Lietzmann HJ (2014) Politische Partizipation und Verantwortungsteilhabe an der Energiewende: Ein soziales Ereignis. In: Cluster EnergieForschung NRW (Hrsg) Transformationsforschung NRW. Wege in ein nachhaltiges Energieversorgungssystem. Düsseldorf

Loske RU (2016) Politik der Zukunftsfähigkeit. Konturen einer Nachhaltigkeitswende. Frankfurt am Main

Ministerium für Wirtschaft, Innovation, Digitalisierung und Energie des Landes Nordrhein-Westfalen (2019) Wasserstoffstudie Nordrhein-Westfalen. https://www.wirtschaft.nrw/sites/default/files/asset/document/bericht_wasserstoffstudie_nrw-2019-04-09_komp.pdf. Zugegriffen 09.07.2020

NOAA Global Monitoring Laboratory (2021) Atmoshperic CO_2 at Mauna Loa Observatory. https://www.esrl.noaa.gov/gmd/ccgg/trends/mlo.html. Zugegriffen am 28.01.2021
Öko-Institut e.V (2019) Modellbasierte Szenarienuntersuchung der Entwicklungen im deutschen Stromsystem unter Berücksichtigung des europäischen Kontexts bis 2050. Freiburg, Bremen, Berlin. https://www.oeko.de/fileadmin/oekodoc/BMWi-FKZ_03ET4031-Endbericht.pdf. Zugriffen 24.06.2020
Our world in data (2020a) Global primary energy consumption. https://ourworldindata.org/grapher/global-primary-energy. Zugegriffen am 15.06.2020
Our World in Data (2020b) GDP per capita. https://ourworldindata.org/grapher/average-real-gdp-per-capita-across-countries-and-regions. Zugegriffen am 07.06.2020
Our World in Data (2021) Global CO_2 atmospheric concentration. https://ourworldindata.org/grapher/global-co-concentration-ppm. Zugegriffen am 28.01.2021
Plöger S, Böttcher F (2016) Klimafakten. Bonn (Bundeszentrale für politische Bildung)
Potsdam-Institut für Klimafolgenforschung (2020) Berliner Konferenz zu Klima und Sicherheit gibt Startschuss für neue Risikoanalyse. Pressemitteilung vom 23.06.2020
Renn O (2015) Aspekte der Energiewende aus sozialwissenschaftlicher Perspektive. München
Robinius M et al (2017) Linking the power and transport sectors—Part 1: The principle of sector coupling. Energies 10:956. https://doi.org/10.3390/en10070956
Schellnhuber HJ (2015) SELBSTVERBRENNUNG. Die fatale Dreiecksbeziehung zwischen Klima, Mensch und Kohlenstoff. München
Schneidewind U (2018) Die Große Transformation: Eine Einführung in die Kunst gesellschaftlichen Wandels. Fischer, Frankfurt am Main
Statista (2020a) Weltweites Bruttoinlandsprodukt. https://de.statista.com/statistik/daten/studie/159798/umfrage/entwicklung-des-bip-bruttoinlandsprodukt-weltweit. Zugegriffen am 07.07.2020
Statista (2020b) Weltweiter CO2-Ausstoß in den Jahren 1960 bis 2018. https://de.statista.com/statistik/daten/studie/37187/umfrage/der-weltweite-co2-ausstoss-seit-1751/. Zugegriffen am 07.07.2020
Statistisches Bundesamt (2020) Stromerzeugung im 1. Quartal 2020: Windkraft erstmals wichtigster Energieträger. Wiesbaden. Pressemitteilung Nr. 189 vom 28. Mai 2020
United Nations (2019) World Population Prospects 2019. New York. https://population.un.org/wpp/Download/Standard/Population/. Zugegriffen am 08.06.2020
Urmersbach B (2020) Weltweites Bruttoinlandsprodukt (BIP) bis 2024. https://de.statista.com/statistik/daten/studie/159798/umfrage/entwicklung-des-bip-bruttoinlandsprodukt-weltweit/. Zugegriffen am 07.06.2020
Van Nuffel L et al (2018) Study sector coupling: how can it be enhanced in the EU to foster grid stability and decarbonise? Policy Department for Economic, Scientific and Quality of Life Policies, Directorate-General for Internal Policies. http://www.europarl.europa.eu/supporting-analyses. Zugegriffen am 01.07.2020
Verbraucherzentrale NRW (2016) Blockchain. Kurzstudie, Düsseldorf
Virtuelles Institut "Transformation – Energiewende NRW" (2019) Ergebnisse der Gemeindestudie zum Thema "Energiewende und globale Megatrends in NRW". Wuppertal. Stuttgart. Schriftenreihe Energiesysteme der Zukunft
WBGU (2011) Welt im Wandel. Gesellschaftsvertrag für eine große Transformation. Berlin
WBGU (2014) Zivilisatorischer Fortschritt innerhalb planetarischer Leitplanken. Berlin. Politikpapier 8
Zentrum für Sonnenenergie- und Wasserstoff-Forschung Baden-Württemberg (2020) Tabelle "Bestand an Elektro-PKW weltweit". https://www.zsw-bw.de/mediathek/datenservice#c8590. Zugegriffen am 06.07.2020

Mining 4.0

Elisabeth Clausen, Aarti Sörensen, and Karl Nienhaus

Contents

1 Introduction

Mining is indispensable for economic growth and welfare. Despite efforts to reduce material intensity in advanced economies through increased **resource efficiency** and **recycling**, primary resource extraction is going to remain central to economic development and growth. In fact, an increasing number of **minerals and metals** play a crucial role not only in high-tech products but also in making the shift to a greener, **low-carbon economy** possible.

According to the European Environment Agency (EEA 2015), the total world economic output is projected to triple in the period from 2010 to 2050. Population growth in conjunction with economic development, a growing global middle class and a gradual convergence of living standards are the drivers behind the rapid increase of **global demand for raw materials**. The use of **metals** is projected to grow the fastest, from 7 Gt in 2011 to 19 Gt per year until 2060. The rapid increase in the demand for metals is mostly driven by Brazil, Russia, India, Indonesia, China, South Africa (BRIICS), as well as other developing countries (OECD 2019).

E. Clausen (✉) · A. Sörensen · K. Nienhaus
Institute for Advanced Mining Technologies, RWTH Aachen University, Aachen, Germany
e-mail: eclausen@amt.rwth-aachen.de; asoerensen@amt.rwth-aachen.de; knienhaus@amt.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_40

There are various efforts and attempts to decouple economic growth from **resource consumption** (OECD 2019), e.g., by promoting **recycling** or increasing material and **resource efficiency**. However, even improved recycling rates may not replace increased mining efforts in the near future. Consequently, mining, defined as the industry that extracts primary resources from the earth, will remain central to feeding the demand of global growth in the foreseeable future (RDB 2018). This is especially true since the growth projections for primary **metals** are not only accounting for structural changes in the advancing economies but also for improving living standards.

Mining as an industry includes a wide range of activities that encompass both artisanal small-scale mining activities as well as large and highly industrialized operations. Worldwide, around 15 million people are directly engaged in small-scale mining activities using simple and mostly manual methods with another 100 million people indirectly depending on those activities to sustain their livelihoods (Wotruba 2004). However, this chapter focuses exclusively on industrial mining operations that are more widely and acutely affected by the transformation towards **Mining 4.0**.

2 What Is Mining 4.0?

The term "**Mining 4.0,**" often used synonymously with "**Smart Mining,**" refers to the application and integration of digital technologies and automation in industrial mining operations.

At the Institute for Advanced Mining Technologies (AMT) of RWTH Aachen University, we define **Smart Mining** as "the meaningful connection and integration of mining machines, processes and other pysical objects with information-processing objects, processes and **communication** technologies to form so-called cyber-physical systems, where the exchange and transmission of data and information takes place via a platform, the IIoT (Industrial Internet of Things)." The intelligent mine of the future thus represents the long-term vision of a **digitally connected, autonomous mine** in which the connected systems are able to reduce the ever-increasing complexity to such an extent that improved decision-making can be realized in real time. (In this context, "real time" is to be understood as "right time," i.e., the information must be available "in time" for the process, i.e., not necessarily in milliseconds depending on the process). The mines of the future will therefore not only be digitally integrated, but also flexible and selective as well as dynamically adaptable, robust and reliable (Sörensen et al. 2020).

In terms of infrastructure, the main components of an intelligent mine include:

- Automated equipment. For instance, excavators and dump trucks, shearer, and conveyors, drilling equipment, crushers, bunkers, skips, etc.;
- Hardware, such as sensors, RFID tags, wireless infrastructure, drones, embedded systems;

- Software, such as cloud and platform solutions, data analytics solutions, 3D imaging and modeling software, and remote management solutions.

The “IT core” of most of these applications are **robust and reliable sensors** that collect data and, in combination with a secure data transmission and **communication** infrastructure and sophisticated algorithms and software programs, transform them into meaningful information—ideally in real time.

In addition, new technologies such as **battery-powered electric vehicles** (BEVs), the integration of **renewable energy** sources or even on-site 3D printing can help to make the operation of a mine smarter (Clausen and Sörensen 2021).

3 Drivers and Objectives for Mining 4.0

While the global demand for **minerals and metals** is projected to increase steadily in the coming years, mining companies around the world continue to be under pressure to be cost-competitive and to generate returns for shareholders, responding to the changing demands of a wide range of stakeholders but also to changing technical and economic conditions.

On the one hand, declining **ore grades**, increasing **mining depths** and mining in **remote regions** are leading to longer, energy-intensive production distances. The demand for **water and energy** is increasing and, as a result, so are the overall production costs. It has been observed for several decades that overall productivity has been declining and has only recently stabilized or, in some areas, improved. These trends are also leading to longer periods for the development of new mining projects (Durrant-whyte et al. 2015).

On the other hand, the values of stakeholders and social expectations are changing. (End-)consumers increasingly demand that companies ensure that products come from traceable sources and are produced in a responsible manner, that they offer fair and secure jobs to their employees, and that they protect the environment and support the communities in which they operate. Investors are increasingly judging companies on factors such as their **environmental footprint**, **carbon footprint**, **greenhouse gas emissions** and **energy** consumption, as well as their safety record and the benefits to employees (Burns 2020). In summary, there is a growing demand for companies to align their performance with the triple bottom line of **sustainability**, measuring and evaluating their social, environmental and economic performance (Burns 2020).

This means, mining companies need to ensure and further improve **occupational safety and health** while minimizing the use of resources, such as operating resources, capital, **energy**, **water** or **land use** and thus minimizing the use of and damage to the environment. At the same time, they have to increase efficiency and productivity while ensuring a high degree of **flexibility**, **selectivity** and **adaptability** of extraction activities due to changing framework and deposit conditions.

To meet future demands for raw materials, new deposits that are not yet technically and/or economically mineable today, need to be made commercially available through the development and deployment of advanced technologies. **Digitalization and automation**, it can be said with certainty, are important keys to realizing these objectives and meet the current and future demands of markets and stakeholders.

After productivity has been declining for several decades, the last couple of years confirmed that digital technologies and automation offer the potential to deliver significant improvements by improving the quality and availability of information and overall operational efficiency, which in turn can be the key to significant productivity gains (Lala et al. 2016). **Sensor-based applications** can be used to monitor **key performance parameters of machines and equipment** in real time and improve **failure prediction**, while at the same time enabling the monitoring and control of process parameters such as optimized **material flow**, best possible **equipment utilization**, improved processing, optimized **ventilation on demand** and improved monitoring of tailing dams.

In addition, digital technologies can help to reduce **environmental impacts** while automation can significantly improve health and safety of workers. It is therefore assumed that digital technologies and automation have the potential to be the key to the **sustainability** of the industry. McKinsey estimates a potential economic impact of introducing **Mining 4.0** at about 370 billion dollars worldwide by 2025, which would account for 17% of the industry's projected cost base globally. By far the greatest economic impact is seen in the area of operations management. In addition to a better understanding of the deposit and the **real-time monitoring** of process and machine data, this particularly includes improved logistics and increasing mechanization of extraction through automation (Durrant-Whyte 2015).

4 Challenges of Mining 4.0: How Is It Different from Industry 4.0?

As a peculiarity of the industry, mining companies have always had to deal with a high degree of uncertainty and imperfect information in their operations, regardless of where in the world they operate.

In mining, more than in most other industries, operations always have to deal with a relative uncertainty and imperfection of information because in a mining operation, both the location and the conditions of the production activity change continuously. In addition, despite thorough exploration and **modeling** of the deposit, knowledge about the actual structure and condition of the deposit is always limited and, consequently, information about the deposit is imperfect and uncertain until the time of the actual extraction activity. The oftentimes daily changes of the conditions for the extraction have major implications for downstream processing (Clausen et al. 2020).

Furthermore, the general conditions for mining companies can be extremely challenging as equipment and workers are exposed to extreme conditions such as dust, dirt, difficult lighting conditions, sometimes confined spaces or vastness, potentially falling rocks and uneven ground, vibration, and shock loads, possibly extreme temperatures (ranging from -20 °C to 80 °C) as well as potentially explosive atmospheres. In addition, hazardous situations can lead to safety-related power cuts, so that the systems used must be insensitive to interruptions in the power supply (Clausen et al. 2020).

The extreme conditions for production can be further exacerbated by the location of the deposit. Since the production activity of a mining company is determined by the geographical location of the deposit, it means that mining operations can be several hundreds of kilometers away from urban structures, access to **water**, **energy**, **communication,** or skilled labor.

In contrast to an industrial manufacturing operation, where the end product is usually made up of a large number of individual parts that are processed, manufactured, and assembled in the course of the process, this process is exactly the opposite in a mining operation, where the end product is separated from the whole. This implies that the mineral deposit is continuously reduced until fully extracted. While the lifetime of a mining operation depends on a variety of factors, such as size of the deposit, technology, economic parameters (including operating costs, market price for the raw material), and/or political and social conditions, the fact that any operation has a finite life span is always immanent to mining (Clausen et al. 2020).

In addition to those challenges immanent to the mining environment, there are particular technical challenges, especially in underground mining, as it relates to the implementation of automated **localization** and **navigation** solutions in underground environments. For example, GPS signals and other traditional localization services do not work below the surface so that alternative solutions must be developed in the form of local localization and positioning systems. Furthermore, the **interoperability** of machines and systems is still a challenge in open pit and underground mining. However, it is an important key to integrating machines and processes throughout the mine and the entire value chain. At present, this is still difficult to implement due to a lack of standardization. In a concerted effort, an international group of mining experts is working to establish an Open Platform Communications Unified Architecture (OPC UA) for mining (Hartmann et al. 2021).

These characteristics of mining represent a considerable challenge for a mine and its operation and make mining operations a particularly difficult and challenging environment for implementing **Mining 4.0**. Considering that mining is, on top of all the previously mentioned factors, also notoriously cyclical, with volatile prices for metals and minerals and cyclical investment patterns as a result, it seems more obvious why mining as an industry has had a history of longer lead times compared to other industries. However, over the last few years, the pace has picked up for mining companies and innovation in analytics, mobile solutions, and automation have become more widely deployed, providing significant efficiency gains (Mori et al. 2018).

5 What the Future Holds

Mining 4.0 is under way in all parts of the world across the industrialized mining sector and it will continue to transform the industry significantly in the coming years. This transformation is accelerated by the exponentially increasing amount of data as a result of deploying more sensors and decreasing data storage costs, the increases in computational power and the development of new analytic tools and methods enabling advanced data analytics, mobile technology contributing to more effective field-force management and streamlining of processes as well as increasing levels of automation (Mori et al. 2018).

Yet, the industry has a way to go to realize the vision of digitally connected and autonomous operations that are integrated and optimized along the entire value chain from pit to port. One area that can help to further develop digitally supported autonomous systems will be the increasing use of **artificial intelligence** (AI), **machine learning**, **robotics-based process automation**, **sophisticated systems analysis**, and **modeling**. This will enable operators and machines to collect, analyze, evaluate and understand the data and thus to develop a situation awareness, and to gain insights into the overall processes in near real-time (in time) and determine which possible courses of action need to be considered. The complexity and harsh conditions, especially in underground mining, require further technological developments, supported by application-oriented research, especially with regard to the development of autonomous systems underground. While the localization of personnel and equipment in some mines has been realized with the help of Wi-Fi networks, the underground autonomous **localization**, positioning and **navigation** of machines as well as machine-to-machine **communication** systems still require research and innovation (Clausen und Sörensen 2021).

Another area in which forward-looking research into new methods and technologies can contribute is the further development of **selective and low-impact mining** methods to increase **resource efficiency** while reducing the amount of overburden or waste material produced during extraction. Much progress has been made in primary processing with the aim of consuming less **energy**, **water,** and chemicals while at the same time increasing the recovery of valuable material. In terms of selective extraction, advances in real-time material characterization (e.g., during the conveying process and/or prior to processing) can further contribute to reducing the amount of waste material to be processed, thereby further optimizing resource efficiency and energy consumption.

While research and innovation play an important role in making the mining industry fit for the future, if the benefits of **smart mining** are to be fully realized, they must be accompanied by a clear vision for reducing the environmental impact of operations while improving productivity and safety (Sörensen et al. 2020). In addition, strong operational performance has been shown to be linked to occupational safety and health and employee engagement. Thus, how companies manage their people changes the way those people manage their equipment and, ultimately, the productivity of those assets. Consequently, mining companies need to pay attention not only to technology improvements but to employee engagement to reap the full benefits of **Mining 4.0** (Flesher et al. 2018). In addition, it is likely that the threat of climate change and difficulties in **water** supply will lead to even more rapid transformations in the way modern mining is conducted.

The bottom line, one can say with certainty, is that digitalization will change the benchmark of value creation in the industry. In future, data will gain increased value compared to the physical product, and the top value creators will be those that effectively capture and leverage data. The stakes are high: McKinsey research suggests "metal producers that harness the full potential of a digital transformation can increase their EBITDA margins by up to 6 to 8 percentage points." (Mori et al. 2018)

6 Conclusion

Mining companies today are challenged to change, adapt, and innovate to remain competitive and economically viable in the future. At the same time, moving towards more responsible and sustainable business practices is now considered essential for mining businesses to survive in the long term (Ellis 2021).

Digitalization and automation, the cornerstones of **Mining 4.0**, are considered important keys to make mining safer, more efficient and productive and more socially acceptable by lowering the environmental impact. However, success is not guaranteed by simply introducing new technology. Being smart about Mining 4.0 includes developing a clear vision and roadmap for transforming the business and taking the people along by engaging them from the onset. People, technology, and the environment need to be considered equal. Only then can technology fulfill its full potential for improving people's lives while contributing to safeguarding our planet for generations to come.

References

Burns D (2020) Thinking ahead (Mining Magazine). Online verfügbar unter https://www.miningmagazine.com/innovation/news/1383087/thinking-ahead. zuletzt aktualisiert am 17.03.2020, zuletzt geprüft am 17.08.2021

Clausen E, Sörensen A (2021) The future of mining is smart. In: VDMA Supplement "Best of Germany"

Clausen E, Nienhaus K, Bartnitzki T, Baltes R (2020) Bergbau 4.0. In: Walter Frenz (Hg.): Handbuch Industrie 4.0: Recht, Technik, Gesellschaft. Heidelberg: Springer, S. 919–937.

Durrant-Whyte H (2015) How digital innovation can improve mining productivity. Unter Mitarbeit von Ryan Geraghty, Ferran Pujol, and Richard Sellschop. Hg. v. McKinsey&Company. Online verfügbar unter https://www.mckinsey.com/industries/metals-and-mining/our-insights/how-digital-innovation-can-improve-mining-productivity, zuletzt geprüft am 19.02.2019

Durrant-whyte H, Geraghty R, Pujol F, Sellschop R (2015) How digital innovation can improve mining productivity. In: McKinsey & Company, 01.11.2015. Online verfügbar unter https://

www.mckinsey.com/industries/metals-and-mining/our-insights/how-digital-innovation-can-improve-mining-productivity, zuletzt geprüft am 17.08.2021

EEA (2015) The European environment—state and outlook 2015: Synthesis report: Publications Office of the European Union Luxembourg. In: European Environment Agency, Copenhagen

Ellis I (2021) Why sustainability will be a key issue to the mining industry's future. Online verfügbar unter https://www.nsenergybusiness.com/features/sustainability-mining, zuletzt aktualisiert am 17.08.2021, zuletzt geprüft am 17.08.2021

Flesher N, Parsons J, Rai PK, Thibert J (2018) The missing link in mining-company productivity: people. In: McKinsey & Company, 23.05.2018. Online verfügbar unter https://www.mckinsey.com/business-functions/operations/our-insights/the-missing-link-in-mining-company-productivity, zuletzt geprüft am 17.08.2021

Hartmann T, Lehmann J, Mavroudis F (2021) Automatisierung und Digitalisierung im Bergbau – OPC UA als Schlüsseltechnologie zur Interoperabilität. GeoResources 2:42–44. Online verfügbar unter https://publications.rwth-aachen.de/record/820246

Lala A, Moyo M, Rehbach S, Sellschop R (2016) Productivity at the mine face: pointing the way forward. In: McKinsey & Company, 04.08.2016. Online verfügbar unter https://www.mckinsey.com/industries/metals-and-mining/our-insights/productivity-at-the-mine-face-pointing-the-way-forward, zuletzt geprüft am 17.08.2021

Mori L, Saleh T, Sellschop R, van Hoey M (2018) Unlocking the digital opportunity in metals. In: McKinsey & Company, 25.01.2018. Online verfügbar unter https://www.mckinsey.com/industries/metals-and-mining/our-insights/unlocking-the-digital-opportunity-in-metals, zuletzt geprüft am 17.08.2021

OECD (2019) Global Material Resources Outlook to 2060: OECD. Online verfügbar unter https://www.oecd.org/environment/waste/highlights-global-material-resources-outlook-to-2060.pdf, zuletzt geprüft am 17.08.2021

RDB (2018) Rohstoffgewinnung und Ressourceneffizienz – ein Widerspruch? *bergbau* 10:456–457

Sörensen A, Nienhaus K, Clausen E (2020) Smart mining-today and tomorrow. Mining report 156 (3)

Wotruba H (2004) Kleinbergbau und seine Umweltauswirkungen- von traditionellen Methoden zu angepaßter Technologie. Geoökologisches Kolloquium WS 2004/05. Bayreuther Zentrum für Ökologie und Umweltforschung, 25.11.2004

Final Disposal 4.0

Frank Charlier and Klaus Fischer-Appelt

Contents

1 Introduction

In 2013, the German federal government and the federal states decided to restart the **site selection** for a repository for high-level radioactive waste. For this purpose, the Site Selection Act (StandAG) was enacted in 2013 [Act on the Search and Selection of a Site for a Repository for Heat Generating Radioactive Waste (Site Selection Act - **StandAG**) 2013] and amended in 2017 [(Site Selection Act - StandAG) 2017]. It aims at the search for suitable sites for a repository for high-level radioactive waste in

F. Charlier · K. Fischer-Appelt (✉)
RWTH Aachen University, Aachen, Germany
e-mail: charlier@els.rwth-aachen.de; fischer-appelt@els.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_41

the **host rocks** of **rock salt**, **crystalline** and **clay stone**. In this process, the site with the relatively best safety is to be selected using predefined **criteria**. This procedure is divided into three phases.

In the first phase sites will be determined for surface exploration in phase 2 by applying defined criteria. After that, sites for underground exploration will be selected based on criteria. Phase 3 ends with the determination of the selected site.

In addition to the **site selection project** for a repository for high-level radioactive waste, there are **other repository projects** in Germany, which are described in the chapter "Who Owns the Data". When considering these projects, '4.0 potentials' in terms of the safety of personnel and operating resources can be identified in several areas:

- **Construction** of the repository mine
- **Operation** of the repository mine
- **Transport logistics** of rock during excavation of the caverns
- **Transport logistics** of waste containers
- Transport logistics of backfill material
- **Retrieval** of waste
- **Accident analyses**
- **Monitoring** of the repository mine

When analysing the entire life cycle of a repository project, the topic of **Building Information Modelling** (BIM) can be considered beyond the classic aspects of the 4.0 concept. BIM is a method of optimised planning, execution and management of buildings and other structures with the help of software. Consistently thought through further, a repository mine can be regarded as such a structure for which BIM would offer the following advantages at many points, e.g. for operators, licensing authorities, experts, but also for the public:

- Direct and continuous availability of all current and relevant data for all parties involved. This would ensure transparency for all parties involved in the procedure, as all parties would have access to a common, consistent database.
- Improved exchange of information between those involved in the process and those involved in planning.
- Continuous documentation and data processing during the entire operating phase.
- Documentation of the repository and the stored waste in the sense of information preservation after closure of the repository.

In the authors' view, the goals of Mining 4.0, Industry 4.0, Repository 4.0 and Building Information Modelling (BIM) should be consistently linked for this purpose.

2 Repository Projects and Interim Storage of Radioactive Waste in Germany

There are currently four repository projects or plans in Germany:

- **Selection procedure** for a repository site for high-level radioactive waste
- **Konrad mine**: construction and subsequent operation of a repository for negligible heat-generating waste
- **Morsleben repository**: decommissioning of the repository for low- and intermediate-level waste
- **Asse II mine**: retrieval of the stored low- and intermediate-level radioactive waste

2.1 Site Selection Procedure

The search for an area or site for high-level radioactive waste is explained later in this paper.

2.2 Konrad Mine

The Konrad Mine is a former iron ore mine in the Salzgitter region that is currently being converted into a repository for 303,000 m^3 of radioactive waste with negligible heat generation. Approximately 90% of the low and intermediate level radioactive waste in Germany is to be disposed there. This contains about 1% of the radioactivity of all radioactive waste produced in Germany. In 2007, after several administrative court proceedings, a legally binding, incontestable licence for the construction and operation of the repository was granted. The emplacement of waste is scheduled to begin in 2027.

2.3 Asse II Mine

In parts of the 'Asse II mine', a former salt mine (Wolfenbüttel area in Lower Saxony), 47,000 m^3 of low and intermediate level radioactive waste were emplaced until 1978. In addition, research work without waste disposal was carried out until 1995.

Due to a brine inflow detected in 1988 and instabilities in some areas of the mine, the 'Act to Accelerate the Retrieval of Radioactive Waste and the Decommissioning of the Asse II Mine' (Lex Asse) stipulated in 2013 that the radioactive waste must be retrieved from the facility. This provided the preconditions to enable and accelerate the retrieval of the waste from the Asse II mine.

Currently, the retrieval of the waste is being planned by the Bundesgesellschaft für Endlagerung (BGE) as the operator of the Asse II mine. This concerns the investigation of the emplacement chambers and the radioactive waste contained therein, the planning of a largely remote-controlled retrieval technology and the infrastructure required for retrieval operations, including the construction of an *additional retrieval shaft, as well as an interim storage facility for the retrieved waste.*

2.4 Morsleben Repository

The Morsleben Repository for Radioactive Waste (ERAM) was constructed in the former Bartensleben rock salt and potash mine. It is located at the border of the federal states of Saxony-Anhalt and Lower Saxony. The ERAM was built and used in the former GDR as a repository for low- and intermediate-level radioactive waste. In the course of reunification, the Federal Office for Radiation Protection (BfS) became the operator of the facility in 1990. Radioactive waste was disposed there until 1998. Currently, the facility is being prepared for closure by BGE as the current operator of the ERAM. The ERAM is thus the first repository for radioactive waste in Germany to be decommissioned following a plan approval procedure according to nuclear law.

2.5 Gorleben Salt Dome

In 1979, the Federal Government decided to investigate the Gorleben salt dome for its suitability as a repository for high-level radioactive waste. The exploration of the Gorleben salt dome was terminated by BMU in the run-up to the Site Selection Act on 7 November 2012. In 2020, the Gorleben site was assessed as less favourable as a result of the BGE's sub-area report (BGE 2020b) and was excluded from the further site selection procedure.

2.6 StandAG 2017 (Amended)

After the final report of the 'Kommission Lagerung hochradioaktiver Abfallstoffe' (Commission on the Storage of High-Level Radioactive Waste) was handed over to the German Bundestag, the main points of the report were implemented in the Act on the Search and Selection of a Repository Site for High-Level Radioactive Waste (Site Selection Act - StandAG) of 05.05.2017.

According to StandAG [(Site Selection Act - StandAG) 2017], the purpose of the selection procedure is to identify a site with the best possible safety for a repository for high-level radioactive waste in Germany. The StandAG defines the site with the best possible safety as the site that is determined in the course of a comparative procedure from the sites suitable in the respective phase in accordance with the relevant requirements of this Act and that ensures the best possible safety for the long-term protection of human beings and the environment from ionising radiation and other harmful effects of this waste for a period of one million years'.

2.7 Responsibilities, Organisational Structures

The overarching responsibility for final and interim storage of radioactive waste in Germany lies with the **Federal Ministry for the Environment, Nature Conservation and Nuclear Safety** (BMU). It is responsible for

- nuclear safety
- repository projects and
- radiation protection

Responsibility for the German repository projects as implementer lies with the **Federal Company for Radioactive Waste Disposal** (Bundesgesellschaft für Endlagerung BGE). The BGE was formed in 2017 from parts of the Federal Office for Radiation Protection, ASSE GmbH and DBE GmbH incl. DBE TECHNOLOGY GmbH. It is responsible for operating the German repository projects. The sole shareholder of the BGE is the BMU.

The **Federal Office for the Safety of Nuclear Waste Management** (BASE), which is under the technical and legal supervision of BMU, was also newly constituted. The BASE has the following main tasks:

- Regulatory, licensing and supervisory tasks in the field of final disposal, interim storage and the handling and transport of high-level radioactive waste.
- Regulation of the site selection procedure in the search for a repository for high-level radioactive waste: BASE examines the proposals of the BGE, organises the public participation procedures and submits the examined proposals to BMU.
- Technical support of BMU.
- Task-related research in these areas.

A National Accompanying Body (**Nationales Begleitgremium** NBG) was established as a further institution in the site selection procedure. According to Section 8 of the StandAG, the task of the NBG, which is composed of recognised public figures and randomly selected citizens, is to provide mediating and independent support for the site selection procedure, in particular for public participation, with the aim of enabling confidence in the implementation of the procedure. It can deal independently and scientifically with all issues relating to the site selection procedure, question the responsible institutions at any time, issue statements and make recommendations to the German Bundestag on the site selection procedure.

2.8 The Road to a Site Decision

Site selection for a repository for high-level radioactive waste is a multi-phase procedure involving the public and the legislature. The public affected regionally and supraregionally will be involved in the procedure.

The **site selection procedure** according to StandAG is divided into the following steps:

- Identification of sub-regions (Section 13 StandAG)
- Identification of siting regions for surface exploration (Section 14 StandAG)
- Decision on surface exploration and exploration programmes (Section 15 StandAG)
- Surface exploration and proposal for underground exploration (Section 16 StandAG)
- Decision on underground exploration and exploration programmes (Section 17 StandAG)
- Underground exploration (Section 18 StandAG)
- Final site comparison and site proposal (Section 19 StandAG)
- Site decision (Section 20 StandAG)

The **Federal Company for Repository Disposal** (BGE), as the project implementer, proposes subareas, regions and sites as well as site-specific exploration programmes and assessment criteria. The **Federal Office for the Safety of Nuclear Waste Management** (BASE) reviews these proposals, is responsible for public participation and, after review, passes on the results to the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU).

At the end of each of the three phases, the **Bundestag and the Bundesrat** decide on the sites or areas to be further explored and on the final site. By examining the geological prerequisites for surface and underground exploration, the area or number of suitable areas or sites is gradually reduced.

At the end of the selection procedure, BASE proposes a final repository site to the BMU. If approved, this proposal is then forwarded by the BMU to the Bundestag and the Bundesrat. The Bundestag and Bundesrat then determine the final repository site by law.

The selection process is consistently criteria-based. The safety-oriented **geoscientific criteria** include

- exclusion criteria
- minimum requirements
- geoscientific weighing criteria

In the first step, **exclusion criteria** defined in the StandAG are used to identify areas that are obviously unsuitable as repository sites and are therefore excluded from the further selection process. The exclusion criteria relate to large scale uplift ratios, active fault zones, seismic or volcanic activity and the presence of young groundwater at repository depth. As soon as one of these exclusion criteria applies, the area in question is excluded from further selection.

In the next step, the remaining areas are examined to determine whether they meet certain **minimum requirements** also laid down in the StandAG. These include areas with rock salt, clay stone or crystalline as host rocks in the subsurface, whose rock permeability is sufficiently low and the effective containment zone has a minimum thickness of 100 m as well as a surface area sufficient for the realisation of a repository for high-level radioactive waste. The surface of an effective containment zone must be at least 300 m below the ground surface. Furthermore, there must be no findings or data that cast doubt on the integrity of the effective containment zone over a period of one million years.

In a third step, the areas identified in the two above-mentioned steps are investigated based on the 11 **geoscientific weighing criteria specified** in the StandAG. The aim is to assess whether there is a favourable overall geological situation that makes a suitability as a repository site for high-level radioactive waste appear promising. The geoscientific assessment criteria relate to

1. the transport of radioactive substances by groundwater movements in the effective containment zone
2. the composition of the rock bodies
3. the spatial characterisability
4. the long-term preservation of favourable conditions
5. the geomechanical properties
6. the tendency to form fluid pathways
7. the gas formation
8. temperature tolerance
9. the retention capacity in the effective containment zone
10. hydrochemical conditions, and
11. the protection of the effective containment zone by the overburden

The site selection procedure is to be completed in 2031 with the determination of the final site.

3 A Look at Industry and Mining 4.0

Digitisation, networking, the Internet of Things, automation, smart mining—these are just a few of the terms one immediately encounters when dealing with the topic of Industry 4.0 and Mining 4.0.

The Industry 4.0 platform of the German Federal Ministry for Economic Affairs and Energy defines as follows: ‘Industry 4.0 refers to the intelligent networking of machines and processes in industry with the help of information and communication technology' (Plattform Industrie 4.0 2019).

In mining, the 4.0 idea can be found in the following topics or keywords:

- Economic efficiency
- Safety and security
- Localisation, environment recognition, collision avoidance
- Autonomous systems
- Intelligent sensors
- Machine-to-machine communication
- Automation of processes
- Acceptance of projects

For environment detection, there are approaches with a variety of sensor technologies. Some examples are sensors, ultrasonic, (multi-)camera, LIDAR, RADAR. Depending on the application environment and requirements, the techniques are better or worse suited for mining or repository projects. For the automation of a repository, these techniques must be tested in advance for their suitability in possible host rocks.

In the case of vehicles, localisation is often carried out outdoors using the Global Positioning System (GPS), while indoors it is usually based on landmarks or existing maps with LIDAR or cameras. A special approach for the case that no map is available a priori is simultaneous localisation and mapping (SLAM) (Nüchter et al. 2004). To calculate a trajectory from the determined to a desired pose (position and orientation), there are different methods of motion planning (LaValle 2006; Laumond 1998).

In recent years, the results of this field of research have increasingly been used to contributed to the automation of industrial applications with (partially) autonomous mobile machines (Charlier et al. 2017).

Current examples of this are autonomous forklifts in intralogistics and semi-autonomous combine harvesters and tractors in agricultural technology (Viereck et al. 2013; Reinecke et al. 2013).

Many of the assistance functions in modern automobiles, such as automatic parking, also originate from mobile robotics. Even underground automation solutions already exist, for example Sandvik Mining's AutoMine (Chadwick 2010).

For underground localisation, LIDAR sensors and cameras are the most widely used. However, these optical sensors are very susceptible to faults in the underground environment. The use of RADAR sensors, which are much less sensitive to interference, for the navigation of mobile systems is the subject of research (indurad GmbH 2015).

4 Transferability of Industry and Mining 4.0 to Repository 4.0

4.1 Boundary Conditions for '4.0' in Final Disposal

The '**Repository Safety Requirements Ordinance**' (EndlSiAnfV 2020) specifies the safety level to be ensured by a repository for high-level radioactive waste in deep geological formations to meet the requirements of nuclear law.

With regard to the safe operation of a repository for high-level radioactive waste, this includes, among other things, operational safety, radiation protection for the operational phase, safety management in general and, in particular, the safe handling and emplacement of the waste packages. Furthermore, the retrievability of radioactive waste already stored during the operating phase of the repository is to be ensured and the retrievability of the waste after closure of the repository is to be made possible (Charlier et al. 2017).

Procedures must be defined and precautions taken to ensure the protection of both employees and the public in terms of occupational safety and radiation protection. In particular, it must be ensured that no radionuclides enter the mine and the biosphere in unacceptable quantities.

In mining, but also in almost all other branches of industry, processes have been increasingly automated in recent decades. This was done for economic reasons, but also for safety reasons.

For example, the use of fibre optic cables and wireless radio networks (e.g. WLAN) enables the transmission of large amounts of data in real time over long distances. The large amounts of data are captured by high-resolution sensors and transmitted in real time to the control rooms of the digitised mines, where they are processed and visualised. Increasingly precise actuators enable the robust (remote) control of machines and vehicles from a safe distance (remote control), from control rooms above ground (teleoperation) up to the complete automation of vehicles monitored from the control rooms.

With regard to the requirement for safe operation of a repository mine, the question thus arises to what extent the work processes taking place there can be automated.

In addition to minimising the radiological exposure of employees due to the spatial distance between man and radioactive waste, automated processes offer the clear advantage that accidents and incidents caused by human error, such as those observed in the events at the Waste Isolation Pilot Plant (WIPP) in the USA in February 2014, can be minimised (N.N.; https://wipp.energy.gov 2015).

Advantages of automation with regard to a Repository 4.0, for example, are (Charlier et al. 2017):

- transferring the possibilities of Mining 4.0 for the excavation and construction of the sites to be explored underground and the later selected repository site
- increasing the operational safety and reliability of the emplacement process

- the increase of operational safety and reliability in the conceptual design of the retrieval process
- the possibilities of autonomous retrieval of radioactive waste from the ASSE II mine
- the avoidance of personal injury and damage to the waste packages due to human error
- monitoring and visualisation of all safety-relevant processes, e.g. in a central control room
- complete and paperless documentation and (visual) reproducibility of the safety-relevant processes
- storage according to digital rules: No ad-hoc decisions/'shortcuts'

A new repository to be constructed, e.g. one for high-level radioactive waste, offers good boundary conditions for enabling a high degree of automation in almost all processes considered relevant to safety. In this context, it is possible to demonstrate process safety in advance by testing the programmes in studies and practical trials.

On the way to Repository 4.0, the following questions will have to be answered in the future:

- Which processes in a repository are relevant for increasing operational safety and what automation potential can be identified?
- What are the basic possibilities for process automation in a repository and what are the resulting requirements for process design?
- Which sensor technology is suitable for process automation and how can suitable process monitoring be carried out?
- What are the similarities and differences when considering the possible host rocks clay stone, rock salt and granite?
- What requirements and possibilities for automation can be derived from the demand for retrievability of waste that has already been emplaced?
- What further investigations are to be carried out in a repository with regard to the automation of selected processes?

The answers to these questions form the basis for the automation of repositories.

4.2 Current Status of the '4.0' Concept in Final Disposal

The automation of processes for the emplacement of radioactive waste has so far only been realised to a limited extent in national and international repository projects.

The concept for the final disposal of low and intermediate level radioactive waste in the Konrad repository in Germany has been described for the plan approval procedure according to § 9b AtG. Largely automated emplacement is not envisaged here. For the final disposal of high-level radioactive waste, considerations with varying degrees of detail on the emplacement of waste in rock salt, clay stone and granite (Fischer-Appelt et al. 2013; Mönig 2005; Papp 1997; Filbert et al. 2004, 2009; Pöhler et al. 2008) (selection) were carried out within the framework of research projects. The focus of the work was on geoscientific aspects, container concepts and emplacement technology.

More recent research projects also do not consider the opportunities of digitalisation and automation (Bertrams et al. 2017).

Internationally, approaches to increasing the degree of automation in the final disposal of high-level radioactive waste are discernible. Examples of this are France and Switzerland.

Furthermore, a high degree of automation is planned for the **retrieval of waste from the Asse II mine**.

The preparation for the retrieval of low and intermediate level radioactive waste is being carried out in several parallel subprojects (BGE 2021):

Drillings in the vicinity of the emplacement chambers are to be used primarily to investigate the condition of the surrounding rock, the gas composition in the chambers and the condition of the emplaced radioactive waste. At present, these investigations are being carried out in two emplacement chambers on the 750 m level. The results of the investigations provide the basis for planning the recovery technology to be used or the access to the storage chambers.

A new retrieval shaft is planned for the retrieval of the waste, as the capacities of the existing shaft are not sufficient. A starting point is currently being investigated for the new shaft from which it can be connected to the existing mine.

Retrieval is currently being planned in stages for the radioactive waste at the 511 m, 725 m and 750 m levels. To this end, the necessary retrieval techniques are being identified and aligned with the retrieval plans moving forward.

The planned automated and remotely operated retrieval technology is briefly presented below using the example of the plans for chamber 7 on the 725 m level (ELK 7/725).

After removal of backfill material in the roof area, recutting and securing of the roofs by means of floor-guided technology, remote-controlled roof-guided transport technology (overhead monorail) will be implemented, the salt heap located in the eastern part either consolidated or, if necessary, loosened and backfilled with cohesive material. The locks will be constructed in the existing western and eastern chamber access. North of ELK 7/725, new infrastructure areas will be created in the rock salt.

After locating the closest containers (e.g. visual recognition, metal detection), remote-controlled tripod excavators with appropriate tools (e.g. ripper hooks, hydraulic hammers, drum grippers) are used to uncover, release and load the containers. In terms of functionality, its upper carriage largely corresponds to that of a tunnel excavator. Important design principles of practical radiation protection are implemented on the undercarriage, such as small, easily decontaminable surfaces, crevice-free construction and detachable support claws that serve to support the tripod's own mass and the release force and can remain within the working area if necessary. Particularly for reasons of intervention capability, but also for overcoming heights as the ELK 7/25 is progressively emptied, the entire tripod excavator is moved at the roofs by an overhead monorail with hoist (BGE 2020a).

The retrieval of radioactive waste from the Asse II mine is scheduled to begin in 2033.

In general, research projects on process design are carried out on a sector-specific basis. The main focus is on information technology with the aim of improving production processes and in particular the control of machines and systems and their interaction. But there are also research approaches outside the engineering sciences, for example in the field of communication, with the background of process design.

What they all have in common is that the precise description of processes forms the basis for their execution. In the future, driverless transport vehicles could contribute to the safe operation of a repository for radioactive waste in deep geological formations by keeping humans out of the direct danger zone and additionally avoiding them as a source of error. Integral tasks of such a mobile system in the automated transport process are environment recognition and autonomous navigation, which have been researched in mobile robotics for decades.

4.3 Regulatory Framework Conditions for Increasing the Degree of Automation in a Repository Mine

Special safety precautions apply in the nuclear environment due to the inherent risks. Accidents, such as the one at the Three Mile Island nuclear power plant in Harrisburg on 28 March 1979, motivate special guidelines that apply to safety-related systems. One such standard is EN61508 on the design of electrical, electronic and programmable electronic systems performing a safety function. The aim of this standard is to define procedures for manufacturing products according to the current state of the art without disproportionate or unjustifiable hazards for users and the environment. There are various implementations of the standard for specific areas of application. There is no separate specialisation of the standard for mining; the EN/IEC 62061 standard for functional safety applies here.

Solutions for process automation in repositories are safety-relevant, but there are no regulations on this so far. There are no regulatory requirements for the (automated) operation of a repository for high-level radioactive waste. An examination of the applicability of existing regulations (AtG, KTA, BBergG and sub-legal regulations from other sectors) must be carried out.

5 Summary and Outlook

The global, cross-sector developments in automation technology offer great potential for the storage and retrieval of radioactive waste. For the **transport processes** in particular, the 4.0 idea brings an increase in operational safety and reliability.

Experience from other sectors of the economy shows that incidents and accidents caused by **human error** decrease as the level of automation increases.

An autonomous **transport system** or, for example, the **automation of the disposal process** of waste containers according to digital rules leads to a reduction of personnel underground and thus to a reduction of the probability of personal injuries as well as to a reduction of radiation exposure of the workforce. The human error factor is also reduced.

If we focus specifically on the requirement for **retrievability** during the operating phase of a repository for heat-generating waste and the legally stipulated retrieval of waste from the ASSE II mine, automation or autonomous systems also offer a gain in safety here and lead to advantages in process quality.

In addition, complete **documentation** of the processes in a Repository 4.0 can be implemented very well. At present, all potential host rocks are being considered in the site selection process. A Repository 4.0 can now be designed independently of the host rock. In conclusion, it can be said that Repository 4.0 is a realisable vision that can hardly be ignored in the future.

References

Bertrams N, Herold P, Herold M, Krone J, Lommerzheim A, Prignitz S, Simo Kuate E (2017) Abschlussbericht: Entwicklung eines technischen Konzeptes für ein generisches Endlager für wärmeentwickelnde Abfälle und ausgediente Brennelemente im Kristallingstein in. Deutschland, Peine

BGE (2020a) Plan zur Rückholung der radioaktiven Abfälle aus der Schachtanlage Asse II, PSP-Element 9A 23500000, Stand 19.02.2020, Peine

BGE (2020b) Summary sub-areas interim report according to Section 13 StandAG. As per 28/09/2020. Ref.: SG01101/16-1/2-2020#30–Object ID: 830270–Revision: 000, Peine

BGE (2021) Kurzinformation zur Rückholung der radioaktiven Abfälle. Internetseite der Bundesgesellschaft für Endlagerung. https://www.bge.de/de/asse/kurzinformationen/rueckholung-der-radioaktiven-abfaelle. Zugegriffen am 15.02.2021

Bonn indurad GmbH, UPNS4D+ (2015) Underground 4D+ Positioning, Navigation and Mapping System, Webseite. http://indurad.com/index.php?id=82. Zugegriffen am 20.03.2019

Chadwick J (2010) The autonomous mine. In: International mining, January, pp 46–57

Charlier F, Gutberlet D, Rabel M (2017) Repository 4.0: opportunities of process automation. In: Proceedings 48th annual meeting on nuclear technology, Berlin

Filbert W et al (2004) Gegenüberstellung von Endlagerkonzepten in Salz und Tongestein, (GEIST): Abschlussbericht. DBE TECHNOLOGY GmbH, Peine

Filbert W et al (2009) Optimierung der Direkten Endlagerung durch Kokillenlagerung in Bohrlöchern, Machbarkeitsstudie und Entwurfs- und Konzeptplanung (DENKMAL). DBE TECHNOLOGY GmbH, Peine

Fischer-Appelt K et al (2013) Vorläufige Sicherheitsanalyse Gorleben (VSG). GRS, Köln

Gesetz zur Suche und Auswahl eines Standortes für ein Endlager für Wärme entwickelnde radioaktive Abfälle (Standortauswahlgesetz - StandAG) (2013) Bundesministerium der Justiz und für Verbraucherschutz, Bonn

Gesetz zur Suche und Auswahl eines Standortes für ein Endlager für Wärme entwickelnde radioaktive Abfälle (Standortauswahlgesetz - StandAG) (2017) Bundesministerium der Justiz und für Verbraucherschutz, Bonn

Laumond J-P (1998) Robot motion planning and control. Springer, Berlin

LaValle SM (2006) Planning algorithms. Cambridge University Press, New York

Mönig et al (2005) Überprüfung und Bewertung des Instrumentariums für eine sicherheitliche Bewertung von Endlagern für HAW (ISIBEL), GRS, BGR, DBE

N.N. (2015). https://wipp.energy.gov; http://www.wipp.energy.gov/wipprecovery/accident_desc.html. Zugegriffen am 10.11.2015. Zugegriffen am 20.03.2019

Nüchter A, Surmann H, Lingemann K, Hertzberg J, Thrun S (2004) 6D-SLAM with an application in autonomous mine mapping. In: Proceedings of the international conference on robotics and automation (ICRA'04), New Orleans, pp 1998–2003

Papp R (1997) GEISHA: Gegenüberstellung von Endlagerkonzepten in Salz und Hartgestein. Kernforschungszentrum, Karlsruhe

Plattform Industrie 4.0 (2019) Internetseite des Bundesministeriums für Wirtschaft und Energie. https://www.plattform-i40.de/I40/Navigation/DE/Industrie40/WasIndustrie40/was-istindustrie-40.html. Zugegriffen am 20.03.2019

Pöhler M, et al (2008) Referenzkonzept für ein Endlager für radioaktive Abfälle in Tonstein (ERATO), Abschlussbericht, DBE TECHNOLOGY GmbH, Dezember, Peine

Reinecke M, Grothaus H-P, Hembach G, Scheuren S, Hartanto R (2013) Dynamic and distributed infield-planning system for harvesting. In: 2013 ASABE annual international meeting, Kansas City, Missouri, July 21–24

Verordnung über Sicherheitsanforderungen an die Endlagerung hochradioaktiver Abfälle (Endlagersicherheitsanforderungsverordnung – EndlSiAnfV, 2020) Bundesministerium der Justiz und für Verbraucherschutz, Bonn

Viereck V, Tödter J, Lacher G (2013) Touch the future: Seriengeräte flexibel und intuitiv automatisiert, VDI-Berichte 2201, 17, VDI Flurförderzeugtagung, 24.–25.09.2013. Nomos, Baden-Baden Endlagerung 4.0 © Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature 2020

Geological Modeling 4.0

From Static Models to Dynamic Tools

Florian Wellmann

Contents

1 Summary

Knowledge about the geological subsurface is important for a variety of industrial and technical applications: as a resource of mineral and non-mineral raw materials, as a reservoir, and as a planning ground for underground infrastructure—and in this context also increasingly in the context of urban planning and "Building Information Management" (BIM) extended into the subsurface, often also referred to as GeoBIM (Svensson 2015). The central problem of characterizing significant rock and fluid properties in the subsurface is that, to date, there is no inexpensive and comprehensive way to directly measure these properties. Although (expensive) boreholes provide very precise information at specific points, extrapolation into space is only possible indirectly with the aid of geophysical measurements. In particular, seismic, gravimetric, and electromagnetic measurements are often used for this purpose (e.g., Telford et al. 1990). From all this information, a geometric structural model is

F. Wellmann (✉)
Computational Geoscience and Reservoir Engineering, RWTH Aachen, Aachen, Germany
e-mail: florian.wellmann@cgre.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_42

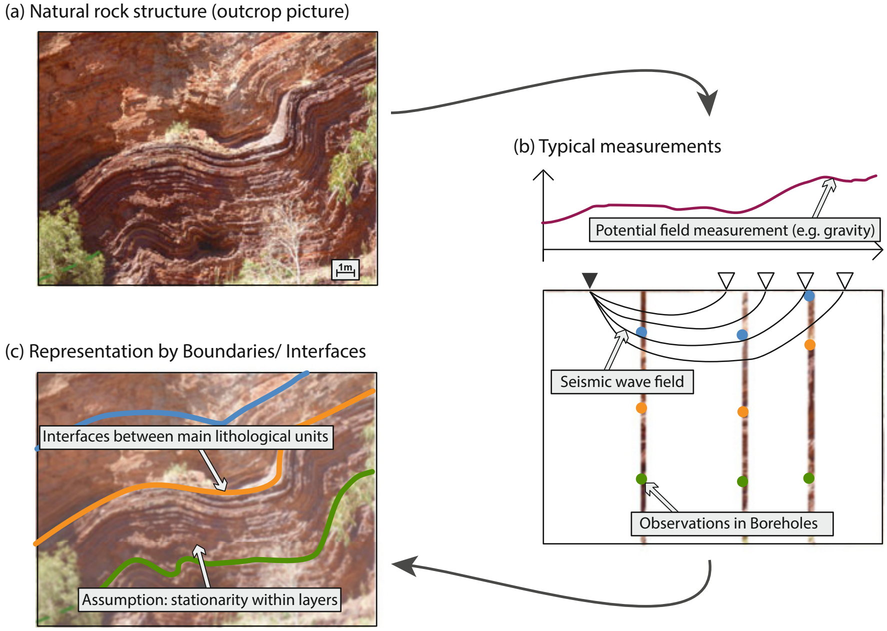


Fig. 1 Geological interpolation and typical measurements. (**a**) Natural rock structure (outcrop picture). (**b**) Typical measurements. (**c**) Representation by boundaries/interfaces

typically created, in which rocks are grouped into formations that show similar properties in the context of the investigation. This view is conceptually illustrated in Fig. 1 for a typical rock sequence, here as an example with a photograph from an outcrop (Fig. 1a). If we now imagine that this rock sequence is not at the surface but in the subsurface, then we often have only very isolated direct observations in the direct vicinity of these boreholes (Fig. 1b). The crucial question then is how these observations can be interpolated in space, between the known points and extrapolated beyond them. Combining geological and geophysical aspects, an attempt is typically made to determine interfaces between units with similar properties (Fig. 1c).

This modeling of boundary layers already found its beginning in the first geological investigations, maps and cross sections through the Earth and is until today one of the corner points of subsurface investigations. The representation of geology as discrete surfaces in 3D space is based on the realization that important events in geological time often manifested themselves in thin transition zones, leading to rock units with distinct properties. What has changed, however, are the methods used today to create these geological models. Currently, we are at a significant step that, similar to Industry 4.0, combines and integrates multiple types of heterogeneous data, allows for a representation of complex geological structures, and additionally enables the quantification and visualization of uncertainties.

In this chapter, the evolution of geological modeling is briefly outlined to highlight the specifics of the advances made in recent years and the potential for a better understanding of the subsurface.

Then, the significant elements of "Geological Modeling 4.0" are described in detail, and finally, connections to other elements of Industry 4.0 are considered, including in particular the better integration of geological modeling into existing workflows, as well as an improved communication of geological settings and uncertainties.

2 A Short History of Geological Modeling

The history of geological mapping and modeling is tightly linked to industrialization, specifically through its relevance for the exploration of mineral resources. This aspect is evident in the history of the Industrial Revolution in England, with the first geological map attributed to William Smith (Smith 1815), which had great significance for the search for coal, as a central element of the Industrial Revolution (Winchester 2001; Sharpe 2015; Wigley 2016). However, geologic mapping with a similar goal existed already before. Of particular importance in this regard are the geological map of the Cesenate sulphur mines by the Italian Count Luigi Ferdinando Marsili, drawn before 1717 (Romano et al. 2016) and the early geologically motivated regional surveys by the state of Saxony (by Von Charpentier 1778).

The first geological models were recorded in drawings and sections within the framework of the available technical capabilities. Early on, block models were already created in perspective views. These methods have developed further—for example by the systematic use of color (Sharpe 2015)—but the technical possibilities in the creation remained limited for a long time. This early phase can be described as the first stage of geological modeling, as *Geological Modeling 1.0*.

A major step forward was then brought about by the development of geophysical measurement techniques to obtain indirect information about the subsurface. In particular, early gravity and seismic measurements led to significant advances in the analysis of the subsurface and thus had a major impact on the development of geological models (see Birett et al. 1974, History of Geophysics). Motivated by this successful application, a wide variety of mathematical algorithms for geological interpolation were used, which in particular greatly simplified the creation of sections and profiles. We refer to his phase here as *Geological Modeling 2.0*. But the algorithms were still limited in their capacity for the representation of complex geological relationships.

In the context of the further availability of numerical methods and especially the rapid developments in computer-aided data processing in the second half of the twentieth century, approaches were then increasingly developed to map even complex geological structures in space in 3D models (see also overview in Jessell et al. 2014). Connection to databases and linkage to geophysical models enabled the incorporation of many data sources (e.g., Zobl et al. 2011). This third phase of geological modeling, *Geological Modeling 3.0*, can be roughly compared to the impact of IT in the *Industry 3.0* developments.

Although methods have been available since this time to create complex geo-logical models in 3D, they still have some deficits. In particular, typical workflows still contain many manual steps. This aspect makes it often difficult to revise geological models once they have been created or to integrate newly acquired data. Furthermore, usually only one model is created—although there is always a multitude of possible models, since the information about the subsurface is always only partially available and geological models therefore often contain significant uncertainties. And especially the rapid developments in the field of collected data and machine learning methods now call for a next step in the development of geological models, towards *Geological Modeling 4.0*. These developments are described in this chapter, especially with regard to the linkage with other areas of *Industry 4.0*.

3 Elements of Geological Modeling 4.0

In the following, significant elements of the current developments in geological modeling are presented, in particular also criteria that will enable a tighter integration into the concurrent developments of *Industry 4.0*.

3.1 Automating Model Generation and Model Updates

The life cycle of a measure affecting the subsurface (e.g., the sustainable extraction of groundwater or mineral resources, or an infrastructure development such as a tunnel project) often lasts for several 10's to over 100 years. Within this timeframe, it can be expected that methods and capabilities for subsurface investigations will develop significantly and, in particular, that new measurements will become possible—the geological subsurface, on the other hand, will remain largely the same (Svensson 2015). With this in mind, it is important that workflows are designed so that predictions from models can be updated quickly and flexibly as new data become available. Automation of geological modeling methods is therefore a significant aspect of *Geological Modeling 4.0*.

In recent years, various methods have been developed that enable this automation of model generation, even for complex 3D geological structures (Wellmann and Caumon 2018). Geometric interpolation algorithms can now handle a large amount of input data of various types, and these algorithms increasingly consider aspects of geological knowledge (e.g., on spatial continuity and the relationship between multiple layers, multiply deformed features, etc.).

In general, explicit and implicit surface representations can be distinguished. Explicit methods enable local adjustments, are overall very flexible and allow the representation of complex structures and relationships. Implicit methods, on the other hand, have been used very successfully in recent years because certain geological conditions can be directly incorporated into the interpolation, for example the important condition that layered surfaces cannot intersect in a continuous sedimentation sequence. Furthermore, implicit approaches can be automated very efficiently due to the global interpolation approach, but local adjustments can be difficult. In recent work, graph neural networks (GNN's) have been proposed as a novel implicit interpolation approach, that enables both global and local adjustments (Hiller et al. 2021).

In general, most approaches of geological modeling can be described with an interpolation function, based on the position in space, as well as other parameters of the interpolation method (the used basis functions and their parameters, the coefficients of the spatial discretization, the primary geological observations that are included, a topological description, as well as other regularization methods). An overview of commonly used methods is presented in Wellmann and Caumon (2018).

3.2 Reproducibility, Open Data, Transparency

In the process of geological model development, decisions are made at numerous steps that depend heavily on the experience and prior knowledge of the modeler. This is compounded by the use of different data from multiple sources of varying quality and grade. It is therefore to be expected that two experts with the same input data may nevertheless arrive at different results. The importance of prior knowledge in the interpretation of geological data and the hermeneutic aspect of geology has been described, for example, in Frodeman (1995) and subject to recent studies on bias in geological interpretation and modeling (e.g., Bond et al. 2007; Schaaf and Bond 2019).

In connection with the increasing automation of geological modeling, it is also becoming more and more possible to reproduce modeling results of complex workflows. Common commercial software environments allow for the recording and storage of entire modeling workflows and detailed quality control steps. Increasingly, this is also taking into account uncertainties in the entire sequence of model creation (e.g., Singh et al. 2013).

In the context of current developments in research data management and data science, it can be expected that there will be increasing opportunities to track the entire life cycle of raw geological and geophysical data through processing steps, right into model creation; finally leading to an increased transparency in the complex decision-making processes in which the geological models are used.

3.3 Integration of Heterogeneous Data

Geological models have always had to be created based on only a few direct measurements, which have to be combined with prior geological knowledge and many other indirect measurements to create a complete 3D geological model representation. This problem is illustrated in Fig. 1b, where partially available direct observations from boreholes are combined with geophysical measurements. This combination of very heterogeneous data with different aspects of geological knowledge has always been a part of geological modeling (see Wellmann and Caumon 2018, for more details). However, this aspect is now becoming increasingly significant with the acquisition of more and more data: both through advances in geophysical measurements, increasingly also airborne measurements based on drones, and through the combination with more satellite-based measurements, an increasing amount of (digitized) borehole measurements, and the integration of additional sensor data (e.g., from tunnel boring machines). This gives further importance to the efficient integration of these heterogeneous data, and most geological modeling environments are only at the beginning of including his rich set of information.

Another significant aspect is that not all data and knowledge can be directly incorporated into any interpolation or modeling algorithm. Some observations and data can only be compared based on the generated model. Typical examples are some geophysical measurements, but also certain geological observations. With increasing availability of different heterogeneous data, it becomes more and more challenging to bring all (and possibly even contradictory) information together in one geological model. Some approaches exist to overcome this problem through an integration in probabilistic models (e.g., Wellmann et al. 2017). However, these developments are only at the beginning, but it can be expected that model-based inversions will allow an increasingly better and more clearly structured integration of heterogeneous data sets in geological modeling workflows.

3.4 Estimation and Communication of Uncertainties

Geological models often contain significant uncertainties because there is no comprehensive method to measure significant rock properties directly with a high spatial resolution. The description and typification of geological uncertainties has therefore long been the subject of scientific consideration (e.g., Mann 1993). However, until now it has hardly been possible to comprehensively represent relevant uncertainties in geological models. In the context of automating geological modeling, recent years have seen some new approaches to estimating these uncertainties (Holden et al. 2003; Wellmann et al. 2010; Wellmann and Caumon 2018), their importance in applications (e.g., Schneeberger et al. 2017; Schweizer et al. 2017), and their representation and communication in 3D geological models (Wellmann and Regenauer-Lieb 2012; Lindsay et al. 2012; Wellmann and Caumon 2018).

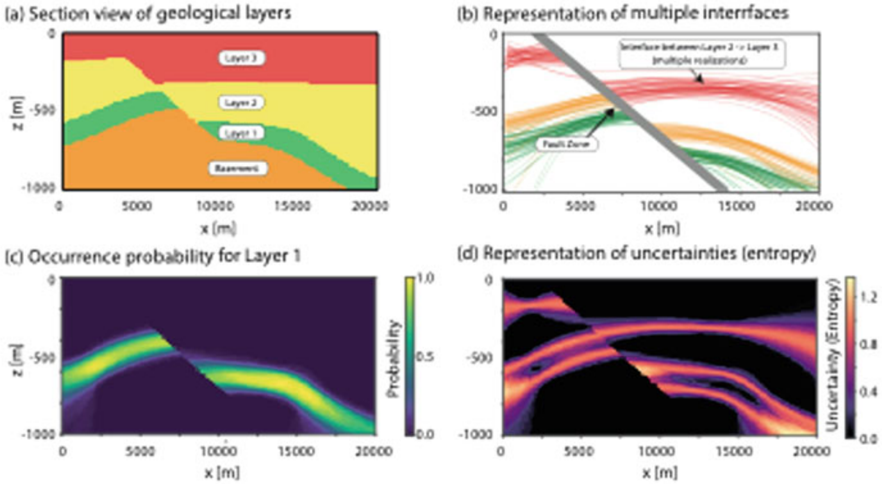


Fig. 2 Representation of uncertainties in geological models. (**a**) Section view of geological layers. (**b**) Representation of multiple interfaces. (**c**) Occurrence probability for Layer 1. (**d**) Representation of uncertainties (entropy) (adopted after De la Varga et al. 2019)

Significant uncertainties lie in the creation of the conceptual model itself and these are difficult to estimate. One option here is to create multiple conceptual models. Uncertainties from input data and interpolation parameters can partly be estimated, based on the measurements themselves, or on expert knowledge (e.g., Wellmann and Caumon 2018), although a general framework for the estimation of parameter uncertainties is still lacking. Still, keeping this restriction in mind, stochastic realizations of geological models can be created and analyzed using appropriate automated modeling methods.

A simple example of this is shown in Fig. 2. In Fig. 2a, in a vertical section through a geological model. The model considered here consists of a total of four geologic units in a concordant sequence of layers, offset along a fault. Figure 2b shows several realizations of this model in the same section, here now only the interfaces between the layers are shown for more clarity. Figure 2c now shows the probability of encountering a specific layer of this layer stack (here: “layer 1”) at each location in the background. This plot is suitable for the analysis of a single layer. One way to calculate the uncertainty of the combined probabilities of all layers is possible using the concept of entropy from information theory (Shannon 1948), applied to spatially discrete model domains (Wellmann and Regenauer-Lieb 2012). This quantity is shown in Fig. 2d, and it is readily apparent that areas around the interfaces themselves, in particular, are highly uncertain.

These methods for representing uncertainty are increasingly being applied, both in scientific papers (e.g., Schweizer et al. 2017) and by geological services and organizations, for example in the nationwide geological model of the Netherlands (Stafleu et al. 2020).

3.5 Machine Learning and Artificial Intelligence

The combination of the automated geological interpolation algorithms described above, in combination with the currently rapid developments in the field of machine learning, is already leading to interesting developments. Machine learning methods have been successfully applied for several years to process and interpret geophysical data and measurements from remote sensing (e.g., Caté et al. 2017; Huang et al. 2017; Maxwell et al. 2018). These developments will mean that measurements can be incorporated more directly into geological models in the future. In particular, this will also automate the labor-intensive and potentially error-prone steps of manual processing and interpretation of data and lead to a more objective integration. However, there are still significant steps to be taken on the way to further application of these methods. Many of the currently developed methods from the field of deep learning are optimized for the analysis of image and audio information and thus not applicable to all areas of geoscientific data analysis and geological modeling. Above all, these methods have so far lacked the integration of physical conditions and an estimation of uncertainties. Here, so-called hybrid modeling methods currently promise interesting developments (e.g., Reichstein et al. 2019), which will also be applicable to the field of geological modeling and already show first successful possibilities to combine heterogeneous data in probabilistic models in this way (de la Varga et al. 2019). Another promising way forward is the integration of neural network methods in the geological interpolation step itself (e.g., Hillier et al. 2021).

To what extent advances from the large field of artificial intelligence can be applied to geological modeling remains to be seen. However, in the context of the rapid developments, it can be assumed that interesting opportunities will also arise here, especially also in the integration of different data sources and automated sampling, for example coupled to autonomous drones equipped with geophysical measuring devices (Stoll and Moritz 2013), which could acquire data in such a way that a geological model is increasingly optimized through continuous integration.

An illustration of a probabilistic machine learning model with an integrated geomodeling step is shown in Fig. 3 with a priori distributions over (uncertain) geologic observations, the geologic modeling itself as a forward model, and Bayesian inference based on likelihood functions that incorporate further measurements and observations. These complex models can now be solved numerically through an incorporation into efficient machine learning programming frameworks (e.g., TensorFlow, Abadi et al. 2016).

4 Applications of Geological Modeling 4.0

The outlined developments of *Geological Modeling 4.0* will directly benefit those application areas in which these models are already widely used—for example, in groundwater exploration, the optimized search for raw materials, and in the field of

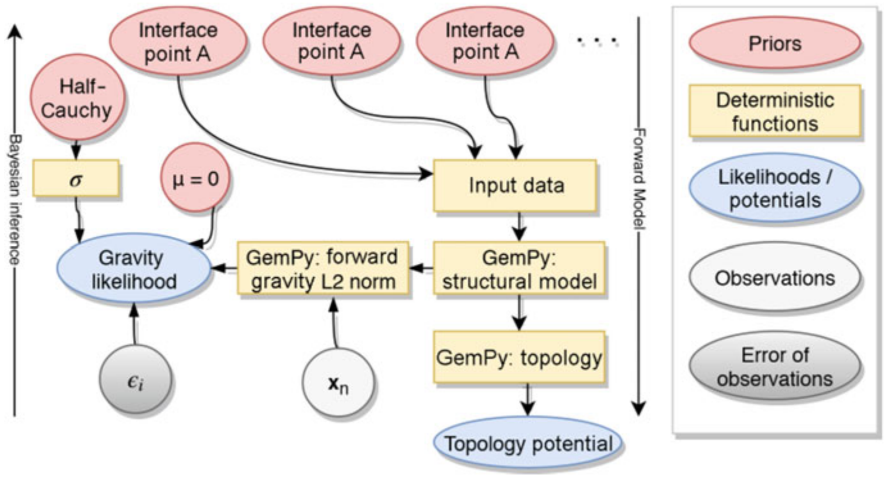


Fig. 3 Example of a probabilistic geological modeling set-up (de la Varga et al. 2019)

engineering geology. However, the future developments will also open up new areas of application, in which the link to other aspects of *Industry 4.0* is particularly interesting.

The widely used commercial software programs in the field of geological modeling already use some of the aspects described here. In particular, methods for automated modeling and the integration of heterogeneous data are at least partially implemented in many programs. Methods for stochastic simulation and estimation of uncertainties are already available. However, some of the algorithms and settings in commercial programs are not published and not even clearly described in white papers, and thus the models created are only partially reproducible. In the area of data formats, there has been some progress in recent years (e.g., Castronova et al. 2013), but it is still not foreseeable that commercial products will move to schemas that allow the basic algorithms, parameter options and settings, as well as the input data used, to be communicated in such a way that the models can be used in another system.

In this context, developments are interesting in which individual aspects of the entire modeling workflow are considered separately and developed further. Especially in the area of machine learning and artificial intelligence, the integration of very powerful packages (e.g., TensorFlow, PyTorch) into existing open-source programming languages such as Python or R has led to significant progress in recent years, especially in combination with other aspects of *Industry 4.0*. In the field of geological modeling, similar advances can be expected if methods allow geological modeling and interpolation, as well as geophysical forward models, to be integrated into these environments. This insight has already led to open-source initiatives in geological modeling (e.g., de la Varga et al. 2019, Ailleres et al. 2018), and geophysical data processing and inversion (e.g., Cockett et al. 2015; Rücker et al. 2017), and it can be expected that these methods can be increasingly connected to the already very advanced and well-developed open-source packages for data analysis, machine learning, and artificial intelligence.

The described splitting of monolithic software environments into individual sub-aspects connected by open data structures is a similar effect observed in *Industry 4.0* under the term "unbundling," namely the splitting of complex workflows, previously performed monolithically by one company, into individual sub-aspects often taken over by small companies. A similar development can be expected with the "unbundling" of workflows in *Geological Modeling 4.0*, if open data formats allow a suitable communication between the components.

An important development in *Industry 4.0* is the increasing use of so-called Digital Twins, i.e., digital representations of real objects. Cyber-physical systems (CPS), in which there is direct communication between real and virtual objects in a system, occupy a special position in this area. These systems are also increasingly being developed in the field of geosciences—for example, for the spatial prediction of climatic and hydrological conditions. These systems are increasingly extended to integrated components of the subsurface, and the aspects of *Geological Modeling 4.0* described here will make an integration possible in the foreseeable future.

Under the aspects described above, new fields of applications and business will also arise with the developments of *Geological Modeling 4.0*. For example, it has already been shown that, by combining an efficient modeling algorithm in an open-source environment, it is possible to create geological models directly from observations in GoogleEarth (Wellmann et al. 2019), making them comprehensible and implementable for interested parties. Also significant are the deployments of geological data and models from geological services (e.g., Kessler et al. 2009; van der Meulen et al. 2014). These modeling methods now allow users to view and possibly even modify their own models of the local subsurface. On the other hand, due to the possibility of automations, data obtained from all applications can also be fed back into the models to successively optimize them. Of course, this raises new questions regarding the quality assurance of the data, but in principle this possibility exists. It is known from other fields that public involvement can lead to very positive developments (e.g., in the field of crowdsourcing of climatic and atmospheric data, Muller et al. 2015), and a transfer to the field of geological modeling could lead to a better and more positive perception of geoscientific methods and measurements—and a better understanding of the important role of the geological subsurface for our society.

5 Conclusion and Outlook

The importance of geological modeling for the representation of the subsurface is gaining more and more importance—especially in view of an increasing use of the underground space. Especially, new developments in the field of automation of modeling and the integration of heterogeneous data from different measurements, as well as the possibility of continuous model updates allow a variety of new applications. These developments are summarized here under the umbrella term *Geological Modeling 4.0*.

The combination with other elements of the *Industry 4.0* developments is particularly interesting. Geological systems have been components of simple Digital Twins for some time now—for example in groundwater supply. These models are then continuously updated by observations. The expansion of the development of Digital Twins in other areas of *Industry 4.0* offers exciting opportunities to better link progress on both sides.

References

Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M et al (2016) TensorFlow: a system for large-scale machine learning. In: 12th USENIX symposium on operating systems design and implementation (OSDI 16). USENIX Association

Ailleres L, Grose L, Laurent G, Armit R, Jessell M, Caumin G, de Kemp E, Wellmann F (2018) LOOP – an open source platform for 3D geo-structural simulations. In: 2018 Resources for Future Generations meeting, June 16–17, Vancouver, Canada

Birett H, Helbig K, Kertz W, Schmucker U (eds) (1974) Zur Geschichte der Geophysik. Springer, Heidelberg

Bond CE, Gibbs AD, Shipton ZK, Jones S (2007) What do you think this is? "Conceptual uncertainty" in geoscience interpretation. GSA Today 17(11):4

Castronova AM, Goodall JL, Elag MM (2013) Models as web services using the Open Geospatial Consortium (OGC) Web Processing Service (WPS) standard. Environ Model Softw 41:72–83

Caté A, Perozzi L, Gloaguen E, Blouin M (2017) Machine learning as a tool for geologists. Lead Edge 36(3):215–219

Cockett R, Kang S, Heagy LJ, Pidlisecky A, Oldenburg DW (2015) SimPEG: An open source framework for simulation and gradient based parameter estimation in geophysical applications. Comput Geosci 85:142–154

De la Varga M, Schaaf A, Wellmann F (2019) GemPy 1.0: open-source stochastic geological modeling and inversion. Geosci Model Dev Discussions 12(1):1–32

Frodeman R (1995) Geological reasoning: geology as an interpretive and historical science. Geol Soc Am Bull 107(8):960–968

Hillier M, Wellmann F, Brodaric B, de Kemp E, Schetselaar E (2021) Three-dimensional structural geological modeling using graph neural networks. Math Geosci 53(8):1725–1749

Holden L, Mostad P, Nielsen B, Gjerde J, Townsend C, Ottesen S (2003) Stochastic structural modeling. Math Geol 35(8):899–914

Huang L, Dong X, Clee TE (2017) A scalable deep learning platform for identifying geologic features from seismic attributes. Lead Edge 36(3)

Jessell MW, Ailleres L, de Kemp E, Lindsay M, Wellmann F, Hillier M et al (2014) Next generation three-dimensional geologic modeling and inversion. Soc Econ Geol Spec Publ 18(18):261–272

Kessler H, Mathers S, Sobisch H-G (2009) The capture and dissemination of integrated 3D geospatial knowledge at the British Geological Survey using GSI3D software and methodology. Comput Geosci 35(6):1311–1321

Lindsay M, Ailleres L, Jessell MW, de Kemp E, Betts PG (2012) Locating and quantifying geological uncertainty in three-dimensional models: analysis of the Gippsland Basin, South-eastern Australia. Tectonophysics 546-547:10–27

Mann JC (1993) Uncertainty in geology. Computers in geology—25 years of progress. Oxford University Press, pp 241–254

Maxwell AE, Warner TA, Fang F (2018) Implementation of machine-learning classification in remote sensing: an applied review. Int J Remote Sens 39(9):2784–2817

Muller CL, Chapman L, Johnston S, Kidd C, Illingworth S, Foody G et al (2015) Crowdsourcing for climate and atmospheric sciences: current status and future potential. Int J Climatol 35(11): 3185–3203

Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, Prabhat. (2019) Deep learning and process understanding for data-driven Earth system science. Nature 566(7743):195

Romano M, Cifelli RL, Vai GB (2016) The first geological map: an Italian legacy. Italian J Geosci 135(2):261–267

Rücker C, Günther T, Wagner FM (2017) pyGIMLi: An open-source library for modelling and inversion in geophysics. Comput Geosci 109:106–123

Schaaf A, Bond CE (2019) Quantification of uncertainty in 3-D seismic interpretation: implications for deterministic and stochastic geomodeling and machine learning. Solid Earth 10(4): 1049–1061

Schneeberger R, de la Varga M, Egli D, Berger A, Kober F, Wellmann F, Herwegh M (2017) Methods and uncertainty estimations of 3-D structural modelling in crystalline rocks: a case study. Solid Earth 8(5):987–1002

Schweizer D, Blum P, Butscher C (2017) Uncertainty assessment in 3-D geological models of increasing complexity. Solid Earth 8(2):515–530

Shannon EC (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423

Sharpe T (2015) The birth of the geological map. Science 347(6219):230–232

Singh V, Yemez I, Sotomayor J (2013) Integrated 3D reservoir interpretation and modeling: lessons learned and proposed solutions. Lead Edge 32(11):1340–1353

Smith W (1815) A delineation of the strata of England and Wales, with part of Scotland; exhibiting the collieries and mines, the marshes and fen lands originally overflowed by the sea, and the varieties of soil according to the variations in the substrata, illustrated by the most descriptive names. John Cary, London

Stafleu J, Maljers D, Gunnink J (2020) Visualisation of uncertainty in voxel models of the shallow subsurface of the Netherlands. In: 82nd EAGE Annual Conference & Exhibition Workshop Programme, vol 2020, no 1, pp 1–3. European Association of Geoscientists & Engineers

Stoll J, Moritz D (2013) Unmanned aircraft systems for rapid near surface geophysical measurements. In: 75th EAGE conference and exhibition – workshops

Svensson M (2015) GeoBIM for optimal use of geo data. In: Conf. Proc., Sustainable exploitation of the subsurface: the geology, engineering and environment of our underground asset, 20–21 May 2015, London

Telford WM, Telford WM, Geldart LP, Sheriff RE (1990) Applied geophysics. Cambridge university press, p 770

Van der Meulen MJ, Doornenbal JC, Gunnink JL, Stafleu J, Schokker J, Vernes RW et al (2014) 3D geology in a 2D country: perspectives for geological surveying in the Netherlands. Netherlands J Geosci 92(04):217–241

Von Charpentier J (1778) Mineralogische Geographie der Chursächsischen Lande. Leipzig

Wellmann F, Caumon G (2018) 3-D structural geological models: concepts, methods, and uncertainties. Adv Geophys 59:1–121

Wellmann F, Regenauer-Lieb K (2012) Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models. Tectonophysics 526 IS:207–216

Wellmann F, Horowitz FG, Schill E, Regenauer-Lieb K (2010) Towards incorporating uncertainty of structural data in 3D geological inversion. Tectonophysics 490(3-4):141–151

Wellmann F, de la Varga M, Murdie RE, Gessner K, Jessell MW (2017) Uncertainty estimation for a geological model of the Sandstone greenstone belt, Western Australia – insights from integrated geological and geophysical inversion in a Bayesian inference framework. Geological Society, London, Special Publications, 453

Wellmann F, Schaaf A, de la Varga M, von Hagke C (2019) From Google Earth to 3D geology problem 2: seeing below the surface of the digital earth. Dev Struct Geol Tectonics 5:189–204
Wigley P (2016) The development and evolution of the William Smith 1815 geological map from a digital perspective. GSA Today 26(7)
Winchester S (2001) The map that changed the world: William Smith and the birth of modern geology. HarperCollins, New York
Zobl F, Chmelina K, Faber R, Kooijman J, Marschallinger R, Stoter JE (2011) Multidimensional aspects of GeoBIM data: new standards needed. In: IAMG 2011: Proceedings of the international association of mathematical geosciences: mathematical geosciences at the crossroads of theory and practice, 5–9 September 2011, Salzburg, Austria

Part VI
Medicine and Healthcare

Telemedicine

Gernot Marx and Robert Deisz

Contents

1 Background

The **digital revolution** and Medicine 4.0 are already clearly noticeable in everyday clinical practice and will continue to profoundly change our health care system in the coming years. Increasing amounts of medical data are already being recorded

G. Marx (✉) · R. Deisz
UK RWTH Aachen, Aachen, Germany
e-mail: gmarx@ukaachen.de; robert.deisz@philips.com

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_43

digitally. A well-known example is the introduction of electronic patient files in doctors' surgeries and hospitals as a replacement for traditional, analogue paper files. In the home environment, app-based applications on smart phones (health apps) allow chronically ill patients to monitor their therapy more closely and become more actively involved in their treatment.

The development of digital health applications has great potential to contribute to high-quality health care and to promote **patient autonomy**. According to the results of a representative survey by the Bertelsmann Foundation, 80% of the population would like to jointly decide on their therapy together with the doctor providing treatment (Krüger-Brand 2019). However, the digital recording of treatment parameters is only the beginning. For more effective use, the data must not only be recorded automatically, but must also be interoperable. Only then can they be systematically analysed, processed, and presented in a way that meets needs—for example, to support in decision making. It must also be ensured that the information is available around the clock for those affected and the doctors treating them. Only then can digital data recording add value to the quality of treatment for patients.

At the same time, the extensive amounts of data in medicine are an ideal basis for Big Data analyses (BDA) and machine learning (ML), as well as pattern recognition and the development of artificial intelligence (AI). Intelligent decision support systems based on these methods can assist doctors in making diagnoses and planning therapies. The technical possibilities seem unlimited. However, we must always be aware that the focus is on benefit for the patient—or, more precisely, the patient's well-being. New technologies should not—and must not—be a substitute for individual treatment from a doctor and interpersonal contact, but should rather be developed and used as an enhancing method so that the quality of treatment can be improved through improved diagnostics and therapy, or by increasing the availability of specialist expert knowledge.

Demographic change with increasing patient numbers and concurrently decreasing numbers of employees will pose new challenges to the health care system in practically all areas. For intensive care medicine, this specifically means that the number of critically ill patients will continue to increase. This increasing demand for intensive care services is being met by a shortage of young medical professionals, because it is estimated that there will be a shortfall of 111,000 doctors in Germany by 2030 (German Medical Association 2014). There is also an increasing shortage of nurses in the ICU, which affects the quality of intensive care treatment (Karagiannidis et al. 2019). The restructuring of the hospital landscape will also lead to a reduction of smaller, local hospitals and to specialisation in centres. The treatment of complex disease patterns in rural areas with patient care close to home will no longer be guaranteed in future (Bölt and Graf 2012). New strategies and solutions must therefore be found to ensure high-quality, guideline-based patient care and to make it available as needed at any place and at any time.

The answer to the relevant and increasing shortage of urgently needed experts for needs-based and comprehensive round-the-clock care lies, among other things, in telemedical co-operation structures. The expansion of telemedical care and intelligent digital networking offers huge innovation potential for our health care system.

To make the medical care structure more flexible, individualised and efficient, this expansion must be accelerated. In intensive care medicine, for example, numerous studies have demonstrated that telemedical co-operation can significantly improve the care of critically ill patients.

2 Telemedicine and Tele-Intensive Care

The term **telemedicine** is not uniformly defined. The German Medical Association (Bundesärztekammer 2014) describes telemedicine as a 'collective term for various medical care concepts that share the fundamental approach that medical services in the areas of diagnostics, therapy, and rehabilitation, as well as advice on medical decision making, are provided to the population over distances or with a time delay. Information and communication technologies are used in the process'.

Telemedicine has a diverse range of applications. It can be used either between doctor and patient (doc2patient) or between two doctors, for example, GP and specialist physician (doc2doc). Roughly systematised, there are three different areas of application:

- Teleconsultations for interdisciplinary exchange between medical colleagues (e.g. tele-tumour board reviews)
- Remote patient monitoring for continuous monitoring of vital parameters by medical experts (e.g. blood pressure monitoring)
- Teletherapy (e.g. Internet-based psychotherapy) (Marx and Beckers 2015)

In this sense, **tele-intensive care** is a combination of teleconsultations or tele-ward rounds between a tele-intensive care centre and connected ICUs as well as remote patient monitoring. The term remote patient monitoring is broadly defined here. Vital data (e.g. blood pressure, pulse and respiratory rate), laboratory values (e.g. inflammation and infection markers), device or treatment data (e.g. data from ventilators), information about medication administered, and complex treatment algorithms are transmitted. Teleconsultations or tele-ward rounds via specially secured data connections are usually carried out daily (and additionally around the clock if required).

The aim of telemedical services is to improve the patient's state of health through interdisciplinary exchange—regardless of time and place. Telemedicine is not meant to replace a doctor, but rather to provide an additional method of ensuring special expertise through co-operation between doctors in the area, thereby improving the quality of treatment.

The organisational structures of telemedical collaborations in intensive care are variable. The establishment of a telemedical network can be achieved through co-operation between a tele-intensive care centre and affiliated ICUs **(hub-and-spoke structure**) or as a collaboration between equal ICUs. The most common form in Germany is currently the hub-and-spoke structure (Fig. 1).

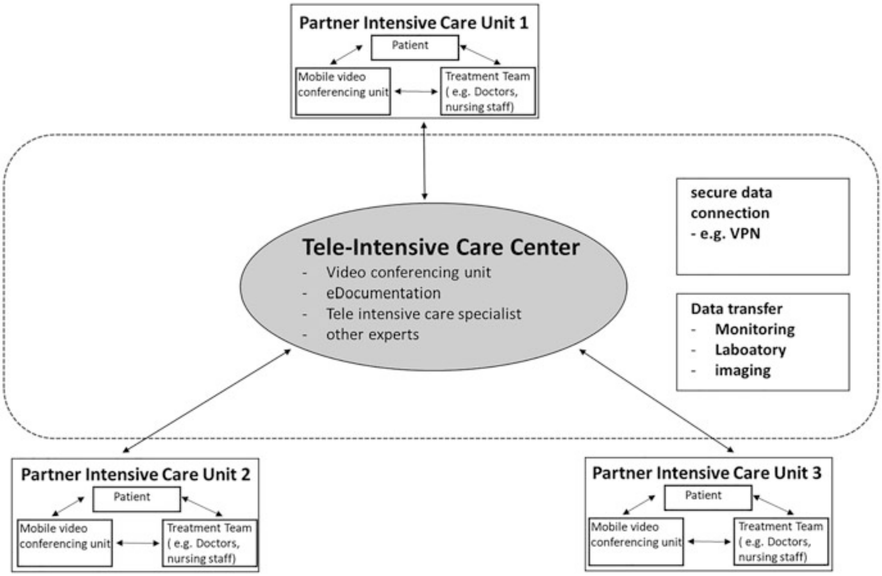


Fig. 1 Hub-and-spoke structure of a tele-intensive network

The core of telemedical collaborations is regular, daily, cross-institutional **tele-ward rounds** in the form of audio-video conferences, during which any vital parameters are transmitted in real time and electronic documents (e.g. laboratory data, imaging and doctor's letters) are transferred via a secure data-exchange platform. The rounds take place just like in the conventional context, as a conversation between the on-site treatment team and the tele-intensive care specialists directly at the patient's bedside. Apart from exchange of experience and the dual-control principle, a crucial point is the standardisation of the ward rounds, with daily systematic screening for severe infections and lung failure, and with earlier optimisation of sepsis therapy or adaptation of ventilation that is gentle on the lungs.

As far as possible, the patient is actively involved in the ward rounds and is continually informed about the course of treatment as well as any upcoming diagnostic or therapeutic measures. At the patient's request, relatives can also take part in the rounds. The daily rounds are supplemented by additional consultations in the case of acute problems or by ad hoc consultations in the event of special clinical questions. In addition, specialists from different disciplines (e.g. pharmacologists and infectiologists) can complement the treatment team and participate in the rounds in an advisory capacity.

The data transmission in telemedical collaborations can be asynchronous or synchronous with patient contact, ward rounds, or consultations: in the case of asynchronous data transmission (store-and-forward concept), findings are stored and, if required, retrieved and evaluated later (e.g. during joint rounds). The increasingly frequent use of synchronous data transmission **(real-time concept**) enables the continuous transmission of vital parameters and monitoring and laboratory data in real time, thus enabling close monitoring of the patient. Monitoring algorithms and early warning systems can also be used for early intervention (Deisz and Marx 2016).

3 Effects of Tele-Intensive Care

Telemedical collaborations can ensure the provision of care in rural areas and improve the quality of treatment over the long term through co-operation between facilities. With the technical means of tele-intensive care, the specialised expertise and experience of a centre can be made available even in remote regions, and around the clock, thus supporting care of the patient close to home.

3.1 Influence on Mortality and Length of Stay

Numerous prospectively planned longitudinal studies compare mortality and length of stay in hospital and ICU in a pre-post design. The inherent methodological weakness of a purely pre-post comparison and the difficult comparability of many studies should be noted, because randomised controlled studies at the patient level are not feasible.

More than 15 observational studies and meta-analyses show a positive effect of tele-intensive care on mortality and length of stay (Kruklitis et al. 2014; Cummings et al. 2007; Breslow et al. 2004; Kohl et al. 2012; Zawada et al. 2009; Ruesch et al. 2012; Sadaka et al. 2013; Rosenfeld et al. 2000; Kahn et al. 2016; Wilcox and Adhikari 2012; Young et al. 2011).

In 2010, Lilly et al. found a **reduction in mortality and length of stay** in 28,429 patients in 26 ICUs as a result of tele-medical interventions (Lilly and Thomas 2010). In the largest multi-centre study to date with 118,990 patients in 56 ICUs in 32 hospitals, the authors also investigated reduced hospital mortality and ICU mortality as well as reduced length of stay in the ICU and hospital after 3 years (Lilly et al. 2014).

In this study, the following individual factors were identified that were associated with the positive treatment outcome:

- More frequent case discussions with an intensive care specialist within one hour of the patient's admission to the ICU
- Earlier analysis and implementation of clinical key figures
- Higher adherence to treatment guidelines
- Quicker response to alarms.

In a pre-post comparison, McCambridge et al. demonstrated a reduction in hospital mortality from 21.4% to 14.7%, as well as a 29.5% reduction in

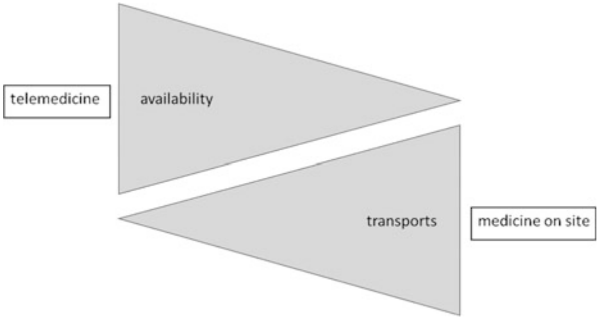


Fig. 2 Exemplary presentation of the interdependencies of telemedical and conventional care structures

risk-adjusted mortality (McCambridge et al. 2010). The authors explain these results, among other things, by the continuous availability of a tele-intensive care specialist in the ICU. In the control group, only core working hours were covered, with an additional call-out service.

Nassar et al. investigated the effect of telemedical supplementary care dependent on structural characteristics and disease severity of the patients. The authors suspected that certain facilities may benefit little from telemedical care because of a highly favourable starting position (e.g. the patients being treated have only a low disease severity). However, according to the authors, adapted care concepts such as coverage of peak workloads, weekends, and night shifts could also offer further advantages for these institutions (Nassar et al. 2014).

A recent meta-analysis confirms the positive effect of tele-intensive care on ICU mortality. Based on the 13 studies examined, it was shown that telemedical care reduced ICU mortality (odds ratio [OR] 0.75; 95% CI 0.65–0.88, $p < 0.001$) (Fusaro et al. 2019).

In addition, the number of **instances of transfer transport** was significantly reduced by telemedical care (Fig. 2). This effect has been demonstrated both for patients with a moderate and a high disease severity as well as for non-surgical patients (Fortis et al. 2018). This is particularly relevant because certain patient groups do not benefit from being transferred. In patients with severe sepsis and septic shock, Faine et al. demonstrated that transport delayed adequate therapy (Faine et al. 2015).

3.2 Influence on the Implementation of Treatment Guidelines

The positive treatment results achieved by tele-intensive care can be attributed, among other things, to **improved treatment quality** in the ICU as well as a **reduction in complications and treatment errors** (Cummings et al. 2007; Kohl et al. 2012; Rosenfeld et al. 2000).

Improved adherence to guidelines has been noted for prophylaxis of deep vein thrombosis, prevention of stress ulcers and cardiovascular protection, as well as prevention of ventilator-associated pneumonia and catheter-associated bloodstream infections (Lilly et al. 2011). Furthermore, telemedical ward rounds increased the rate of daily breaks in sedation by optimising sedative medication and through adherence to sedation protocols, facilitating coordination with spontaneous breathing trials (Forni et al. 2010). Several studies determined a positive impact on ventilation with improved adherence of lung-protective ventilation with low tidal volumes (Peterson et al. 2014; Kalb et al. 2014). Both ventilation time and the rate of ventilation were also reduced (McCambridge et al. 2010). The reduction of ventilator-associated pneumonia (Ruesch et al. 2012) led to a reduction in mortality in the ICU (Kalb et al. 2014).

Through telemedicine, specialists not available on site can be called in to complement the treatment team and improve patient safety. The additional involvement of a clinical pharmacologist in telemedical rounds led to more frequent detection of drug interactions and the reduction of prescription errors (Cummings et al. 2007; Amkreutz et al. 2020). Clinical effectiveness was enhanced through tele-intensive care and telepharmacy (Breslow 2007). Overall, tele-intensive care promoted the standardisation of processes and increased compliance with evidence-based medicine, thus positively influencing patient safety (Fuhrman and Lilly 2015).

3.3 Influence on Sepsis Mortality

Every year, more than two million people are treated in the ICU in Germany, about one third of them in university hospitals or maximum-care hospitals. Of the patients treated, 11% develop severe sepsis; in Germany, this is associated with a hospital mortality rate of 30–50% and costs of up to €4.5 billion. With approximately 75,000 deaths annually, sepsis is the third most common cause of death in Germany (Engel et al. 2007; SepNet Critical Care Trials Group 2016). Early diagnosis and guideline-based treatment can significantly improve survival (Damiani et al. 2015).

Promising projects in the field of tele-intensive care have also been carried out in Germany in recent years. The expansion of existing projects and nationwide networking are milestones on the way to introducing tele-intensive care into standard care. Germany’s first tele-intensive care project ‘**Telematics in Intensive Care Medicine**’ (TIM) has been carried out since October 2012 under the leadership of the University Hospital RWTH Aachen in co-operation with several ICUs in the region. The pilot project funded by the state of North Rhine-Westphalia successfully demonstrated that tele-intensive care is also feasible in Germany and that the care of critically ill patients can be significantly improved (Marx et al. 2015; Deisz et al. 2019).

This multi-centre observational study, the first to date in Germany, investigated the influence of telemedical care on the **treatment of sepsis** (blood poisoning) in 1168 patients. Sepsis is a medical emergency that requires immediate and targeted therapy. The guidelines of the Surviving Sepsis Campaign (SSC) recommend implementing defined sets of measures within specified time intervals (after 3 h and after 6 h). These include taking blood samples and testing for infections with bacteria or fungi, as well as administering life-saving antibiotics. Through daily tele-ward rounds, the recommended bundles of measures were implemented much more frequently: the implementation of the recommended measures within 3 h after diagnosis increased from 35% to 76.2%. The implementation of the package of measures within 6 h increased from 50% to 95.2%. Mortality fell from 50% in the first quarter to 33.3% in the sixth quarter. Thus, tele-ward rounds in the ICU were able to significantly improve survival for patients with sepsis (Deisz et al. 2019).

3.4 Algorithms and Early Warning Systems

A valuable addition to telemedical care is provided by **algorithm-based monitoring systems** for the safe and early detection of organ dysfunction. Immediate diagnosis and guideline-based therapy are essential here for the treatment outcome and thus the survival of the patient. Herasevich et al. studied the accuracy of detection of acute respiratory failure in 3795 critically ill patients in nine multidisciplinary ICUs. Bedside doctors detected only 86 (26.5%) out of 325 patients who developed acute respiratory failure. The lack of implementation of the recommended lung-protective ventilation strategy with low ventilation volumes was problematic (median tidal volume 9.2 ml/kg compared with 8.0 ml/kg; $p = 0.001$). On the other hands, computer-assisted detection of acute respiratory failure showed a very good sensitivity of 96% (Herasevich et al. 2009). Such algorithm-based monitoring therefore leads to diagnosis being able to be made earlier and more frequently, and survival is improved through the subsequent implementation of an adapted ventilation strategy that is gentle on the lungs.

The use of telemedical **early warning systems** is also possible in the outpatient setting and was recently investigated in patients with chronic cardiac insufficiency at the Charité. As part of the study, daily, non-invasively measured vital signs (ECG monitoring, oxygen saturation, blood pressure and body weight) were transmitted to a telemedical centre and monitored around the clock by a team of doctors and nurses. The early detection of acute decompensation of cardiac insufficiency and timely adjustment of medication dispensed with the need for hospitalisation and improved the survival chances of patients (Koehler et al. 2018).

Previous telemedicine studies with invasive monitoring demonstrated a positive effect on the quality of treatment for patients with chronic cardiac insufficiency. In the CHAMPION study from 2011, patients were implanted with a telemedical pressure sensor in the pulmonary artery. By controlling medicinal therapy based on the pressure values transmitted, it was possible to reduce the rehospitalisation rate by over 30% (Abraham et al. 2011). The IN-TIME study based on the daily telemonitoring of data from implanted pacemakers or defibrillators demonstrated a reduction in mortality (Hindricks et al. 2014). The early use of telemedical monitoring can thus reduce the rate of patients potentially requiring intensive care.

3.5 Tele-Emergency Medicine

In the pre-clinical sector, the tele-emergency system developed in Aachen has been successfully used in the rescue service since 2014. The rescue team on site has the option of connecting to the tele-emergency doctor in the tele-emergency doctor centre ‘at the push of a button’ via a mobile communication and transmission unit. Communication takes place via an audio or, if required, video connection. In addition to the transmission of all vital data in real time, the transmission of images (e.g. medication plans, doctor’s letters) is also possible via smart phone. The two main advantages of the system are a reduction in the number of emergency doctor interventions in cases of diagnostic and therapeutic uncertainty for non-life-threatening diseases and shortening of the interval without therapy or a doctor in the case of life-threatening emergencies. In the case of non-life-threatening diseases, the tele-emergency doctor can delegate medical measures to support the non-medical rescue service personnel. In acute emergencies, the time until the arrival of the emergency doctor can be used meaningfully (Brokmann et al. 2017).

In the case of acute ST elevation myocardial infarction (STEMI), pre-clinical ECG transmission has been shown to significantly reduce ‘door-to-reperfusion time’ and thus to accelerate therapy, with an improved treatment outcome (Adams et al. 2006). A meta-analysis showed that telemedical ECG diagnostics could even almost halve the time to treatment of acute ST elevation myocardial infarction (Brunetti et al. 2017).

3.6 Tele-Stroke Care

In stroke care, several promising pilot projects have been carried out in recent years due to insufficient care in rural areas. The TEMPiS project (Telemedical Pilot Project for Integrated Stroke Care in the South-East Bavarian Region) was incorporated into standard care in as early as 2006. Upon arrival of a patient with an acute stroke in the emergency room, stroke experts are immediately connected via an audio-video conference. The telemedical doctor can question the patient directly and examine him or her neurologically together with the doctor providing treatment on site. At the same time, the computed tomography images are sent to the telemedical centre for assessment. This way, a decision as to whether lysis therapy is indicated for the patient can be made as quickly as possible (Völkel et al. 2017).

Care in the TEMPis network was associated with improved quality of treatment as well as a more favourable prognosis (mortality, living situation, degree of disability) for the patients. After 3 months and 1 year, the probability of a poor treatment outcome (death, institutionalised care or severe need for assistance) was 38% lower than in the comparison hospitals. The lysis rates increased significantly (Audebert et al. 2006). The STENO project (Stroke Network with Telemedicine in

Northern Bavaria) also showed a significant increase in lysis rates within 3 years (from 8.2% to 12.8%) (Handschu et al. 2014).

3.7 Smart Phone Apps

Another interesting development is constituted by smart phone apps for a wide variety of medical fields. The transitions between health and lifestyle apps are fluid. Although the quality and usefulness of these apps cannot be assessed here because of the diverse areas of application, the smart phone could certainly offer a useful application in prevention, the treatment of chronically ill patients, or extended follow-up care after a hospital stay.

After intensive care treatment, some patients develop post-intensive-care syndrome (PICS), which is often not recognised and thus not treated or rehabilitated. The syndrome, which includes cognitive and physical changes such as neuromuscular dysfunction and a reduction in strength, means a significant reduction in the quality of life for these patients and their families (Rawal et al. 2017).

Although follow-up by the general practitioner or specialist physician should be a matter of course and is essential, smart phone apps in combination with audio-video conferencing and electronic documentation could be a first step in systematically questioning patients and maintaining contact with the treatment team and specialists. Improved diagnosis and active involvement of the patient in the exchange of medical data (patient empowerment) could further improve the quality of treatment.

Complementary telemedical care can thus contribute to high-quality care and the best possible treatment outcome in many areas. Both the monitoring of chronically ill patients at home and tele-medical support in acute emergencies can, in the best case, reduce severe disease progression through early detection and needs-based therapy. In the ICU, promoting guideline-based treatment can positively affect mortality and complication rates.

4 Legal and Organisational Framework Conditions

4.1 E-Health Act

An important step towards strengthening telemedical services is the act on secure digital communication and applications in health care, which came into force on 29 December 2015 (the **E-Health Act**). The aim of the act is to ensure secure, digital communication between doctors, patients, hospitals and health insurers. The law sets out a concrete roadmap for the nationwide **introduction of telematics infrastructure** and contains a bundle of incentives, deadlines and sanctions (Bundesgesundheitsministerium 2018). The inclusion of telemedical services and the facilitation of intersectoral networking promotes the expansion of telemedical networks.

4.2 Prohibition of Remote Treatment

For years, the **prohibition on remote treatment** according to Section 7, paragraph 4 of the model professional code of conduct for doctors was seen as an obstacle to the further expansion of tele-medical collaborations. However, the prohibition explicitly referred to the exclusion of telemedical treatment via the Internet or print media. This exclusionary framework condition that justifies prohibition applies neither to the tele-emergency doctor nor to tele-intensive care as a doctor-to-doctor collaboration, because either a health professional or a doctor is in contact with the patient.

This view was not only confirmed for the first time in the explanations of the Federal Medical Association from 11 December 2015, but also extended there to other health professions in the interests of clarification.

With the resolution of the 121st Medical Congress in 2018 to expand the treatment principles for remote treatment and to relax the prohibition on remote treatment, legal certainty was created in this field (Bundesärztekammer 2018; ZTG Zentrum für Telematik und Telemedizin GmbH 2018).

4.3 Responsibility for Treatment

Like in the traditional treatment context, the **responsibility for treatment** remains with the local doctor with whom a concrete treatment relationship exists. This essential question of responsibility for treatment is also preserved by the consultative nature of tele-intensive supplementary care.

4.4 Data Protection

By definition, all health data is sensitive data requiring special protection. Their collection, processing and use therefore requires special care (e.g. a secure IT infrastructure with end-to-end encryption or VPN connections for data transmission between the ICUs). In addition, organisational and contractual arrangements must be put in place. Hospital admission contracts between hospital providers and patients must be adapted in an appropriate form so that patients can either consent to or refuse telemedical collaborations. Depending on the technical implementation, consent for commissioned data processing may also be required. Furthermore, contractual regulations for data collection and data use should also exist between the ICUs involved.

4.5 Structural Recommendations of the DGAI

Telemedicine is not a means of circumventing reasonable standards of care but rather a means of ensuring a consistently high quality of care. The German Society for Anaesthesia and Intensive Care Medicine (DGAI) has therefore defined the framework conditions for the transfer of existing telemedical projects into standard care by founding a permanent commission on telemedicine and **drawing up structural recommendations** (Marx and Koch 2015).

These structural recommendations are a decisive step towards ensuring high structural and process quality because they not only define minimum standards on the part of the participating intensive care units but also make clear statements on the supply side regarding the qualifications and availability of the telemedical centres.

5 Outlook

5.1 Robotics as Assistance in Care

The development of robotics is advancing rapidly in many areas of medicine. Robot-assisted surgical systems that are a further development of traditional laparoscopies have already been firmly established for more than 20 years. For rehabilitation after an acute stroke, for example, robotic-assisted rehabilitation procedures can also be used to help patients relearn movement sequences with high intensity, or robotic exoskeletons can be used in gait training for spinal cord injuries.

In the area of inpatient and outpatient care, the problems of demographic change and the shortage of care are driving the development of care robots. Care robots support or replace individual tasks of human caregivers. For example, they can bring medication and food to the patient's bedside or help with positioning or mobilising patients. Some robots have natural language abilities or are capable of learning and intelligent to a certain extent. So far, mainly prototypes of these robots are in use in test environments or at trade fairs. Examples of these developments include the Care-O-bot 4 (Fraunhofer IPA), which can fetch and take away objects and move safely in spaces occupied by other people, the TUG (Aethon), which transports medications and materials, can ride the lift independently, and has natural language abilities, the HOBBIT (TU Vienna, EU project), which is designed with a friendly face and is intended to help older people by increasing their sense of security or picking up objects from the floor, the robot Cody (Georgia Institute of Technology), which can move and wash bedridden patients, and the Robear (Riken), as a successor to RIBA and RIBA-II, which assists caregivers in transferring patients and sitting them upright (Bendel 2018). Further advantages of nursing robots are their continuous availability, the consistent quality of the service, and relief of the nursing staff in certain ranges of duties. Disadvantages are the currently high costs and the loss of interpersonal contact. A milestone would be the development of multifunctional robots that can provide assistance in the care of elderly and impaired people by adapting to changing situations.

5.2 Intelligent Networking in the Health Care System

An important task of the health care system is to digitally network the existing decentralised structures across sectors. With the E-Health Act, the legislator has set a concrete schedule for the nationwide development of a secure digital infrastructure for the health care system. Doctors, dentists, psychotherapists, hospitals and pharmacies are to be successively connected to this telematics infrastructure. Within the network, all participating institutions can communicate securely and quickly with each other. If the patient has given his or her consent, emergency medical data can be accessed. The comprehensive and complete documentation of disease patterns and therapies is also advantageous. With the patient's consent, his or her medical history can be viewed by the doctors providing treatment and other health professionals without them having to first form their own picture from the multitude of old doctor's letters. However, the use of the telematics infrastructure depends on acceptance by all parties involved. This acceptance can become established among patients and doctors, in particular, through a benefit that is recognisable to them as well as the usability of the applications. Important technical cornerstones for intersectoral digital networking include the expansion of cloud services and the interoperability of the IT systems involved.

For the field of intensive care medicine, the inter-clinic monitoring of patients would be another milestone based on this intersectoral networking. Treatment data are already recorded in a **patient data management system** (PDMS) and kept for further use (e.g. treatment planning, billing). However, it is currently not possible to obtain an overview of a large group of patients because the different PDMS systems are not interoperable, and the structured data cannot be used for communication or decision support across system or facility boundaries. A central telemedical platform for linking different PDMS systems would be desirable. A **cockpit solution** (Fig. 3) enabled a team of intensive care specialists to identify deterioration in the conditions of a large group of critically ill patients at an early stage and to initiate treatment as quickly as possible in accordance with guidelines. Here, linking with algorithm-based monitoring and decision support systems for automated detection of organ dysfunction is essential. By means of '**algorithmic surveillance**', data is analysed by high-performance computing, thereby creating the prerequisite for early alerting, which enables rapid diagnostic and therapeutic intervention.

The development of the **Internet of Medical Things** (IoMT) goes beyond the networking of patient-specific treatment data and refers to medical devices and applications that are connected via a cloud.

The principle of the Internet of Things (IoT) is to network physical and virtual objects and make them work together through information and communication technologies. The goal is to automatically capture information from the real world,

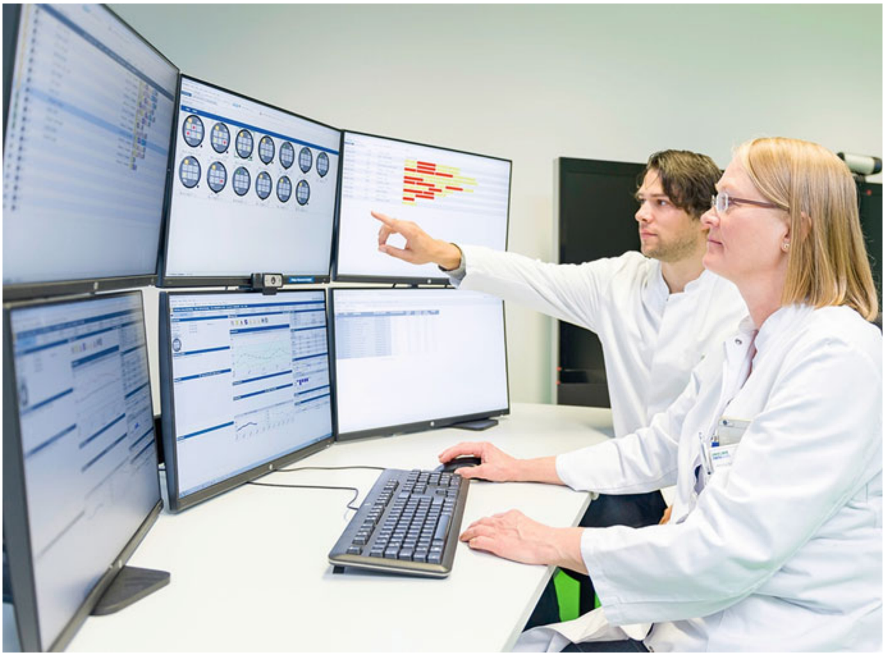


Fig. 3 Presentation of a cockpit solution for tele-intensive care monitoring of a larger group of patients

link it together, and make it available on the network. One example of industrial use is the monitoring of large-scale systems, which allows operators to carry out predictive maintenance on individual devices. In production, IoT technologies enable the seamless tracking of goods in the manufacturing process.

In medicine, the use of **wearables** in which miniaturised computers with various sensors are incorporated directly into items of clothing, watches or jewellery lends itself particularly well. Patients with chronic diseases could thus be continuously monitored and receive seamless therapy. Examples of current developments are automated blood pressure measurement, smart contact lenses for blood glucose measurement, and continuous heart rate and ECG recording (Dimitrov 2016). If an abnormal value is registered, the patient could ideally receive a warning directly and be reminded to take their medication. In the case of critical values, a telemedicine specialist could be called in directly, or an emergency call could be made.

5.3 Use of Artificial Intelligence

New scientific findings in medicine are multiplying at a rapid rate. It is estimated that medical knowledge will double every 73 days in 2020 (Densen 2011). This makes it difficult or impossible for doctors to keep up with the latest scientific findings. The solution lies in computer-based systems that collect and analyse the data and, for example, make them available in the specific patient or treatment context. In addition, **artificial intelligence** and **decision-support systems** will become of great interest to doctors and patients.

The combination of published knowledge with Big Data and Big Data analyses offers the possibility of analysing the individual characteristics of patients together with information from the electronic patient record, as well as general information such as guidelines, and thus being able to match diverse patterns. This makes it possible to achieve more personalised care. To make use of these advantages, the necessary digital infrastructures must be created. In the field of intensive care medicine, through targeted interoperability projects such as the SMITH project (www.smith.care), the **Medical Information Technology Initiative of the Federal Ministry of Education and Research** will contribute to connecting local digital documentation systems and making them usable to improve health care in Germany (Winter et al. 2018).

Clinical decision-making support systems and artificial intelligence can assist doctors in making diagnoses or treatment decisions. Here, case-specific patient data such as symptoms and laboratory data are introduced into a system and matched with diseases within the interference engine of the system. This generates possible diagnoses or detects changes earlier. In the field of radiology, automated analysis of the results of imaging procedures (X-ray, CT, and MRI) will in future enable reliable diagnoses to be made by computer programs.

The use of decision-making support systems can also make it easier for patients themselves to interpret health data they have themselves collected. Sensors in wearables such as watches, rings or clothing can be used to obtain data on heart rate, blood pressure, ECG, EEG, body temperature, oxygen saturation and physical activity. Special parameters (e.g. blood sugar, blood pressure in the pulmonary circulation or intra-ocular pressure) can be determined via insidables (i.e. implanted sensors).

The automated transmission of this data in combination with audio and video conferencing with the GP or a specialist enables close monitoring of the health status. The early detection of deteriorating conditions and the immediate initiation of an appropriate therapy can immediately improve patient care.

The **smart4health** project (www.smart4health.eu) is therefore developing a citizen-centred patient record based on an EU initiative on interoperable electronic patient records. This is not only interoperable from a technical point of view, but also considers the cross-border mobility of EU citizens through the option of automated translation. This translation function makes relevant treatment information from the citizen’s home country readable and usable for health care there, even during stays abroad.

6 Conclusion

The digitisation of the health system and telemedical collaborations offer enormous potential. Improved diagnosis and targeted therapy can measurably improve the quality of treatment by reducing complications or increasing survival rates and thus improving the quality of life of patients. Increased availability of medical expertise—also in rural areas—can ensure comprehensive medical care and treatment close to home for all patients—also given the challenges of demographic change. The autonomy of patients can be strengthened through direct access to their health data as well as more intensive involvement in treatment. In addition, economic advantages can result from more efficient use of resources.

However, the success of telemedical applications stands and falls with its acceptance among users. There must be a clear benefit for patients and doctors. This is illustrated by improved treatment quality and nationwide care as well as needs-based support in everyday clinical practice. It is essential to ensure data protection. In addition, innovative technologies must be designed to be user-friendly and function without interference. Finally, the remuneration of telemedical services must be regulated by law to ensure permanent use beyond the project stage.

The central task of medical practitioners, engineers and computer scientists in implementing the digital revolution is to increase the **benefit for the patient**. Even the early detection of deteriorating conditions in the outpatient setting can prevent hospitalisation and improve quality of life for patients by reducing inpatient hospitalisation. In the ICU, early warning systems and the availability of specialised expertise through telemedical collaborations reduce mortality and contribute to the best possible treatment outcome. In all these measures, sight must not be lost of central aspects of the doctor–patient relationship, such as personal consultation, because they help to establish and maintain the relationship of trust between the patient and the doctor. To a greater degree, telemedical procedures are intended to provide a meaningful and supportive addition to everyday clinical practice.

Finally, it should be emphasised that **patient autonomy** and digitisation are not contradictory; rather, the introduction of electronic documentation gives patients improved access to their own treatment data. The monitoring of one's own health parameters and the use of health apps contribute to improved information about one's own disease and enable needs-based therapy. In addition, improved availability of medical advice through tele-ward rounds and the possibility of resolving outstanding issues or problems in a timely manner can increase patient satisfaction.

In conclusion, medicine and technology must continually work to achieve the best possible benefit for the patient by improving existing processes. The further expansion of tele-medical collaborations and integration of new technologies can make a valuable contribution to this.

References

Abraham WT et al (2011) Wireless pulmonary artery haemodynamic monitoring in chronic heart failure: a randomised controlled trial. Lancet 377(9766):658–666

Adams GL et al (2006) Effectiveness of prehospital wireless transmission of electrocardiograms to a cardiologist via hand-held device for patients with acute myocardial infarction (from the Timely Intervention in Myocardial Emergency, NorthEast Experience [TIME-NE]). Am J Cardiol 98(9):1160–1164

Amkreutz J et al (2020) Medication safety in a German telemedicine centre: implementation of a telepharmaceutical expert consultation in addition to existing tele-intensive care unit services. J Telemed Telecare 26(1–2):105–112

Audebert HJ et al (2006) Effects of the implementation of a telemedical stroke network: the Telemedic Pilot Project for Integrative Stroke Care (TEMPiS) in Bavaria, Germany. Lancet Neurol 5(9):742–748

Bendel O (2018) Roboter im Gesundheitsbereich. In Pflegeroboter, pp 195–212

Bölt U, Graf T (2012) Statistisches Bundesamt. 20 Jahre Krankenhausstatistik. https://www.destatis.de/DE/Publikationen/WirtschaftStatistik/Gesundheitswesen/20JahreKrankenhausstatistik.pdf?__blob=publicationFile. Accessed 12 Feb 2019

Breslow MJ (2007) Remote ICU care programs: current status. J Crit Care 22(1):66–76

Breslow MJ et al (2004) Effect of a multiple-site intensive care unit telemedicine program on clinical and economic outcomes: an alternative paradigm for intensivist staffing. Crit Care Med 32(1):31–38

Brokmann J et al (2017) Telemedizin: Potenziale in der Notfallmedizin. Notfallmedizin up2date 12(03):307–317

Brunetti ND et al (2017) Pre-hospital electrocardiogram triage with telemedicine near halves time to treatment in STEMI: a meta-analysis and meta-regression analysis of non-randomized studies. Int J Cardiol 232:5–11

Bundesärztekammer (2014) Ergebnisse der Ärztestatistik zum 31.12.2014. https://www.bundesaerztekammer.de/ueber-uns/aerztestatistik/aerztestatistik-der-vorjahre/aerztestatistik-2014/. Accessed 12 Dec 2018

Bundesärztekammer (2018) 121. Deutscher Ärztetag. Beschlussprotokoll. Erfurt, 8. bis 11. Mai 2018. https://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/121.DAET/121_Beschlussprotokoll.pdf. Accessed 12 Feb 2019

Bundesgesundheitsministerium (2018) E-Health-Gesetz. https://www.bundesgesundheitsministerium.de/service/begriffe-von-a-z/e/e-health-gesetz.html. Accessed 12 Feb 2019

Cummings J et al (2007) Intensive care unit telemedicine: review and consensus recommendations. Am J Med Qual 22(4):239–250

Damiani E et al (2015) Effect of performance improvement programs on compliance with sepsis bundles and mortality: a systematic review and meta-analysis of observational studies. PloS One 10(5):e0125827

Deisz R, Marx G (2016) Telemedizin in der Intensivmedizin – Möglichkeiten und Grenzen einer Innovation. Medizinische Klinik, Intensivmedizin und Notfallmedizin 111(8):723–728

Deisz R et al (2019) Additional telemedicine rounds as a successful performance-improvement strategy for sepsis management: observational multicenter study. J Med Internet Res 21(1): e11161. https://doi.org/10.2196/11161

Densen P (2011) Challenges and opportunities facing medical education. Trans Am Clin Climatol Assoc 122:48–58

Dimitrov DV (2016) Medical Internet of Things and big data in healthcare. Healthcare Inf Res 22(3):156–163

Engel C et al (2007) Epidemiology of sepsis in Germany: results from a national prospective multicenter study. Intensive Care Med 33(4):606–618

Faine BA et al (2015) Interhospital transfer delays appropriate treatment for patients with severe sepsis and septic shock: a retrospective cohort study. Crit Care Med 43(12):2589–2596
Forni A et al (2010) Evaluation of the impact of a tele-ICU pharmacist on the management of sedation in critically ill mechanically ventilated patients. Ann Pharmacother 44(3):432–438
Fortis S et al (2018) ICU telemedicine reduces interhospital ICU transfers in the veterans health administration. Chest 154(1):69–76
Fuhrman SA, Lilly CM (2015) ICU telemedicine solutions. Clin Chest Med 36(3):401–407
Fusaro MV, Becker C, Scurlock C (2019) Evaluating tele-ICU implementation based on observed and predicted ICU mortality. Crit Care Med 47(4):501–507
Handschu R et al (2014) Feasibility of certified quality management in a comprehensive stroke care network using telemedicine: STENO project. Int J Stroke 9(8):1011–1016
Herasevich V et al (2009) Validation of an electronic surveillance system for acute lung injury. Intensive Care Med 35(6):1018–1023
Hindricks G et al (2014) Implant-based multiparameter telemonitoring of patients with heart failure (IN-TIME): a randomised controlled trial. Lancet 384(9943):583–590
Kahn JM et al (2016) ICU telemedicine and critical care mortality: a national effectiveness study. Med Care 54(3):319–325
Kalb T et al (2014) A multicenter population-based effectiveness study of teleintensive care unit-directed ventilator rounds demonstrating improved adherence to a protective lung strategy, decreased ventilator duration, and decreased intensive care unit mortality. J Crit Care 29(4) 691.e7–e14
Karagiannidis C et al (2019) Intensivmedizin: Versorgung der Bevölkerung in Gefahr. Dtsch Arztebl 116(10):A-462/B-378/C-374
Koehler F et al (2018) Efficacy of telemedical interventional management in patients with heart failure (TIM-HF2): a randomised, controlled, parallel-group, unmasked trial. Lancet 392(10152):1047–1057
Kohl BA et al (2012) The effect of ICU telemedicine on mortality and length of stay. J Telemed Telecare 18(5):282–286
Krüger-Brand HE (2019) Digitale Gesundheitskompetenz: Datensouveränität als Ziel. Dtsch Arztebl 116(10):A-468/B-382/C-378
Kruklitis RJ, Tracy JA, McCambridge MM (2014) Clinical and financial considerations for implementing an ICU telemedicine program. Chest 145(6):1392–1396
Lilly CM, Thomas EJ (2010) Tele-ICU: experience to date. J Intensive Care Med 25(1):16–22
Lilly CM et al (2011) Hospital mortality, length of stay, and preventable complications among critically ill patients before and after tele-ICU reengineering of critical care processes. JAMA 305(21):2175–2183
Lilly CM et al (2014) A multicenter study of ICU telemedicine reengineering of adult critical care. Chest 145(3):500–507
Marx G, Beckers R (2015) Telemedizin in Deutschland. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz 58(10):1053–1055
Marx G, Koch T (2015) Telemedizin in der Intensivmedizin: Strukturempfehlungen der DGAI. Anästhästhesiologie & Intensivmedizin 56:257–261
Marx G et al (2015) Telekooperation für die innovative Versorgung am Beispiel des Universitätsklinikums Aachen. Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz 58(10):1056–1061
McCambridge M et al (2010) Association of health information technology and teleintensivist coverage with decreased mortality and ventilator use in critically ill patients. Arch Intern Med 170(7):648–653
Nassar BS et al (2014) Impact of an intensive care unit telemedicine program on patient outcomes in an integrated health care system. JAMA Intern Med 174(7):1160
Peterson L et al (2014) Telemedicine improves hospital system compliance with lung-protective ventilating strategies. Crit Care Med 42:A1555

Rawal G, Yadav S, Kumar R (2017) Post-intensive care syndrome: an overview. J Transl Intern Med 5(2):90–92
Rosenfeld BA et al (2000) Intensive care unit telemedicine: alternate paradigm for providing continuous intensivist care. Crit Care Med 28(12):3925–3931
Ruesch C et al (2012) Using nursing expertise and telemedicine to increase nursing collaboration and improve patient outcomes. Telemed J E Health 18(8):591–595
Sadaka F et al (2013) Telemedicine intervention improves ICU outcomes. Crit Care Res Pract 2013: 456389
SepNet Critical Care Trials Group (2016) Incidence of severe sepsis and septic shock in German intensive care units: the prospective, multicentre INSEP study. Intensive Care Med 42(12): 1980–1989
Völkel N, Hubert GJ, Haberl RL (2017) Telemedizin: “TEMPiS”-Netzwerk für Schlaganfallbehandlung in der Regelversorgung. Anaesthesiologie, Intensivmedizin, Notfallmedizin, Schmerztherapie: AINS 52(2):127–136
Wilcox ME, Adhikari NKJ (2012) The effect of telemedicine in critically ill patients: systematic review and meta-analysis. Crit Care 16(4):R127
Winter A et al (2018) Smart Medical Information Technology for Healthcare (SMITH). Methods Inf Med 57(S 01):e92–e105
Young LB et al (2011) Impact of telemedicine intensive care unit coverage on patient outcomes: a systematic review and meta-analysis. Arch Intern Med 171(6):498–506
Zawada ET Jr et al (2009) Impact of an intensive care unit telemedicine program on a rural health care system. Postgrad Med 121(3):160–170
ZTG Zentrum für Telematik und Telemedizin GmbH (2018) Lockerung des Fernbehandlungsverbots – Hintergrundgespräch mit den Ärztekammern in Nordrhein-Westfalen. https://www.ztg-nrw.de/2018/05/lockerung-des-fernbehandlungsverbots-hintergrundgespraech-mit-den-aerztekammern-in-nordrhein-westfalen/. Accessed 7 Jan 2019

Clinic 4.0: The Digital Hospital

Christian Juhra

Contents

1 Clinic 1.0 to Clinic 3.0

New technologies have led to and continue to lead to drastic changes in medical care. For example, the introduction of the stethoscope. Developed in 1816 by René Laennec in France, it made possible a new technique of auscultation and eventually replaced the ear trumpet in virtually all areas of application. The introduction of anaesthesia a few years later (around 1846) paved the way for modern medicine by making possible operations that had previously been unthinkable due to their duration and the associated pain.

At the beginning of the twentieth century, diagnostics were revolutionized using X-ray machines. And just as technology changed and evolved, so did the hospital. Until well into the nineteenth century, hospitals were essentially facilities for the lower class, while wealthier patients were treated at home. However, these

C. Juhra (✉)
Münster University Hospital, Münster, Germany
e-mail: christian.juhra@ukmuenster.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_44

technological innovations made it increasingly necessary to treat patients at a central location outfitted with the necessary equipment. Moreover, complex operations now made possible by modern anaesthesia required inpatient follow-up care.

The monitoring stations introduced by Martin Kirchner and Ferdinand Sauerbruch in the 1930s were the first of their kind; here newly operated patients could now receive intensive, post-operative care. In 1952 a polio epidemic in Denmark led to the establishment of the first respiratory care unit by Björn Ibsen (Bär 2011). At the time, ventilation was still performed manually, and the first ventilators were not to be available until a few years later. For the interested reader, the “Historia Hospitalium” series of journals by the German society for Hospital History (Deutsche Gesellschaft für Krankenhausgeschichte e.V) is recommended reading for a deeper understanding of the literature.

Innovations in information technology also influenced the development of hospitals. The invention of the computer is generally attributed to Konrad Zuse, who in 1941 developed the first programmable, electronic computer, the Z3. Through advances in technology and the shrinking size of analog switches, smaller, more powerful computers became possible. ENIAC (Electronic Numerical Integrator and Computer) required 85 cubic meters of space, and weighed almost 30 tonnes, while equipped with 1 Kbit of memory. Today, a single computer chip but a few millimetres in size can accommodate billions of transistor functions (Arnold 2010).

That computers would also find their use in healthcare was an inevitability. Peter L. Reichertz first mentioned the term “medical informatics” in an article in 1970 (Reichertz 1970), and in 1972 the first medical informatics program was offered at Heidelberg/Heilbronn. A detailed account of the history of medical informatics until 1980 can be found under (Köhler 2003).

Analogous to industrial development, Medicine 1.0 refers the introduction of new processes (anaesthesia), the introduction of new technologies (X-ray technology) is referred to as Medicine 2.0, and the introduction of information processing and storage devices as Medicine 3.0. Medicine 4.0 refers to the introduction of networking technology that enables real-time communication (Mittelstaedt 2015).

2 The Fourth Revolution in the Hospital?

What does the introduction of networked technologies mean for medicine, and especially hospitals? The aim of this integration is to make any relevant information available to the right people at the right time, so that decisive action can be taken and the proper follow-up procedures can be put into action. To illustrate this, take for example the logistics of drinking water: The average two-person household drinks three litres of water every day. So, a crate of 12 0.75 l bottles should last three days and two crates are purchased every week in the supermarket, more if there are any planned visitors. However, if any spontaneous additional demand arises, for example an unplanned visit, the water will not last until the next scheduled trip to the supermarket. Either the household takes an unscheduled trip to the supermarket or it runs low on water. What would this look like if the household was somehow linked to the supply chain? Suppose there were crates that could measure and transmit how full they were. For this example, the system would detect when the remaining amount of water is below a certain level, and then order more. The manufacturer would know when and where how much water is needed. Moreover, the delivery process could also be optimized, as it would already be known which households would need more water and when.

Modern companies such as Amazon use similar technologies to try to determine future demand for products in a given region as accurately as possible. How could hospitals benefit from networked logistics? And when does a hospital become a digital hospital?

2.1 Framework

The Healthcare Information and Management Systems Society (HIMSS) Europe has developed an Electronic Medical Record Adoption Model (EMRAM), which has been the worldwide standard since 2015. The hospital is assigned to a developmental stage according to its level of digitalization, with the highest achievable stage being stage 7. The stages defined by HIMSS (HIMSS EMRAM 2019) are as follows:

Stage 0: The organization has not installed all of the three key ancillary department systems (laboratory, pharmacy, and radiology).

Stage 1: All three major ancillary clinical systems are installed (i.e., pharmacy, laboratory, and radiology). A full complement of radiology and cardiology PACS systems provides medical images to physicians via an intranet and displaces all film-based images. Patient-centric storage of non-DICOM images is also available.

Stage 2: Major ancillary clinical systems are enabled with internal interoperability feeding data to a single clinical data repository (CDR) or fully integrated data stores that provide seamless clinician access from a single user interface for reviewing all orders, results, and radiology and cardiology images. The CDR/data stores contain a controlled medical vocabulary and order verification is supported by a clinical decision support (CDS) rules engine for rudimentary conflict checking. Information from document imaging systems may be linked to the CDR at this stage. Basic security policies and capabilities addressing physical access, acceptable use, mobile security, encryption, antivirus/anti-malware, and data destruction.

Stage 3: 50 percent of nursing/allied health professional documentation (e.g., vital signs, flowsheets, nursing notes, nursing tasks, care plans) is implemented and integrated with the CDR (hospital defines formula). Capability must be in use in the ED, but ED is excluded from 50% rule. The Electronic Medication Administration Record application (eMAR) is implemented. Role-based access control (RBAC) is implemented.

Stage 4: 50 percent of all medical orders are placed using Computerized Practitioner Order Entry (CPOE) by any clinician licensed to create orders. CPOE is supported by a clinical decision support (CDS) rules engine for rudimentary conflict checking, and orders are added to the nursing and CDR environment. CPOE is in use in the Emergency Department, but not counted in the 50% rule. Nursing/allied health professional documentation has reached 90% (excluding the ED). Where publicly available, clinicians have access to a national or regional patient database to support decision making (e.g., medications, images, immunizations, lab results, etc.). During EMR downtimes, clinicians have access to patient allergies, problem/diagnosis list, medications, and lab results. Network intrusion detection system in place to detect possible network intrusions. Nurses are supported by a second level of CDS capabilities related to evidence-based medicine protocols (e.g., risk assessment scores trigger recommended nursing tasks).

Stage 5: Full physician documentation (e.g., progress notes, consult notes, discharge summaries, problem/diagnosis list, etc.) with structured templates and discrete data is implemented for at least 50 percent of the hospital. Capability must be in use in the ED, but ED is excluded from 50% rule. Hospital can track and report on the timeliness of nurse order/task completion. Intrusion prevention system is in use to not only detect possible intrusions, but also prevent intrusions. Hospital-owned portable devices are recognized and properly authorized to operate on the network, and can be wiped remotely if lost or stolen.

Stage 6: Technology is used to achieve a closed-loop process for administering medications, blood products, and human milk, and for blood specimen collection and tracking. These closed-loop processes are fully implemented in 50 percent of the hospital. Capability must be in use in the ED, but ED is excluded from 50% rule. The eMAR and technology in use are implemented and integrated with CPOE, pharmacy, and laboratory systems to maximize safe point-of-care processes and results. A more advanced level of CDS provides for the "five rights" of medication administration and other 'rights' for blood product, and human milk administrations and blood specimen processing. At least one example of a more advanced level of CDS provides guidance triggered by physician documentation related to protocols and outcomes in the form of variance and compliance alerts (e.g., VTE risk assessment triggers the appropriate VTE protocol recommendation). Mobile/portable device security policy and practices are applied to user-owned devices. Hospital conducts annual security risk assessments and report is provided to a governing authority for action.

Stage 7: The hospital no longer uses paper charts to deliver and manage patient care and has a mixture of discrete data, document images, and medical images within its EMR environment. Data warehousing is being used to analyze patterns of clinical data to improve quality of care, patient safety, and care delivery efficiency. Clinical information can be readily shared via standardized electronic transactions (i.e., CCD) with all entities that are authorized to treat the patient, or a health information exchange (i.e., other non-associated hospitals, outpatient clinics, sub-acute environments, employers, payers and patients in a data sharing environment). The hospital demonstrates summary data continuity for all hospital services (e.g., inpatient, outpatient, ED, and with any owned or managed outpatient clinics). Physician documentation and CPOE has reached 90% (excluding the ED), and the closed-loop processes have reached 95% (excluding the ED). [HIMSS EMRAM 2019]

Currently (January 2021), only one hospital in Germany reached level 6 (Medius Klinik Nürtingen), while none reached level 7. Worldwide, 187 hospitals reached level 7 (153 in the USA, 7 in Brazil, 6 in China, 3 in Turkey, Canada and Saudi Arabia, 2 in the UK and the Netherlands, and 1 in Portugal, Thailand, and Australia) and 2346 hospitals reached level 6. However, it should be noted that the classification by HIMSS is voluntary and therefore not every hospital goes through this process. The inclined reader can make themselves a picture of which stage the hospital in which they work would reach based on the criteria above. The advantages that advanced networking could have in different areas are described in detail below.

In 2021, Germany started a new funding initiative to support new technology investments in hospitals (Krankenhauszukunftsgesetz/Hospital Future Law). Thus, hospitals shall—among other things—be able to improve their IT security as well as their electronic data exchange with other hospitals.

2.2 Logistics 4.0

Optimal provision of all manner of goods (medicine, implants, consumables, cleaning agents etc.) is a major challenge in hospitals. On the one hand, the provision of goods should be as efficient as possible (e.g., only those medications that are actually being taken by patients should be stocked). On the other hand, the need for more niche medications may spontaneously arise, medications which may rarely be used, but cannot be delayed by external shipping times when the need arises and need to be readily available. Pacemakers are another example. The implantation of a pacemaker can often be planned; however, a quick implantation may be necessary. Or maybe the device that had originally been planned to be implanted needs to be switched out for another, requiring the hospital to keep a number of devices in stock at any given time. It also needs to be possible to keep track of which patient had what device implanted.

What might this networked future look like? Take endoprosthetics for example. The installation of an artificial hip joint is often a plannable operation (in the case of joint wear) but may also be necessary in the case of an acute fracture. These prostheses are often available in a modular design, where the prosthesis consists of multiple combinable elements, each of which is available in difference sizes. For scheduled installations, this is planned in advance using special software. That way, not only is it clear what prosthesis and parts are needed, but also which should be kept on hand should the operation deviate from the plan. Afterwards it can run a check as to whether the prostheses, or in this case, all of the required parts for the modular prosthesis, are in stock at the clinic or need to be ordered in. Should the prosthesis or its components not be in stock, this needs to be accounted for when planning the operation. If the implantation is successful, replacements for any expended materials can be ordered automatically afterwards if necessary. In addition, the patient can receive an electronic implant badge in their electronic patient file containing any relevant information about the prosthesis.

2.3 Doctor 4.0

The first stroke-network in Germany (TEMPiS) was founded in 2003 in Bavaria. Since then the network has treated over 7000 patients annually. The rate of stroke-related permanent disabilities as well as the mortality rate of stoke patients could be drastically reduced (Tempis 2019). The success of this as well as other networks nationwide depends on close cooperation between hospitals of different care levels. A stoke is a time-critical emergency, where important decisions regarding treatment must be made very quickly. The diagnosis of a stroke is made using medical imaging (CT and possibly MRI) and patient examination. There are several treatment options available, some of which can only be carried out at specialized treatment facilities. However, there are also treatments options available at less specialized clinics. If one were to send all stroke patients to those specialized treatment facilities, they would become quickly overwhelmed and some patients would then have to spend unnecessary time in transit.

If a patient suffers a stroke in the area covered by the TEMPiS network, an emergency physician will admit them to the nearest in-network clinic. There, all necessary diagnostics are carried out and the nearest specialized treatment center contacted if that clinic is not one itself. The images are transmitted electronically, so with the help of a video teleconferencing solution not only the doctor at that clinic but also a specialist in the stroke center can examine the patient. The experts can then decide together as to whether or the not the patient needs to be brought to the specialized stroke center. Since 2018, trials are being conducted to determine whether specialists from the stroke center could simply be brought to the clinic where the patient is being treated via helicopter and treat the patient on site to avoid having to transfer them. The TEMPiS network, which is also one of the first recipients of the German Telemedicine Award (2009) is a sterling example of a successful networking and information exchange between hospital doctors.

Another example of successful hospital networking is the German Society for Trauma Surgery's nationwide Trauma Network initiative. Similar to a stroke, some injuries, such as serious road traffic injuries, require quick and effective treatment (Trauma Network 2019). As of January 2019, 677 clinics have joined together to form a total of 53 certified trauma networks. Each clinic is certified as a local, regional or superregional clinic depending on its treatment capabilities. Similar to the stroke networks not every patient can be treated at a superregional treatment facility. To ensure the same standard of care at each clinic, they exchange patients amongst each other. X-ray and CT images of critically ill and injured patients are sent to a superordinate hospital, decisions can be made together as to how to proceed and the patient transferred if necessary. This can be done immediately upon admission in acute cases as well as during treatment, for example, when complication arise.

Unfortunately, the current division of medical care into family doctors, specialists in smaller practices and doctors working in larger clinics often leads to gaps in information. For instance, it is not yet possible for a family doctor to send a digital copy of a patient's medical history to a hospital for the doctors there to see. If this were possible, the doctors in the hospital would be able to obtain valuable information about the patient before they had even made first contact. The usual anamnesis (previous illnesses, medication, etc.) could be considerably shortened. The same problem arises when the patient is discharged. An electronic doctor's letter which is sent directly to the patient's family doctor or specialist upon discharge, while technically possible and in some cases possibly already in use, is still not standard.

What might Medicine 4.0 look like in the coming years? Let us say Mr. Müller feels a sudden chest pain at home in the countryside. Mr. Müller calls an emergency physician who can access his data online, which he had his family doctor upload. The emergency physician then sees that Mr. Müller was treated a year ago for a heart attack at the university hospital (50 km away). He then performs an ECG and deliberates over which hospital Mr. Müller should be sent to: To the nearby district hospital, or the university hospital which is further away. He sends the ECG results to the university hospital and talks to a cardiologist there. Because of his pre-treatment, and current condition allowing for a lengthier transport time, the decision is made to send Mr. Müller to the university hospital. There everything is already being prepared for a surgical intervention. The clinic can track the ambulance carrying Mr. Müller and knows exactly when he will arrive. On arrival at the clinic Mr. Müller says that his family doctor had just done an ECG test a week ago. The ECG test results are available in his electronic patient file and the doctors in the clinic have immediate access to them. Mr. Müller is now being treated at the hospital and his family doctor is informed. The materials required for his care, such as the stents needed to improve blood flow, are automatically reordered after use, ensuring a constant supply is available at all times. Unfortunately, complications arise and Mr. Müller receives pacemaker. Here too, a replacement is ordered automatically for restock. The pacemaker is also entered into Mr. Müllers electronic patient file. After a few days in intensive care Mr. Müller's condition has improved such that his treatment can continue in the district hospital nearer his home. The university hospital informs them of all relevant findings and Mr. Müller is transferred and can be discharged in a week. Upon discharge, his family doctor receives all important information about his treatment and Mr. Müller can make an appointment for a check-up online. The medication he has to take for the next few days is ordered electronically by the hospital, and delivered to his doorstep. Since the pacemaker is also network-enabled they can check on it at any time.

The above example serves to illustrate how information gaps and the resulting wasted time can be avoided through systems integration. Unfortunately, preliminary findings such as the ECG performed the family doctor from the example are often difficult or impossible to find which results in decisions being made with incomplete information.

The increasing integration of systems in all sectors, across all boundaries, means change is coming to the medical profession. Increasing specialization means no physician today is capable on their own of offering each patient optimal care for every illness. Rather, doctors are increasing becoming team players who take joint care of the patient. These teams are always different depending on the patient and illness, and their compositions change as the illness progresses. However, these teams can only function with a leader. In the future, the family doctor's role will increasingly consist of putting together an individual network of medical experts and healthcare professionals for their patients which they can adapt dynamically according to the patient's needs. Such networks will be far more than just networks of doctors. Finally, it is also essential to integrate the patient's nursing care into their network. Nursing care will therefore also change considerably as networking expands both in and outside of hospitals.

The current COVID pandemic has led to a huge increase in the uptake of telemedicine in Germany and worldwide. Because many patients were not able to safely attend a hospital, other means of connecting patients and physicians had to be found. Safe and easy to use online video-conferencing applications, which were developed especially for the healthcare sector, enabled physicians to see and talk to the patient, saving them a personal visit. However, online video consultation cannot be used for all patients, especially not for those who need a physical examination or any other procedure that requires physical attendance.

2.4 Nursing 4.0

Inpatient care will also change permanently due to digitalization. Not least because the increasing shortage of doctors and nursing staff demands that solutions be found for providing adequate care to hospital patients.

In addition to the actual nursing care that patients require nursing staff are burdened with a variety of additional, usually bureaucratic tasks. Digitalization can help make processes more efficient, if used incorrectly however, it can have the opposite effect. Take for example the morning rounds on a surgical ward. There are often as little as 20–30 min to see patients before the daily surgical program is discussed. While looking at wounds and discussing how treatment should proceed orders are given which need to be carried out over the course of the day (e.g., X-ray, planning of discharge, information regarding new surgery, start of physiotherapy, etc.) and involve a wide array of professional groups.

Since it takes too long to input these orders into the hospital information system, slips of paper are often used instead, which often wind up in the coat pocket of a doctor who spends the rest of the day in the operating theatre and the orders can only be carried out late in the afternoon when they get out. This is where mobile documentation systems with intelligent task control can help. These systems enable fast access to any important information right at the patient's bed. They can also be used to create and share to-do lists, which can then be worked through together. When the senior physician comes out of the operating room late in the afternoon, they can see at a glance whether or not all of their orders have been carried out and which still need to be done.

Another problem nurses and other healthcare professionals face is knowing where to find their patients. For example, the patient may be taking X-Rays or an ECG, or have gone to the cafeteria with their relatives. One solution available in a Hospital 4.0 would be the implementation of a tracking system, such as an RFID-chip in the patient's wristband. This would make it easy to know where any patient is at any time. Other staff, such as doctors would also be much easier to locate this way, for example when a doctor is needed for a pre-operative examination.

An electronic order-entry-system, i.e., the electronic input and recording of all manner of requests and orders must be implemented in a digitally enabled hospital. This would enable physicians to, for example, request an X-ray be made and then be able to see it in the hospital information system once it has been done. Such systems require the integration of various subsystems (in this case HIS and PACS) and usually the exchange of messages between these systems. Other processes can also be better tracked and monitored using an order-entry-system. For example, an order-entry-system for the prescription (physician) and administration (nursing) of medication, which can be linked to the hospital's pharmacy so that an up-to-date list of available medication is available at all times and supplies can be restocked if necessary.

Ideally, Hospital 4.0 is also integrated with the services responsible for providing nursing care. Patients dependant on nursing care also need to be provided for after inpatient treatment. This is particularly challenging if the patient did not need nursing care before their admission to the hospital (e.g., in the case of an acute stroke). In addition to organising their further care, structural changes may need to be made in the patient's home and general environment, such as a nursing bed, a wheelchair or walking aid and sometimes renovations may even need to be made. If this process is initiated only once the patient has been discharged, they may have to cope with what were perfectly avoidable waiting times, during which the patient may even have to remain in the hospital.

2.5 Administration 4.0

In addition to the core process of providing medical care to patients, hospitals must manage a number of supporting processes. The potential for digital integration of logistical processes has already been explored in detail, however other areas such as infrastructure management, human resources, payroll and accounting, communication and IT can also benefit massively from integration in intelligent networks. For example, recruiting is increasingly taking place on social media platforms such as Facebook, Twitter, Instagram, as well as recruiting websites, meanwhile traditional recruitment tactics such as advertisements in newspapers are become less important. Applications are sent online, which means they can be easily processed and shared through the proper channels within the hospital. Pay slips, holiday applications, specialized training applications, etc. can all be submitted electronically, which, because everyone has an overview of any planned absences, allows for smoother absence-planning within each work group.

Communication between staff can also be done electronically. Intranet solutions can be used to speed up processes greatly, for example, medication might be ordered for the ward electronically in a similar fashion to how one might apply for holiday time. It can then be approved by a superior and forwarded to whatever department is responsible for processing the application (in this case either the pharmacy or personnel department).

Other sections of this book examine the changes to the working world with much greater depth.

2.6 Patient 4.0

The digital integration and networking of hospitals must ultimately serve the goal of optimal care for the patient. As previously described, the strict sectoral boundary between inpatient and outpatient care needs to go. With the help of modern telemedicine, patients will be able to receive optimal care from experts at far away clinics at their local hospitals. Electronic patient files can ensure that patients always have access to information about their condition, and with home monitoring and wearables the patient can gather and collect data at home, during their everyday lives, which may be critical to planning their future treatment.

Patients will be transformed from a “sufferer” (according to the original Latin meaning of the word) to an active member of their own care team. It will also transfer responsibilities to patients that they never had before. If X-ray images are no longer sent between doctors, but uploaded into their electronic patient file, the patient now suddenly becomes involved in the exchanging of those images. Patients will then have a full overview and complete control of who sees their X-ray images.

Patients can now also be involved in their inpatient treatment, thanks to digital integration. They can see their upcoming appointments on a bedside display, see who their attending doctors and nurses currently are, order meals or access the hospitals entertainment program. Patients can also call up information about previous operations and gather information before the actual clarification discussion takes place. Patients can also make appointments for follow-up care online and communicate with their doctors in between examinations.

2.7 Research and Training 4.0

University hospitals, and increasingly other hospitals as well, contribute to the advancement of medicine with their deep commitment to research and training. Here too, digitalization and network integration will lead to considerable changes. Young doctor’s residencies often take place in multiple hospitals. With the use of video-conferencing solutions, even students from other hospitals can virtually participate in patient presentations. Virtual and augmented reality can also be used to train new doctors. The first steps of many operations can now be performed virtually, the procedures simulated before being performed on living patients. Now new doctors can be trained without endangering patients at all. Even experienced physicians can benefit from simulated training, similar to aviation, where mastery over the solutions to rare but extremely dangerous complications can be gained.

Learning algorithms (artificial intelligence/machine learning) open up never before available options to researchers. For such a system to be able to learn and recognize patterns, for example, to identify the early symptoms of a heart attack, mountains of data are required from the corresponding patient group (in this case patients with cardiovascular problems). Other disciplines face similar challenges.

Oncologists want to choose the best option for their patients out of the increasing number of treatment options available for certain tumours, for example. Intelligent systems could greatly aid in the decision-making process. These systems too require mountains of data from a multitude of different patients.

The data required is often already available in hospitals, however no normed and uniform format exists which would allow the data from different hospitals to be combined. Moreover, this data must be anonymous for the patient's protection. The German Ministry of Education has provided incentives for the digital integration of hospitals with its Medical Informatics Funding Initiative. Research data are often just snapshots at the moment. "Through their integration, a range of new possibilities in patient care, as well as biomedical research are made available. (. . .) In combination with treatment information collected over longer periods of time by doctors, the course and complexity of a given condition can be described with a greater precision. This enables researchers to gain a better understanding of diseases, something urgently needed for the development of new individualized methods of prevention, diagnosis and treatment. The new framework aims to further medical research and improve patient care. Innovative IT solutions will enable the exchange and intelligent application of data from patients, clinical and biomedical research.

The Federal Ministry of Education and Research will support high-performance, interdisciplinary consortia in medical informatics with approximately 120 million euros over the next four years. The keystone of this framework is establishment of so-called "Data Integration Centres" at German university hospitals and partner organizations. These centres are to demonstrate how data, information and knowledge from patients, clinical and biomedical research can be linked together across various locations." (BMBF 2019)

2.8 Risks 4.0

In 2016, a hospital in Neuss fell victim to a so-called "ransomware" virus (Heise 2016). These viruses encrypt the data on the hard drives of computers so the data remains inaccessible until the user enters a password. The password however can only be acquired by paying a "ransom." Although a backup was available the estimated damages totalled around one million euro, 15% of all operations could not take place.

An estimated two-thirds of German hospitals fell victim to a series of attacks in May of 2017. Despite this the IT budgets in German hospitals are usually quite small. German hospitals spend 2% of their turnover on IT. In comparison, American hospitals spend five times as much. IT experts, given the mounting threat of cyberattacks, generally view the 2015 IT security law (IT-Sicherheitsgesetz) as a step in the right direction. The law aims to protect against IT security threats by defining what critical infrastructures are (KRITIS) and who operates and maintains them. Now, with the BSI-Kritis ordinance (BSI-Kritis Verordnung), which came into effect in June of 2017, the health sector also counts as essential, critical infrastructure.

In the hospital in Neuss, the virus infected the hospital's IT systems via an email attachment. According to EMRAM criteria, hospitals must ensure that attacks on digital infrastructure are avoided entirely, or swiftly dealt with. This will become increasingly important as digitalization and integration progress.

The difficulty here is managing mobile devices; especially those of the employees (bring your own device). Portable and mobile devices need to be set up properly and managed, which requires expertise and personnel that many hospitals lack.

Data protection means that only those who are explicitly allowed to access a piece of information actually gain access to it. Which also means that third parties cannot be allowed to modify that data or gain unauthorized access to any devices. In 2018 for example, Reuters reported that pacemakers could become a target for hackers in the future (Reuters 2018). Although no such cases have been reported so far, the danger seems quite real, especially when considering that then US vice president Dick Chaney received a specially made pacemaker in 2007, one that had all wireless functions disabled. The attending cardiologist Dr. Jonathan Reiner said: "It seemed to me to be a bad idea for the vice president of the United States to have a device that maybe somebody . . . might be able to get into, hack into," Turning to Cheney, the cardiologist added, "I worried that someone could kill you." (CNN 2013)

New technologies bring with them new opportunities, and new threats. When the new electronic health record VIVY was unveiled in 2018 for example, it took only a few short days to identify multitude of security flaws. Theoretically, because it is a joint project between several insurance companies, over 13 million people can use VIVY. Although several security flaws have been patched (according to the manufacturer) and some of the vulnerabilities deemed unrealistic, the patient's confidence in the security and integrity of their data has suffered (Zeit 2018). It is however worth mentioning that VIVY is the first electronic patient file to be published in Germany. Other providers will follow suit and learn from their mistakes. VIVY will also continue to evolve and try to regain some of that lost trust.

2.9 The Digital Hospital: Living Dream or Waking Nightmare?

The unstoppable march of technological progress is changing many aspects of our lives. It will lead to extensive changes in hospitals, and more broadly, changes to medicine. Hospitals now have to both embrace modern business practices as economic enterprises, and further develop themselves as medical service providers. Patients will rightly expect that their diagnosis and treatment correspond with the current best practices according to science, and that their stay in the hospital be as efficient as possible.

In the future, increasing personnel shortages in nursing and other medical services will force hospitals to integrate their processes digitally to keep staffing requirements low. Unfortunately, this also involves a definite risk of alienation. In other branches, there exist already many processes completely devoid of human contact (online banking, online booking, in-app train tickets).

Many hospitals in Germany are still transitioning towards Hospital 4.0. Hopefully, they can succeed in embracing digitalization and using its full potential for the benefit of their patients without losing the human touch. The current COVID pandemic has also shown that modern technologies can enable the contact between patients and physicians/nurses and increase the safety in healthcare during a pandemic. While the COVID pandemic will hopefully end once the majority of the population is vaccinated, it will have a lasting effect on how we will use modern technology to perform healthcare in a safe and effective way.

References

Arnold J; Geschichte der Informatik, Script zur Vorlesung Grundkonzepte der Informatik, PH Bern, 2010.: http://www.tigerjython.ch/download/InformatikGeschichte.pdf (21. Januar 2019)

Ärzteblatt, IT-Sicherheit im Krankenhaus: Cyberrisiken als Herausforderung https://www.aerzteblatt.de/archiv/194044/IT-Sicherheit-im-Krankenhaus-Cyberrisiken-als-Herausforderung (30.01.2019)

Bär AS; Intensivmedizin im Wandel der Zeit; Dissertation an der Universität Regensburg, 2011; Epub: https://epub.uni-regensburg.de/21743/4/Dissertation_Endversion.pdf (21. Januar 2019)

BMBF: Digitalisierung in der Medizin: https://www.bmbf.de/de/medizininformatik-3342.html (30.01.2019)

CNN, October 24 2013, Cheney's defibrillator was modified to prevent hacking; https://edition.cnn.com/2013/10/20/us/dick-cheney-gupta-interview/index.html (30.01.2019)

Heise online 2016: Ransomware-Virus legt Krankenhaus lahm; https://www.heise.de/newsticker/meldung/Ransomware-Virus-legt-Krankenhaus-lahm-3100418.html (30.01.2019)

HIMSS EMRAM 2019.: https://www.himss.eu/sites/himsseu/files/analytics/emram/Europe_HIMSS_Analytics_EMRAM_Criteria_2018.pdf (28.01.2019)

Köhler CO; Historie der Medizinischen Informatik in Deutschland von den Anfängen bis 1980; 2003.: http://www.informierung.de/cokoehler/HistorieMI_Koehler_text.pdf (21. Januar 2019)

Mittelstaedt E; Gesundheit 4.0; erschienen in: Versorgung 2030 – eHealth, mHealth, Telemedizin; Gesellschaft für Versicherungswissenschaft und –gestaltung; 2015

Reichertz PL (1970) Requirements for configuration and management of integral medical computer center. Methods Inf Med 9:1–8

Reuters Health News February 20 2018: Pacemakers, defibrillators are potentially hackable; https://www.reuters.com/article/us-health-heart-pacemaker-cyber/pacemakers-defibrillators-are-potentially-hackable-idUSKCN1G42TB (30.01.2019)

Tempis: Zehn Jahre TEMPiS – eine telemedizinische Erfolgsgeschichte; http://www.tempis.de/index.php/pressemeldung-2013-05-07.html (23.01.2019)

Traumanetzwerk.: http://www.traumanetzwerk-dgu.de/de/startseite_tnw.html (23.01.2019)

Zeit Online, 2018; Krankenkassen-App Vivy hatte womöglich erhebliche Si-cherheitslücken; https://www.zeit.de/digital/datenschutz/2018-10/gesundheitsdaten-sicherheit-medizin-app-vivy-datenschutz/komplettansicht (30.01.2019)

Operating Room and Clinic 4.0: The OR.NET Approach

Armin Janß, Stefan Schlichting, and Klaus Radermacher

Contents

1 Introduction

For some time now, a paradigm shift has been taking place in the healthcare sector—away from isolated solutions and towards standardized, open integrated systems. Thus, integration and networking in the area of medical IT and the operating room have been a major topic for years. As of today, however, only manufacturer-specific, proprietary integrated system solutions are available. The BMBF (German Federal Ministry of Education and Research) lighthouse project OR.NET (duration: 2012–2016) has provided the scientific and technical basics for vendor-independent and thus open, dynamic interoperability of networked medical devices and medical IT systems (Golatowski et al. 2018).

A. Janß (✉) · K. Radermacher
Chair of Medical Engineering at the Helmholtz Institute for Biomedical Engineering at RWTH Aachen University, Aachen, Germany
e-mail: janss@hia.rwth-aachen.de; radermacher@hia.rwth-aachen.de

S. Schlichting
Drägerwerk AG & Co. KG, Lübeck, Germany
e-mail: stefan.schlichting@ornet.org

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_45

For many years, the number and complexity of technical systems in the operating room (OR) and in hospitals has been increasing (Rau et al. 1994, 1995; Leape 1994; Hyman 1994; Cook and Woods 1996; Kohn et al. 2000; Merten 2007; Janß et al. 2014; Benzko et al. 2016; Janß et al. 2018). In particular, recent developments in computer-aided diagnosis and therapy devices support the trend towards personalized medicine. This in turn requires information sharing through improved communication, high flexibility, and access to specific minimally invasive therapy options on demand (Lauer et al. 2008; Ibach et al. 2009; Barbe et al. 2010; Strathen et al. 2017; Vossel et al. 2017). The boundaries between medical devices and clinical information technology (IT) are disappearing. The paradigm shift from isolated devices to system interoperability requires consideration of modular-networked devices and safe, effective, and efficient Human-Machine-Interaction (Zimolong et al. 2003; Ibach et al. 2008; Ibach 2011; Ibach et al. 2012; Janß et al. 2014; Benzko et al. 2016; Janß et al. 2018).

Current developments as the Internet of Things (IoT) and Industry 4.0 as well as paradigms such as service-oriented architectures (SOA) already promise open systems today and will continue to do so in the future. However, currently available commercial solutions for integrated operating rooms and the interfaces of the associated medical devices between each other and the medical IT are not yet open, but proprietary; networking or data exchange is usually only possible between products from one manufacturer or between certain devices from different manufacturers (within collaborations) and even then, often only to a very limited extent. These monolithic solutions limit the flexibility of operators and medical users to integrate independent devices and innovative device functionalities into these integrated OR solutions. Functional safety and the protection of sensitive data pose a further problem. At the same time, with “open” networking, medical devices which are not coordinated with each other, pose significant risks regarding incorrect usage and potentially relevant information remains unused (Lauer et al. 2010; Montag 2012; Hölscher and Heidecke 2012; Zeißig et al. 2016). Risk management is also problematic, as clinic operators are responsible and become liable if they combine medical devices in “in-house production” (Janß et al. 2018).

In hospitals and operating rooms, operators and medical users are familiar with the problems of previously missing or insufficient integration due to non-existent communication standards. Operators in the operating room are surrounded by a multitude of isolated solutions consisting of highly specialized technology and medical devices. The devices are often provided by different manufacturers, whose user interfaces usually have different and complex usage concepts (Cook and Woods 1996; Zimolong 2004; Hübler et al. 2007; Blaar et al. 2015; Janß et al. 2016). To avoid an additional burden for the medical staff due to the use of complex medical devices, the Human-Machine-Interfaces must be safe and usable and the work process must be supported by the system integration (Backhaus 2010; Janß et al. 2016). The realization of an integrated intelligent operating room is one of the major challenges in research and development regarding the clinic, especially with regard to the integration of instruments, devices and functionalities from different manufacturers (Feußner et al. 2016; Czaplik et al. 2018).

The surgeon bears responsibility in the operating room and should also be able to control all devices and their parameter settings. For example, the surgeon wants to be able to make settings on the equipment himself, but today the equipment or operating interfaces are usually located in the non-sterile area. Therefore, the surgeon must ask the non-sterile OR nurse or the so-called "circulating nurse"—who is generally responsible for several operating rooms at the same time. The surgeon usually does not have direct control over whether the value is set correctly, since many devices are not directly visible to him. Incorrect usage of the specialized equipment can thus potentially lead to hazards for the patient or the surgical team (Rau et al. 1996; Cook and Woods 1996; Zimolong et al. 2003; Lauer et al. 2009; Hofinger 2009).

Another example of deficient communication between the sterile and non-sterile area is the interaction with the operating table. When adjusting the position of the operating table, the surgeon's instructions are often inadequately realized; instead of "foot-down," for example, the first step is often "head-down" (Matern 2006). In most cases, this does not have any consequences, but it disturbs the workflow in the OR and promotes the surgeon's desire to gain control over the overall OR system himself.

Integrated OR systems focus on higher quality in patient care and patient safety. On the one hand, the intraoperative workflow and its effectiveness and efficiency are improved, resulting in time and cost savings (Nowatschin 2009; Barbe et al. 2010; Wallwiener et al. 2011; Ibach 2011; Ibach et al. 2012). On the other hand, e.g., centralized control and flexible image distribution improve usability and reduce workload (Lauer et al. 2008; Klein 2010; Strauss et al. 2013; Janß et al. 2018).

In addition to the benefits for users and patients, open networked systems bring further advantages for hospital operators and especially small and medium-sized enterprises. Open networking of medical devices enables the clinical operator to combine devices from different manufacturers as required, as these are developed independently of each other and yet are compatible. The operator thus gains technical and economic independence.

However, independence from a system manufacturer applies not only to operators, but also to small and medium-sized enterprises, which are able to develop their own products based on open standards and thus participate more easily in the market. This also means that innovations can reach the market more quickly and do not fail due to resistance from "big player" manufacturers.

The main objective of the BMBF-funded lighthouse project OR.NET (duration: 2012–2016) was to develop the technological, legal and operational basis for modular, dynamic and open interoperability of medical devices and IT systems in future operating rooms and their clinical environment (Kücherer et al. 2013). In addition to the development of the system architecture and the construction of demonstrators, novel approaches and methods were developed with regard to risk management and usability engineering as well as the approval and certification of openly-networked medical systems. Furthermore, the operational suitability of the novel OR systems and components in the daily routine of a hospital was investigated and the international standardization was advanced (Beger et al. 2017; Golatowski et al. 2018).

In this paper, the state of the art as well as further background and results of the OR.NET project will be presented and highlighted.

2 State of the Art

In the following, proprietary integrated OR systems are presented at the beginning. The OR1 Fusion from Karl Storz, Tuttlingen, is an intelligent PC-based system that combines the functionalities of communication and documentation. The modularity of the platform enables an operating room environment that can be adapted to the changing conditions in the OR. Endoscopic devices, video and data sources as well as, for example, the operating table and ceiling-based OR lights can be controlled from a central unit. Opportunities for teleconferencing and telemedicine are possible by integrating the OR with existing hospital information systems. Networking is implemented via the proprietary Storz Communication Bus (SCB).

The Endoalpha system from Olympus, Hamburg, Germany, integrates surgical equipment and support systems to create an environment in the operating room that enables effective, efficient and comfortable surgery. The Endoalpha system can be used to coordinate and control medical equipment in the operating room—from room cameras to surgical lights to 3D endoscopes. In addition, surgeons can call up information directly at the operating table or use video telephony. Furthermore, the documentation is networked with the hospital information system (HIS). Olympus also uses a proprietary network protocol.

The Tegris OR system from Getinge Maquet, Rastatt, Germany, includes a user interface for the integrated operating room. With the Tegris system, the OR staff can control various devices such as cameras and monitors as well as operating tables and OR light systems. Patient data can also be retrieved via interfaces to the HIS system and visualized on central screens for the OR staff.

Research approaches of openly networked systems are described below. The goal of the SCOT project (Smart Cyber Operation Theater) in Japan is to develop an openly integrated operating room (Okamoto et al. 2018). Through the project, OR rooms should be designed more intelligent. Here, the OR room is connected to a computer *system* via a *Cyber Physical System (CPS).* By analyzing and processing the resulting large amounts of data, it is possible to improve safety, effectiveness and efficiency in the treatment. To this end, SCOT also incorporates networked medical devices and high-performance computing. SCOT uses the *Open Resource interface for the Network (ORiN)* as a communication interface, which connects the devices in the OR room. The interface was originally developed for industry and adopted in the SCOT project for OR broom integration.

One focus area of the project deals with robotics. This speaks for the choice of the interface, since ORiN was originally developed for robotics, e.g., to control several industrial robots. In contrast to the MDPnP and OR.NET projects presented below, SCOT does not have a “plug and play” function. Therefore, when new devices are used, the code must be changed by the administrator. However, it is assumed that in the future a “Plug and Play” function will also be necessary for SCOT.

The *Medical Device Plug and Play (MDPnP)* program was established in 2004 in the United States and is part of the *Massachusetts General Hospital (MGH)* and the *Center for Integration of Medicine and innovative Technology (CIMIT)* (Arney et al. 2018). The program aims to eliminate barriers in medical device technology related to interoperability in a multifaceted approach. This includes developing and supporting an appropriate open standard and eliciting, capturing, and modeling clinical use scenarios and engineering requirements for an appropriate platform. The goal of the program is to develop an open standard and define the conditions for a patient-centered *Integrated Clinical Environment (ICE)* under which interoperability will enable device integration that produces a medical device system with high safety and performance.

The OR.NET project (Safe Dynamic Networking in Operating Room and Clinic) was funded by the German Federal Ministry of Education and Research (BMBF) in the period 10/2012–4/2016 with a total of 18.5 million € (Birkle et al. 2012; Kücherer et al. 2013; Benzko et al. 2014) and led by Heidelberg University Hospital (Prof. B. Bergh—spokesperson) and RWTH Aachen University (Prof. K. Radermacher—project coordinator). The network of the project grew from about 50 to 95 project partners by 2016, especially due to further interested industry partners as well as clinics. The aim of the project was to achieve interoperability of medical devices and IT systems from different manufacturers via a defined system architecture and interface standards, so that operators can use medical devices in an interoperable manner regardless of the manufacturer (Feußner et al. 2016). To this end, fundamental concepts for a safe and open dynamic network must be developed, evaluated, and transferred into standardization activities (Janß et al. 2018). At the same time, the interests of both, manufacturers and clinical operators must be balanced and considered (Janß et al. 2018). In addition, Human-Machine-Interaction in particular is being considered in this new environment and, as a result, concepts for safe usage are being developed. In this way, medical devices and IT systems are to be combined in a flexibly networkable manner that can be implemented in everyday clinical practice and enables risk management.

In the OR.NET project, a common data model with open source libraries was developed. In addition, relevant methods and strategies for the risk management of such modularly integrated, clinical system architectures were developed for the certification process. Furthermore, novel user interfaces that consider ergonomic and usability requirements have been developed. The benefit of safe and dynamic networking could be proven with the help of numerous “use cases” (Feußner et al. 2016).

The OR.NET project results have been transferred to the ISO IEEE 11073 SDC (Service-Oriented Device Connectivity) international communication standard (ISO IEEE 11073-10207, -20701, -20702) and have enabled the application in future medical devices since September 2018 (Kasparick et al. 2015; Kasparick et al. 2018a; Janß et al. 2018; Besting et al. 2017).

The OR.NET project, which involved almost 100 project partners from hospital operators and medical users as well as research institutions, industry and standardization bodies, was transferred to a non-profit association OR.NET e.V. in 2016 (www.ornet.org), which in the meanwhile continues and internationalizes the work with over 57% industrial participants.

With the OR.NET initiative and the corresponding communication standards, vendor-independent and dynamic networking of medical devices and medical IT technology is becoming a reality. In contrast to the now available open ISO IEEE 11073 SDC communication standard, closed systems were previously dependent on the devices approved by the manufacturer and tied to them for the device ensemble. Now, all SDC devices available on the market can be networked in the future. In addition, scalable solutions are possible with the SDC standard, from a minimum of two devices to significantly more than 10 devices, which can be centrally controlled and monitored (Beger et al. 2017, 2018; Golatowski et al. 2018).

The industry-led EFRE ZiMT project (certifiable integrated medical technology and IT systems based on open standards in operating rooms and clinics; coordinated by SurgiTAIX AG, Herzogenrath) is a project that follows the OR.NET project. Supported by funds from the EU and the state of North Rhine-Westphalia, the open communication standard is being further developed and established here. In addition to the further development of approval strategies and methods for modular risk analysis, a surgical workstation with novel user interfaces is being developed in particular. This approach is intended to enable secure and dynamic networking in operating rooms and clinics, thus counteracting the declining compatibility caused by different manufacturer systems with their proprietary protocols.

3 The OR.NET Solution Approach

3.1 Standardization

The basis of all innovations described so far is the vendor-independent networking approach, based on open standards. The ISO IEEE 11073 SDC (Service-oriented Device Connectivity) family of standards was developed within the BMBF-funded projects OR.NET and MoVE as well as the EFRE-funded project ZiMT (Kasparick et al. 2018b). The non-profit association OR.NET (www.ornet.org) coordinates these standardization activities. The SDC standard specifies how devices can exchange application-related data without loss of quality. Standardized interfaces facilitate the integration of medical devices into a network. Data transmission is bidirectional, so that an SDC-enabled device can be both, a consumer and a provider. Thanks to end-to-end encryption, SDC offers a high level of safety during data transmission. The ISO IEEE 11073 SDC standard family consists of three sub-standards (Fig. 1).

The Medical Devices Communication Profile for Web Services (MDPWS, ISO IEEE 11072-20702:2016) standardizes safe data transmission between medical

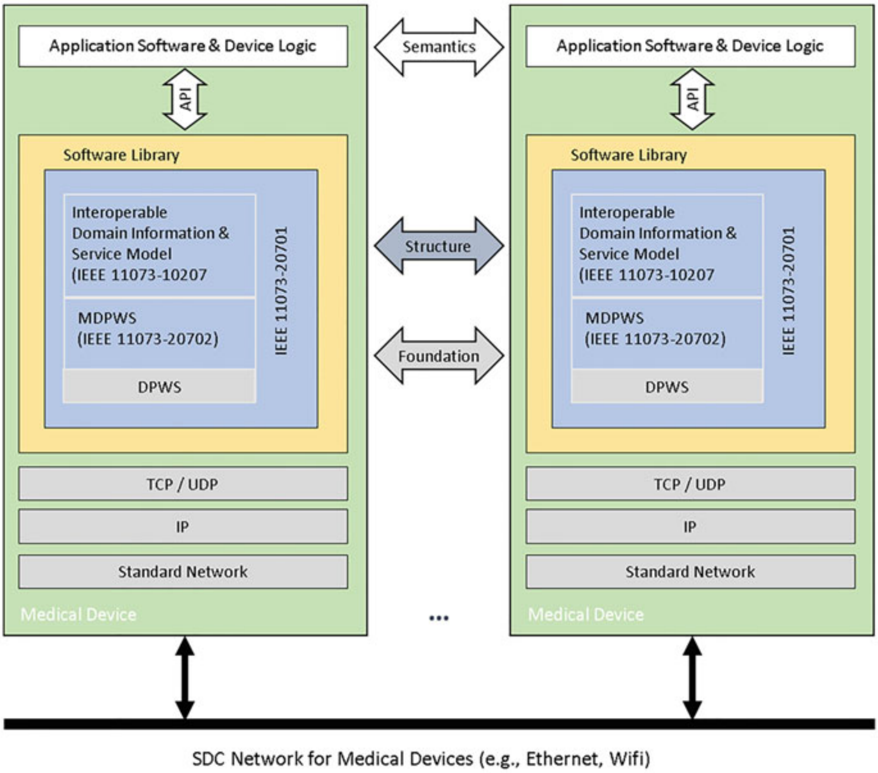


Fig. 1 ISO IEEE 11073 SDC standard family according to Kasparick (Kasparick et al. 2018a)

devices. In addition to the exchange of data, other aspects are specified, such as the dynamic finding of devices at runtime, safety mechanisms (safety) by means of dual-channel or embedding context information, data security (security) by means of encryption, authorization and authentication, transmission of data streams (e.g., for ECG) or efficient data transmission by means of compression.

The IEEE 11073-10207:2017 Domain Information and Service Model defines the self-description of device properties and device state. It also describes which interaction capabilities (services) can be provided to enable interaction between medical devices. Such a device description, using semantic tags based on known nomenclature standards, enables safe semantic interoperability. This means that information and control commands cannot simply be exchanged, but that it can be ensured that they are interpreted correctly. This represents a mandatory prerequisite for safe and manufacturer-independent networking.

To be able to network medical devices with each other and with hospital IT, regardless of manufacturer, they must be semantically interoperable. This is a benefit for patients, OR teams and hospital operators. Let us assume, for example, that the patient's current vital signs are to be displayed and used. For example, a pulse oximeter might provide a value of 70. A displaying device would record this and display it in a suitable form. If the value 70 represents the heart rate, the patient is generally doing well. If, on the other hand, this value represents oxygen saturation, the patient is usually in a very critical condition. It is therefore crucial to know not only the value of a measured value but also its semantics. Therefore, specific codes from so-called coding systems are assigned to each piece of information provided in the network to ensure unambiguous interpretability. IEEE 11073-101XX is used as nomenclature. In addition, the unit of measurement values is also defined. For pressures, for example, it is important to know whether they are given in Pa, bar, psi, mmHg, etc. A comprehensive and unambiguous self-description of medical devices is therefore a basic requirement for safe networking. Since medical devices are often much more complex than a pulse oximeter, which has only a small number of measured values and parameters, the self-description takes the structured form of a so-called containment tree. The interaction options between the medical devices are also defined by means of so-called services. A medical device can provide information, for example, for read access, according to the retrieval and/or subscription principle, but can also permit remote control, if this makes sense and is possible from a risk management point of view.

The so-called ISO IEEE 11073 SDC standard family is completed by the third standard ISO IEEE 11073-20701:2018. This describes the overall system of decentralized networked medical devices based on the service-oriented architecture (SOA) and makes a binding between the two previously described standards (Kasparick et al. 2018a).

Since 2018, the standardization of the ISO IEEE 11073 SDC standard family has been completed. While the ISO IEEE 11073-20702 and -10207 sub-standards were already completed at the end of 2016 and 2017, respectively, the standardization process for the third part IEEE 11073-20701 was completed in September 2018.

Work is currently underway on extensions to the ISO IEEE 11073 SDC standard family that will allow specific clinical system functions to be implemented in a cross-vendor combination of medical devices. Once the technical foundation for open systems has been laid, the next step is to bring such systems to market. In addition, so-called device specializations for endoscopic medical devices (camera, light source, pump/insufflator) and HF surgical devices are being developed, e.g., in the BMWi-funded PoCSpec project. These device specializations represent defined model building instructions for the ISO IEEE 11073 SDC mapping of the medical devices. This approach brings together the manufacturers of the relevant medical devices (SMEs and international market leaders) and research institutions regarding the further developments and standardizations.

3.2 Implementation of the SDC Standard

Due to the high comprehensiveness of the SDC standard documents, easy-to-use software libraries for the usage of manufacturers are essential. Therefore, open-source software libraries (e.g., www.surgitaix.com) were created in the OR. NET project, which manufacturers can use to adapt their existing software to implement standard-compliant interfaces. These developments were continued in the ZiMT project.

The software libraries offer a simple programming interface for medical device participants to find other devices, offer data in the OR network or consume data from other devices and thus follows the service-oriented architecture mentioned earlier. The libraries are now available for several common high-level languages (C++, Java, and .NET).

Further maintenance of the libraries is the responsibility of the OR.NET association *Software Stacks* working group, which is in constant contact with medical device manufacturers to provide support. A further goal is also the close exchange with the standardization work, since new revisions must be incorporated into the software life cycle.

3.3 Approval Strategies and Testing

To use the advantages of open networking, the regulatory framework conditions must be observed and appropriate prerequisites have to be created. In the beginning, the manufacturer-independent open networking of medical devices represented a completely new situation and, among other things, raised legal questions of approvability and liability. The usability engineering process according to DIN EN ISO 60601-1-6 and risk management according to DIN EN ISO 14971 (technical as well as human-centered) must be considered and the methodology for validating networked software components had to be revised (DIN EN ISO 14971) (DIN EN 60601-1-6). The goal was to find a technically feasible and legally unambiguously clarified modular overall approval strategy for the manufacturer.

In addition, procedures and tools must be developed to support the approval and operation process for manufacturers and operators (Janß et al. 2018).

Offering new functionalities in open networking may entail new requirements in the context of the approval process regarding classification in a different (higher) risk class. For example, controlling another device leads to the risk class of the controlled device (Janß et al. 2018).

In the future, manufacturers and operators will be able to incorporate risk- and usability-analyses of the individual components into an overall risk and usability assessment on a modular basis. For this purpose, extended device profiles (incl. user interface profile and risk profile) and test procedures have been developed to support the approval process of modular systems (Janß et al. 2018).

In the declaration of conformity of new networkable medical devices, the manufacturer faces the challenge of potential networking partners being unknown in the approach of completely open networking. Since the manufacturer cannot develop and test against unknown interfaces, a way had to be found to make the “unknown” known.

One possibility has been presented in (Janß et al. 2018) by extended device profiles, which allows a standardized description of a user interface. The standardized description of the interface makes it possible for other manufacturers to use this description as a basis for their development. The description thereby represents a specification of the interface against which development and testing can take place, instead of having to develop against concrete devices (in consultation with other manufacturers). Development can therefore take place without having a foreign physical device available. The extended device profile contains descriptions of the following aspects:

The **extended intended use** describes the functionalities offered beyond the main purpose of the device type, which are now available by the network capability. Here, not only functions are offered in the network (service provider), but also functions of other medical devices are used (service consumer), both of them have be described. In particular, the intended use must describe the application scenarios that can be addressed by the device within the network.

The **technical device profile** sets out the network functions offered or used by the medical device. In addition to the name, the description of the functions offered must also include the purpose of the function and the forms of interaction supported. At this point, simplified reference can also be made to existing device specializations (according to IEEE 11073-104xx).

The technical description also contains information about the properties and functionalities of the medical devices, which they can offer or retrieve via their network interface. It also explains the modeling of the device type and possible extensions and specializations by the medical devices.

Since the networking of medical devices in the OR involves new risk-factors and -sources, the internal risk control measures of the medical devices alone are no longer sufficient to ensure the safety of patients and operators. The internal measures can only relate to an isolated medical device and must therefore be supplemented by **communication assurance measures**. These describe which communication safety measures are required by the devices (cf. ISO IEEE 11073-20702). There are various safety measures that can apply to the entire device or only to individual channels or metrics and can be divided into three categories:

- Integral part of the protocol. The standard-compliant implementation of the protocol automatically covers these measures.
- Optional part of the protocol. A standard-compliant implementation of the protocol provides for configuration and activation of these measures.
- Part of the protocol to be supplemented manually. These measures are supported by the protocol, but must be adapted to the particular characteristics of the medical device.

The required and supported measures must be listed in the device type profile, since compliance with these requirements and thus the safeguarding of communication between devices is the responsibility of the manufacturers. Since these measures (e.g., sequence number, time stamp, encryption, signature, authorization, feedback, dual channel, safety context and idempotent messages) are only meaningful and effective if they are supported by both sides, they must be documented at a central location.

The device profile must describe in each case, whether the measure is required or only supported. At the same time, it must be described whether the measure is required or supported for the entire device, for individual components, or only for individual metrics. For networking to take place, both networking partners must be compatible with regard to the safety measures for the metrics involved in the networking. If a measure is required by one of the partners, the other partner must at least support this measure.

The **Medical Device User Interface Profile** describes the requirements on the Human-Machine-Interface and thus on the representation of the functions offered on other devices (Janß et al. 2014). Among other things, these are guidelines for visualizing metrics and triggering actions, which are based on information from, inter alia, ISO 24752 (Universal Remote Console) and have been expanded in the course of OR.NET developments. However, the concrete design of the user interface is not prescribed, but restrictions and design recommendations are offered (ISO 24752).

If a device uses functions of other devices and displays them, the offering device must also describe in the UI profile which arrangement and appearance are present in the design of the user interface. The UI profile structures and groups individual functions and controls in a tree-like structure. Each of the functions and groupings is described by several attributes (criticality, assignment, spatial arrangement, relevance, elementary task (e.g., show, drag, select), etc.). In addition to the functions of the medical device and their grouping, the UI profile also describes requirements for the input or output device (e.g., minimum resolution and size of a screen).

To test the extended device profiles, the manufacturer has to perform conformance tests (tested by an independent institution), interoperability tests (connectathons or testing against a simulator) and integration tests (validation in the test laboratory).

The verification and validation is divided into three stages, which facilitates the conformity assessment of the medical device. The first stage verifies the basic compliance of the medical device with the procedures and technologies described in ISO IEEE 11073 SDC. The second stage describes interoperability testing between devices from different manufacturers and intraoperability testing that addresses the internal operations and functionalities of a device. As a third aspect of the test procedure, the usability of the medical device is validated with the help of interactive-centered usability tests. In this process, the physical devices are tested by means of a representative user group. The division of the test procedure into three stages represents a logical grouping and does not imply a chronological sequence of the stages (Fig. 2).

The introduction of the Medical Device Regulation also poses further challenges, which are currently being incorporated. The positive evaluation of the approval strategies developed with notified bodies and representatives of the FDA (Food and Drug Administration, USA) shows that the solution approach currently being pursued is promising under both European and US law.

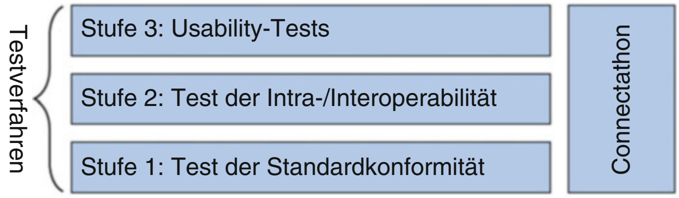


Fig. 2 Overview of the three-stage test procedure (incl. parallel connectathon)

4 Demonstrators

The development and deployment of open integration for the next generation of operating rooms and clinics requires comprehensive test and demonstrator environments. In the following, various challenges that must be considered during the demonstrator development are presented. Overall, the various demonstrators (at different locations including Aachen, Leipzig, and Lübeck) have proven the feasibility of the service-oriented medical device architecture (SOMDA) with a variety of example applications. These application examples demonstrate the potentials of open device interoperability (Rockstroh et al. 2018).

The demonstrator implementations were developed, evaluated, and discussed with medical partners from different disciplines. The evaluation focused on the characteristics of a network of medical devices, e.g., latencies in data transmission, stability as well as workflow processing, safe usage, and aspects of Human-Machine-Interaction. The demonstrator development at the Chair of Medical Engineering at RWTH Aachen University is described below.

A total of 27 partners from industry, research, and clinical practice were involved in the OR.NET demonstrator development at the Chair of Medical Engineering at the RWTH Aachen University. The use cases for the demonstrator were developed in close cooperation with surgeons and anesthesiologists from the University Hospital Aachen. More than 30 technical use cases were embedded in a realistic surgical workflow based on analyses of neurosurgical procedures of cervical spine decompression and fusion surgery. Technical use cases such as the automatic login and logout of devices in the network were shown, but also use cases with immediate clinical relevance, e.g., the display of device and vital parameters in the microscopic image or the documentation of image data and corresponding device parameters.

A major result was the development of central surgical and anesthesia workstation that enable the integration and control of various medical devices in the OR via touch-based graphical user interfaces and other Human-Machine-Interfaces (Fig. 3) (Benzko et al. 2016; Janß et al. 2014). The workstation is able to flexibly adapt its information presentation to the different requirements of both specialties, but maintains a unified user interface. The usability and interoperability as well as the intraoperative workflow were optimized and the information exchange between surgery and anesthesiology was enabled. In addition to information from the connected medical devices, the framework also visualizes general patient and

Fig. 3 OR.NET demonstrator at the conhIT 2017 (Connected Health IT Trade Fair and Conference), Berlin, 2017 including surgical and anesthesia workstation; device networking from over 30 manufacturers and systems

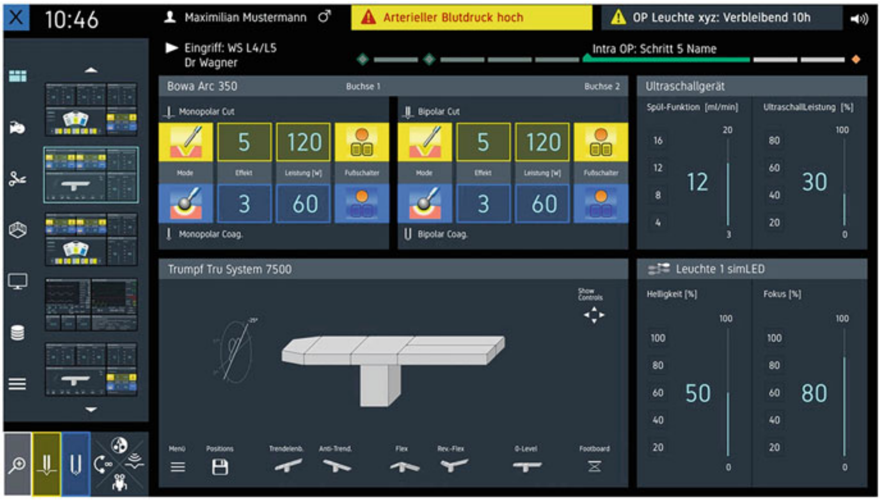


Fig. 4 Process-specific function group view in the OR.NET surgical workstation

intervention data, work steps, and physiological and technical alarms (Zeißig et al. 2016). It also provides access to the hospital information system (HIS) and the Picture Archiving Communication System (PACS).

The OR.NET user interface is also device- and platform-independent and enables simple and intuitive usage and, thanks to its responsive design, adapts to all display sizes in a context-specific manner. The surgeon and the anesthesiologist can compile important functions individually and manufacturer-independently and save this compilation as a process-specific function group view (Fig. 4).

The development of suitable input devices to ensure safe and usable systems was demonstrated in the OR demonstrator in Aachen with, among other things, a height-adjustable footstep and a central universal foot switch (Fig. 5).

Foot switches are used daily in surgery. However, problems such as shifting or slipping of foot switches often occur due to the parallel use of up to

Fig. 5 Universal foot switch (left) and function assignment via the GUI (right)

10 device-specific foot switches, which leads to a significant burden for the surgeon and the patient (Blaar et al. 2015). Currently, no footswitches are available on the market, which allow centralized activation of different devices from different manufacturers and reconfiguration during operation. Therefore, a new concept of a configurable central foot switch unit was implemented (Dell'Anna et al. 2016; Vitting et al. 2017; Janß et al. 2017). This universal foot switch, which has been developed by the Chair of Medical Engineering at the RWTH Aachen University with industrial partners, allows the release of different devices and functions from different manufacturers. The surgical workstation has a configuration option that realizes the assignment of the device functions to the footswitch pedals and displays it via a graphical user interface and allows intraoperative changing of the assignment. This replaces the previous "foot switch organ" (for a variety of different devices) in the OR with a universal foot switch that can be assigned context-sensitively and thus provides the necessary functionalities depending on the scenario and situation. This allows the surgeon to use customized combinations of device functions. Waiting times during the OR are reduced, as there is no need to search for or position the foot control on the floor or on a footrest.

In a user-centered evaluation, nine surgeons used both the configurable central foot control unit and four device-specific foot controls for a crossover experiment in an experimental surgical situation. It was shown that all surgeons were able to operate the configurable foot switch independently and that, in particular, efficiency in surgeon-device communication was increased.

In addition, footsteps frequently used in the operating room can be height-adjustable in the future and are able to adapt their level to the height of the operating table in the framework of interconnection. This will ensure an ergonomic posture and safe handling for the surgeon at all times (Radermacher et al. 1996; Lauer and Radermacher 2007; Dell'Anna et al. 2016; Janß et al. 2017).

In addition to the development of new input devices, a workflow bar with an integrated surgical safety checklist was implemented (Fig. 6). The WHO-recommended Surgical Safety Checklist (WHO 2008) consists of three parts which are used by the anesthesiologist and surgeon during different phases of the intervention. The digitally integrated safety checklist facilitates workflow and supports documentation and automatic transfer to the HIS system.

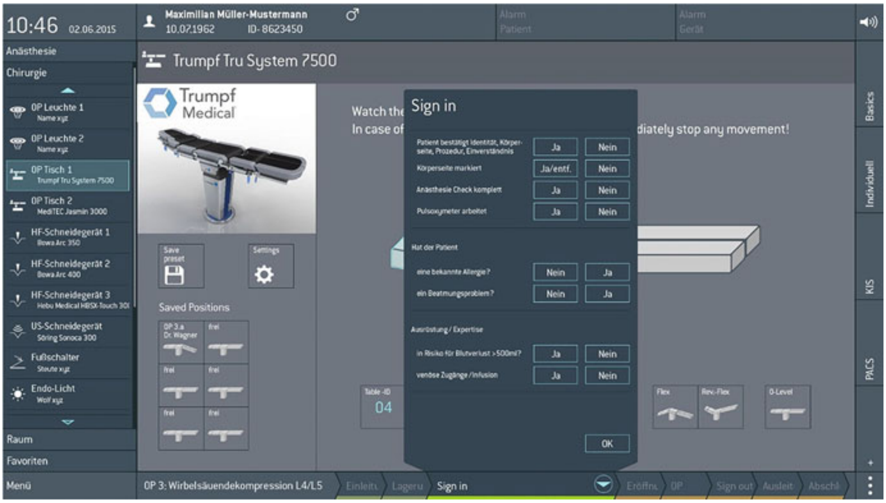


Fig. 6 Integrated WHO Surgical Safety Checklist in the OR.NET workstation

In addition, intelligent context-adaptive decision support has been implemented in the workflow bar. Alerts or notices are displayed depending on the current surgical step, comparable to the surgical safety checklist. This functionality supports the physician in critical situations by displaying standard operations and combining information from the surgical and anesthetic areas and devices.

The architecture implemented in the project forms the basis for novel applications in the operating room. The added value for the various stakeholders, in particular for the surgeons, the OR team and the hospital operators, is provided by novel functions that build on the achieved device interoperability. One of the most frequent requests from clinicians, namely the flexible context-specific configurable display of parameters and measurements of device functionalities directly in the field of view, could be realized. Although the specific configuration strongly depends on the clinical use case, the setups demonstrated the technical feasibility of the approach for a comprehensive solution. Anesthesia data could also be integrated in a compact manner via the workstation for the surgeon as needed (Rockstroh et al. 2018; Benzko et al. 2016).

The configurable assignment of device functions to input devices could also be implemented in various examples of foot switches, microscope hand switches and touch screens or endoscope probes. This allowed access to key device functionalities directly at the situs, giving the surgeon increased control (Rockstroh et al. 2018).

The integration of the interoperable medical device with the HIS is based on the current needs of physicians. Seamless transfer of patient identification and order data from clinical IT to the devices, replacing time-consuming and error-prone manual interventions, contributes to an optimized workflow. Conversely, the postoperative transfer of data from the devices to clinical IT reduces the effort required for documentation and device maintenance.

Overall, the demonstrators have proven the feasibility of the SOMDA approaches. A variety of example applications have been implemented at the demonstrator sites, demonstrating the potential of open device interoperability. Additional concrete applications for various clinical use cases will be implemented in future work. And based on the work on risk management, standardization, regulatory approval, and hospital operator strategies, new products will emerge to improve everyday clinical practice.

5 Outlook

Digitization is creating new network possibilities in the field of the integrated operating room and thus new functionalities (Barbe et al. 2010; Ibach 2011; Ibach et al. 2012; Andersen et al. 2017; Rockstroh et al. 2018). In the future, for example, courses of operations can be documented electronically and information available in the OR network from imaging and networked sensor systems can be used for improved patient positioning systems (Lauer et al. 2008), for navigation of surgical instruments and robotic systems (Ibach et al. 2008, 2009; Nowatschin 2009; Vitting et al. 2017; Vossel et al. 2017; Strathen et al. 2017). This can result not only in new functionalities and reduced costs, but also in better working conditions with less stress due to body postures, optimized workflows in terms of time or significantly reduced X-ray exposure (Barbe et al. 2010).

In an integrated OR, the environment (lights, air flow, etc.), medical devices and video distribution will be connected and controlled via a central OR workstation in the future, so that the OR team can control and view relevant parameters even in the sterile area. Images, videos, and device information can be flexibly distributed to different monitors (Ibach 2011; Birkle et al. 2012).

In the future, the OR team will be able to operate individual medical devices from different workstations (surgical, anesthesia), both on the wall monitor, on the sterile touch display or on a tablet (with sterile overlay). This creates central workstations that enable the surgeon and his team to control all devices and visualize device

Fig. 7 Setting device parameters from the SDC surgical workstation (left) and from the tablet (right)

Fig. 8 Cyber-Physical Systems in the operating room of the future (© meditec, RWTH Aachen 2016)

functions and vital parameters, as well as simplify and optimize the workflow (Fig. 7).

Bundled information about the patient's state of health can thus be made available immediately. The surgeon and his team must be able to use different input modalities in combination and operate safely, effectively, and efficiently.

In the future, mobile tablets can be used to quickly switch between individual device views and individual (process or user) views using the OR.NET approach. In addition to context-specific views, patient-relevant data will be provided in the OR field or transferred to the wall monitors at any time. Due to the interoperability within the SDC standard, for example, preoperative images can be called up and used on the wall display or the OR microscope.

The OR.NET approach (Fig. 8) allows the flexible combination, validation, and integration of working systems based on virtual, dynamically-updatable images of the real physical device world, whereby innovative technical solutions can be flexibly and safely as well as manufacturer-independently integrated into open networks as real device systems without having to make compromises in terms of usability and risk management.

Continuity of interaction concepts and also future device control from the sterile field eliminates waiting times and common sources of error in the device usage (Cook and Woods 1996). Surgeons and OR nurses, who have worked with various integrated OR systems for years, report time savings, reduction of setup errors, less mental stress and confusion and point to increased responsiveness regarding unexpected events (Nowatschin 2009).

In the operating room of the future, the OR.NET approach offers an optimization of the entire workflow for the OR team as well as an effective and efficient communication and interaction with the interfaces to the hospital IT. The OR.NET initiative is already advising hospitals to include the possibility of upgrading their OR systems to the SDC standard in their current calls for tender for new OR tracts and systems, thus creating further pressure on medical equipment manufacturers with regard to SDC integration and usage. In the course of these developments, an expansion and extended usage of the SDC standard, for example, in the hospital areas of OR management, care units and intensive care can also be expected in the next few years. This is also supported by the current dissemination and recognition of the SDC standard.

The non-profit OR.NET association has been founded in 2016 to continue and consolidate the work of the OR.NET project. Current research and standardization projects were also initiated with the support of the OR.NET association. The ZiMT, PriMed, PocSpec, and MoVE research projects are currently addressing issues in the areas of approvability, risk management, testing and certification of ISO IEEE 11073 SDC-compliant medical devices. In these projects, for example, a modular test suite is being developed that can be used by manufacturers during the development phase and for certification by independent institutions. In addition, test procedures and software-supported tools are being developed for hospital operators to enable verification of the device assembly, prior to the (initial) bringing into service.

In the short term, it is to be expected that manufacturers will use SDC interfaces in particular within the framework of bilateral and multilateral cooperations and initially still have one manufacturer bearing the main responsibility for the approval and conformity assessment process. After completion of the approval and test strategies by the OR.NET initiative, however, a completely open and modular networking approach can be expected in the medium term, in which clinics and operators can actually request SDC-compliant devices in their tender calls and procure them independently of one another to finally integrate them into a safe and dynamic integrated OR system.

At Medica 2018 in Düsseldorf, manufacturers already presented the first "pre-series" SDC-integrated medical devices. The first approved ISO IEEE 11073 SDC-compatible devices can be expected in the next two years.

Acknowledgement The presented work has been funded by the European Regional Development Fund (ERDF) within the ZiMT project (grant number ERDF-0800454) and by the German Federal Ministry of Education and Research (BMBF) within the projects OR.NET (grant number 16KT1203) and MoVE (grant number 03VNE1036B) as well as by the non-profit association OR.NET e.V. (www.ornet.org).

References

Andersen B, Kasparick M, Janß A, Rockstroh M, Franke S, Golatowski F, Radermacher K, Neumuth T, Ingenerf J (2017) Unterstützung des Chirurgen durch vernetzte Medizingeräte &

Computer-assistierte Chirurgie. In: Einhäupl KM (ed) Medizin 4.0 – Zur Zukunft der Medizin in der digitalisierten Welt, Berlin, Germany, Hanns Martin Schleyer-Stiftung, vol 88, pp 159–172

Arney D, Plourde J, Goldman JM (2018) OpenICE medical device interoperability platform overview and requirement analysis. J Biomed Eng 63(1):39–47, Verlag De Gruyter, Berlin, Deutschland

Backhaus C (2010) Usability Engineering in der Medizintechnik: Grundlagen - Methoden - Beispiele. Springer, Berlin/Heidelberg

Barbe A, de la Fuente M, Belei P, Buschmann C, Radermacher K, Mueller M, Mumme T (2010) Evaluation of the efficiency of the zero-dose-C-arm navigation approach. Z Orthop Unfall 148(5):554–558

Beger F, Janß A, Kasparick M, Besting A (2017) Der Operationssaal OP 4.0 - Für mehr Sicherheit und Effizienz durch offene Vernetzung in Operationssälen und Kliniken der Zukunft. In meditronic-journal - Fachzeitschrift für Medizin-Technik 4:20–24, beam-Verlag, Marburg, Deutschland, November 2017

Beger F, Janß A, Kasparick M, Bürger S, Clusmann H (2018) Gerätevernetzung im OP 4.0 wird Standard. In meditronic-journal - Fachzeitschrift für Medizin-Technik 5:6–10, Beam-Verlag, Marburg, Deutschland, November 2018

Benzko J, Ibach B, Marschollek B, Köny M, Leonhardt S, Radermacher K (2014) Innovative Kommunikations- und Netzwerkarchitekturen für den modular adaptierbaren integrierten OP-Saal. In: Niederlag W, Lemke HU, Strauß G, Feußner H (eds) Der digitale Operationssaal, 2. De Gruyter, Health Academy, pp 31–47

Benzko J, Krause L, Janß A, Marschollek B, Merz P, Dell'Anna J, Radermacher K (2016) Modular user interface design for integrated surgical workplaces. In Dössel O (ed) Journal of Biomedical Engineering/Biomedizinische Technik, Band 61, Heft 2, Seiten 183–197

Besting A, Stegemann D, Bürger S, Kasparick M, Strathen B, Portheine F (2017) Concepts for developing interoperable software frameworks implementing the new IEEE 11073 SDC Standard Family. In: Radermacher K, Rodriguez F, Baena Y (eds) CAOS2017. EPiC Series in Health Sciences, vol 1, pp 258–263

Birkle M, Benzko J, Shevchenko N (2012) Das Projekt OR.NET – Sichere dynamische Vernetzung in OP und Klinik. Deutsche Zeitschrift für Klinische Forschung, Innovation und Praxis 2012 16(11/12):41–45

Blaar M, Janß A, Dell'Anna J, Höllig A, Radermacher K, Clusmann H (2015) Bottlenecks and needs in human-human and human-machine interaction - a view from and into the neurosurgical OR. In: Dössel O (ed) Journal of Biomedical Engineering/Biomedizinische Technik, Band 61, Heft 2, Seiten 135–146, ISSN (Online) 1862-278X, ISSN (Print) 0013-5585, https://doi.org/10.1515/bmt-2015-0059, September 2015

Cook RI, Woods DD (1996) Adapting to the new technology in the operating room. Hum Factors 38:593–613

Czaplik M, Voigt V, Kenngott H, Clusmann H, Hoffmann R, Will A (2018) Why OR.NET? Requirements and perspectives from a medical user's, clinical operator's and device manufacturer's points of view. Biomed Eng-Biomed Tech 63:5–10

Dell'Anna J, Janß A, Clusmann H, Radermacher K (2016) A configurable footswitch unit for the open networked neurosurgical OR – development, evaluation and future perspectives. i-com. Band 15, Heft 3, Seiten 227–247, ISSN (Online) 2196-6826, ISSN (Print) 1618-162X, https://doi.org/10.1515/icom-2016-0031, December 2016

DIN EN 60601-1-6:2016-02 Medizinische elektrische Geräte - Teil 1-6: Allgemeine Festlegungen für die Sicherheit einschließlich der wesentlichen Leistungsmerkmale - Ergänzungsnorm: Gebrauchstauglichkeit, 2016

DIN EN ISO 14971:2013-04 Medizinprodukte - Anwendung des Risikomanagements auf Medizinprodukte, 2013

Feußner H, Ostler D, Kohn N, Vogel T, Wilhelm D, Koller S, Kranzfelder M (2016) Umfassende Systemintegration und Vernetzung im Operationssaal. Chirurg 87(12):1002–1100

Golatowski F, Janß A, Leucker M, Neumuth T (2018) OR.NET – secure dynamic networks in the operating room and clinic. J Biomed Eng/Biomedizinische Technik 63(1):1–3. https://doi.org/10.1515/bmt-2018-0004, Verlag De Gruyter, Berlin, Deutschland, Januar 2018

Hofinger G (2009) Lernen aus Fehlern im krankenhaus; Systemische Fehlersicht und Zwischenfallberichtssysteme. Der Unfallchirurg 112(6):604–609

Hölscher UM, Heidecke CD (2012) Defizite in der Patientensicherheit, Deutsche Zeitschrift für Klinische Forschung – DZKF, 03/04, pp 102–105, Mediengruppe Oberfranken – Fachverlage GmbH & Co. KG

Hübler M, Möllemann A, Metzler H, Koch T (2007) Fehler und Fehlermeldesysteme in der Anästhesiologie. Anaesthesist 56:1067–1072

Hyman WA (1994) Errors in the use of medical equipment. In: Bogner MS (ed) Human error in medicine. Erlbaum Publ, Hillsdale

Ibach B (2011) Konzeption und Entwicklung einer serviceorientierten Integrationsarchitektur für die Vernetzung von Medizinprodukten im Operationssaal. In: Leonhardt S, Radermacher K, Schmitz-Rode T (eds) 8, Aachener Beiträge zur Medizintechnik (ISBN 978-3-8440-0298-0), Shaker, 2011, (Dissertation RWTH Aachen, 2011)

Ibach B, Kanert A, Radermacher K (2008) Concept of a service-oriented integration architecture for the orthopaedic operating theatre. Proc CARS 3(Suppl 1):S446–S447

Ibach B, Benzko J, Janß A, Radermacher K (2009) Sharing tracking data between different systems using an Open Surgical Communication Bus. In: Dössel O, Schlegel WC (eds) IFMBE Proceedings, 25/V, pp 21–23

Ibach B, Benzko J, Schlichting S, Zimolong A, Radermacher K (2012) Integrating medical devices in the operating room using service-oriented architectures. Biomed Tech (Berl) 57(4):221–228

IEEE 11073-10207:2017 - IEEE Health informatics - Point-of-care medical device communication Part 10207: Domain Information and Service Model for Service-Oriented Point-of-Care Medical Device Communication

IEEE 11073-20701:2018 – IEEE Health informatics – Point-of-care medical device communication Part 20701: Service-Oriented Medical Device Exchange Architecture and Protocol Binding

IEEE 11073-20702:2016 - IEEE Standard for Health informatics--Point-of-care medical device communication Part 20702: Medical Devices Communication Profile for Web Services

ISO/IEC 24752-1:2014 Information technology -- User interfaces -- Universal remote console - Part 1: General framework, 2014

Janß A, Benzko J, Merz P, Dell'Anna J, Strake M, Radermacher K (2014) Development of medical device UI-profiles for reliable and safe human-machine-interaction in the integrated operating room of the future. In: Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics 2014, pp 1855–1860

Janß A, Plogmann S, Radermacher K (2016) Human centered risk management for medical devices - new methods and tools. In: Dössel O (ed) Journal of Biomedical Engineering/Biomedizinische Technik, Band 61, Heft 2, Seiten 165–181, ISSN (Online) 1862-278X, ISSN (Print) 0013-5585. https://doi.org/10.1515/bmt-2014-0124

Janß A, Vitting A, Strathen B, Strake M, Radermacher K (2017) Intraoperative workflow optimization by using a universal foot switch for open integrated OR systems in orthopaedic surgery. In: Radermacher K, Baena FRY (eds) CAOS 2017. 17th Annual Meeting of the International Society for Computer Assisted Orthopaedic Surgery, vol 1, pp 248–253

Janß A, Thorn J, Schmitz M, Mildner A, Dell'Anna-Pudlik J, Leucker M, Radermacher K (2018) Extended device profiles and testing procedures for the approval process of integrated medical devices using the IEEE 11073 communication standard. J Biomed Eng/Biomedizinische Technik 63(1):95–103, Verlag De Gruyter, Berlin, Deutschland

Kasparick M, Schlichting S, Golatowski F, Timmermann D (2015) New IEEE 11073 standards for interoperable, networked point-of-care medical devices. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp 1721–1724

Kasparick M, Andersen B, Ulrich H, Franke S, Schreiber E, Rockstroh M, Golatowski F, Timmermann D, Ingenerf J, Neumuth T (2018a) IEEE 11073 SDC and HL7 FHIR - emerging standards for interoperability of medical systems. In: International Journal of Computer Assisted Radiology and Surgery: CARS 2018 - Computer Assisted Radiology and Surgery: Proceedings of the 32nd International Congress and Exhibition, vol 13, nr 1, pp 135–136, Springer Verlag, Berlin, Deutschland, Juni 2018

Kasparick M, Schmitz M, Andersen B, Rockstroh M, Franke S, Schlichting S, Golatowski F, Timmermann D (2018b) OR.NET: a service-oriented architecture for safe and dynamic medical device interoperability. J Biomed Eng/Biomedizinische Technik 63(1):11–30. https://doi.org/10.1515/bmt-2017-0020, ISSN: 1862-278X, Verlag De Gruyter, Berlin, Deutschland, Januar 2018

Klein M (2010) Psychological and physical stress in surgeons operating in a standard or modern operating room. Surg Laparosc Endosc Percutan Tech 20:237–242

Kohn LT, Corrigan J, Donaldson MS (2000) To err is human: building a safer health system. National Academy Press, Washington, DC

Kücherer C, Benzko J, Bußhaus T (2013) Forschungsprojekt OR.NET: Ein Abschied von Insellösungen. Deutsches Ärzteblatt 110(46):8–10

Lauer W, Radermacher K (2007) Simulation - a valuable tool for prospective usability-assessment of complex OR-components. In: Rötting M, Wozny G, Klosterman A, Huss J (eds) Prospektive Gestaltung von Mensch-Technik-Interaktion, 7. Berliner Werkstatt Mensch-Maschine-Systeme, vol 22(25), pp 47–52

Lauer W, Ibach B, Radermacher K (2008) Knowledge-based OR table positioning assistant for orthopedic surgery. In: Sloten JV, Verdonck P, Nyssen M, Haueisen J (eds) IFMBE Proceedings, vol 22, pp 1676–1678

Lauer W, Janß A, Radermacher K (2009) Risiken der Mensch-Computer-Interaktion im OP. In: Kain S, Struve D, Wandte H (Hrsg) Workshop-Proceedings der Tagung Mensch & Computer 2009, Logos Verlag, S. 305–307

Lauer W, Janß A, Ibach B, Radermacher K (2010) Man-machine-interaction in complex intraoperative orthopedic work systems. In: Duffy VG (ed) Advances in human factors and ergonomics in healthcare, pp 825–833

Leape LL (1994) The preventability of medical injury. In: Bogner MS (Hrsg) Human error in medicine. Erlbaum Publ., Hillsdale

Matern U (2006) Arbeitsbedingungen und Sicherheit am Arbeitsplatz OP, Deutsches Ärzteblatt, 24. November 2006, 103(47):3187–3192

Merten M (2007) Risikomanagement: Den Ursachen auf der Spur. Deutsches Ärzteblatt, Deutscher Ärzte-Verlag GmbH

Montag K (2012) Untersuchung zur Gebrauchstauglichkeit von Medizinprodukten im Anwendungsbereich OP, Dissertation. http://nbn-resolving.de/urn:nbn:de:bsz:21-opus-61107

Nowatschin S (2009) CIO-Computer Integrated Operationroom - Neue Konzepte und Systeme für einen Computer-Integrierten Operationssaal (Dissertation TU München)

Okamoto J, Masamune K, Iseki H, Muragaki Y (2018) Development concepts of a Smart Cyber Operating Theater (SCOT) using ORiN technology. J Biomed Eng 63(1):31–37, Verlag De Gruyter, Berlin, Deutschland

Radermacher K, Pichler C, Erbse S, Boeckmann W, Rau G, Jakse G (1996) Using human factor analysis and VR simulation techniques for the optimization of the surgical worksystem. In: Weghorst SJ, Sieburg HB, Morgan K (eds) MMVR 4 - health care in the information age -future tools for transforming medicine. IOS Press, pp 532–541

Rau G, Radermacher K, Pichler Cv (1994) A helping hand - motivations and perspectives of an increased technical support in surgery. Medical Technology International, 5/1994, Cornhill Publications Ltd., pp 47–52

Rau G, Pichler Cv, Radermacher K (1995) Surgical reality. Medical Technology International, Cornhill Publications Ltd., pp 46–51

Rau G, Radermacher K, Thull B, Pichler C (1996) Aspects of ergonomic system design applied to medical worksystems. In: Taylor RH (ed) Computer-integrated surgery: technology and clinical applications. MIT Press, pp 203–221

Rockstroh M, Franke S, Dees R, Merzweiler A, Schneider G, Dingler M, Dietz C, Pfeifer J, Kühn F, Schmitz M, Mildner A, Janß A, Dell'Anna Pudlik J, Köny M, Andersen B, Bergh B, Neumuth T (2018) From SOMDA to application - integration strategies in the OR.NET demonstration sites. Biomed Tech (Berl) 63(1):69–80

Strathen B, Janß A, Gödde P, Radermacher K (2017) Sensor integrated instruments for surgical quality assurance and process monitoring. In: Radermacher K, Rodriguez F, Baena Y (eds) CAOS2017. EPiC Series in Health Sciences, vol 1, pp 264–268

Strauss G, Gollnick I, Neumuth T, Meixensberger J, Lueth TC (2013) The "Surgical Deck": a new generation of integrated operational rooms for ENT. Laryngo-Rhino-Otologie 92:102–112

Vitting A, Janß A, Strathen B, Strake M, Radermacher K (2017) Further development and evaluation of a universal foot switch for diverse medical disciplines within the framework of an open integration concept for the operation theatre of the future. In: Proceedings Duffy V, Lightner N (eds) Advances in human factors and ergonomics in healthcare and medical devices, 8th International Conference on Applied Human Factors and Ergonomics (AHFE 2017). Springer International Publishing, pp 438–449

Vossel M, Strathen B, Kasparick M, Müller M, Radermacher K, de La Fuente M, Janß A (2017) Integration of robotic applications in open and safe medical device IT networks using IEEE 11073 SDC. In: Radermacher K, Rodriguez F, Baena Y (eds) CAOS2017. EPiC Series in Health Sciences, vol 1, pp 254–257

Wallwiener D, Wallwiener M, Abele H, Rothmund R, Becker S, Zubke W (2011) Integrierte OP - Systeme (IOPS) als Basis für innovative Operationsverfahren in der Gynäkologie. Der Gynäkologe 44:187–195

WHO 2008: World alliance for patient safety - implementation manual surgical safety checklist (first edition), WHO Press, World Health Organization, Geneva, Switzerland, 2008. (https://www.who.int/patientsafety/safesurgery/tools_resources/SSSL_Manual_finalJun08.pdf)

Zeißig EM, Janß A, Dell'Anna-Pudlik J, Ziefle M, Radermacher K (2016) Development and experimental evaluation of an alarm concept for an integrated surgical workstation. In: Dössel O (ed) Journal of Biomedical Engineering/Biomedizinische Technik, Biomedical Engineering/Biomedizinische Technik. Band 61, Heft 2, Seiten 199–209, ISSN (Online) 1862-278X, ISSN (Print) 0013-5585, April 2016

Zimolong A (2004) Verfahren zur Zuverlässigkeitsbeurteilung der Mensch-Maschine-Interaktion bei computergestützten Chirurgiesystemen. (Diss. RWTH-Aachen), Shaker-Verlag

Zimolong A, Radermacher K, Stockheim M, Zimolong B, Rau G (2003) Reliability analysis and design in computer-assisted surgery. In: Stephanides C et al (eds) Universal access in HCI. Lawrence Erlbaum Ass. Publ., pp 524–528 (ISBN: 0-8058-4934-3)

Life Science 4.0

Sensor Technology and Machine Learning in Motion Analysis

Marion Mundt, Arnd Koeppe, Franz Bamer, and Bernd Markert

Contents

M. Mundt
Minderoo Tech and Policy Lab, UWA Law School, The University of Western Australia, Perth, WA, Australia
e-mail: marion.mundt@uwa.edu.au

A. Koeppe
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
e-mail: arnd.koeppe@kit.edu

F. Bamer · B. Markert (✉)
Chair and Institute of General Mechanics, RWTH Aachen University, Aachen, Germany
e-mail: bamer@iam.rwth-aachen.de; markert@iam.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_46

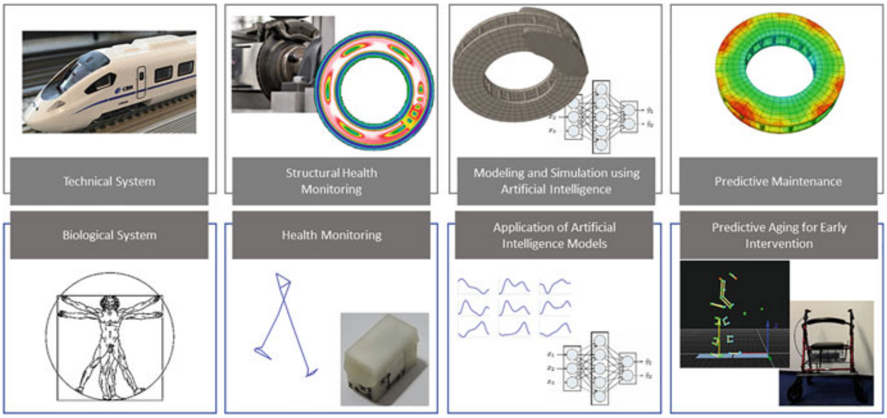


Fig. 1 Industry 4.0 applications in engineering and life science

1 Introduction

Methods, previously established for engineering systems, can be used in life science, or more precisely in biomechanics. Figure 1 shows this connection between applications in engineering and biology exemplarily for a braking disc and motion analysis.

As one part of *Structural Health Monitoring* the heat development of a braking disc is monitored automatically to receive information about possible damage. This is similar to the collection of motion data to supervise motion patterns to detect injury. Computer-assisted modelling and simulation enable the collection of parameters that are very difficult or even impossible to measure. Only for validation of these models and simulations measured data is necessary. Artificial intelligence can be used to further analyse this data. The same workflow can be applied to biomechanics: parameters such as joint moments that cannot be measured are calculated based on multi-body simulations and validated by, e.g. instrumented endoprostheses (Damm et al. 2017). Further, these simulated data can be used to train artificial intelligence models that can be applied to data that can be measured easily. In the field of *Predictive Maintenance* artificial intelligence is not only used to detect but to predict damage before it occurs. The same principle is applied in motion analysis: faulty motion patterns can be detected before injury occurs. Automated feedback systems can inform the user about possibilities to change their motion to a healthier pattern (Reeves and Bowling 2011) or to detect fatigue. Thereby, fatigue induced falls, which are highly relevant in the aging society, can be prevented (Ambrose et al. 2013).

This chapter will provide insight into applications of machine learning techniques to motion analysis. The most relevant methods will be explained, and advantages and disadvantages detailed. Current applications of artificial intelligence in biomechanics will be displayed.

2 Machine Learning to Model Unknown Data

Machine learning describes different statistical method that enable automated data modelling. It can be differentiated between supervised and unsupervised methods. During supervised learning, an algorithm learns the connection between known input and output data. This learned connection can then be applied to unknown input samples to predict their output data. Generally, these algorithms can model the correlation between any kind of data if the algorithms' complexity is high enough (Wernick et al. 2010). One frequently used example of such an algorithm are artificial neural networks. They use hidden layers with a variable number of neurones to model the connection between input and output data. During a training process the neurons' weights are adapted to minimise the error between known and predicted output data (Rumelhart et al. 1986). Further information can be found in Chapter Mechanics 4.0.

3 Motion Analysis

The analysis of motion is important for different fields: in sports it is used to improve performance and understand injury mechanisms (David et al. 2017, 2018; Johnson et al. 2018); in clinical applications it can be used to improve treatment planning (Kay et al. 2000), avoid inappropriate mechanical load before injury by adaptation of new motion patterns (Schülein et al. 2017) or evaluate surgery and rehabilitation outcomes (Ardestani et al. 2014). Due to demographic change and the aging society, mobility and life-long independence gains more and more relevance. Orthopeadic aids such as wheeled walkers are already used frequently (Bradley and Hernandez 2011; Brandt et al. 2003) and will become even more relevant if equipped with a feedback system.

3.1 Measurement Methods

During the last 30 years different methods have been developed to analyse motion. Most of these are based on cameras. For higher accuracy and precision complex methods are used. The current gold standard is based on multiple infrared cameras and retro-reflective markers that are attached to anatomical landmarks of the body to record motion trajectories (Fig. 2).

The recorded marker trajectories are used to calculate joint kinematics such as joint angles over time in all three motion planes. To additionally collect kinetic information, force plates can be used to record the ground reaction force. Based on these, inverse dynamic simulations are performed to determine joint kinetics such as the joint moments over time in all three motion planes (Shahabpoor and Pavic 2017).

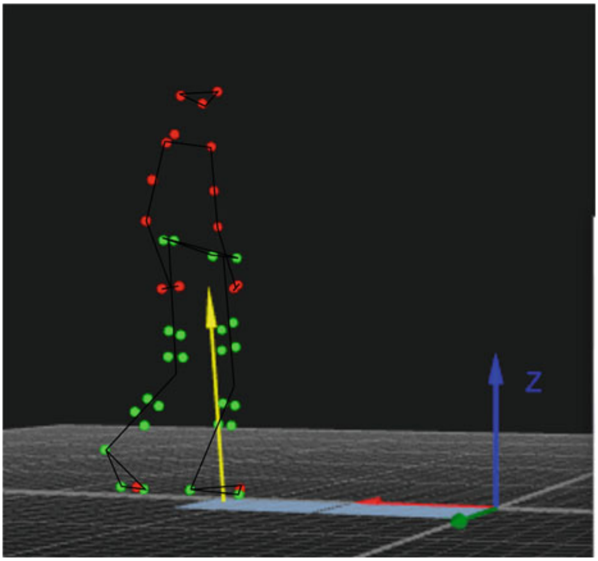


Fig. 2 Standard laboratory setup for motion analysis. Cameras are used to capture retro-reflective markers (green and red) and calculate three dimensional coordinates over time with reference to the global coordinate system. The two force plates in the ground (light grey) measure ground reaction force, which is used in inverse dynamics simulations to determine joint kinetics

Unfortunately, this method shows several limitations: Since this kind of motion capture is relying on cameras and force plates, it is dependent on laboratory conditions to ensure adequate lighting and installation of the force plates. Due to these restrictions, data capture is mostly restricted to a few steps, based on space, and number of cameras and force plates available. The correct and precise attachment of the markers is very important to achieve valid information. A misplacement of only a few millimetres can significantly influence the results (Schwartz et al. 2004). For many applications, the analysis of only a few steps is not sufficient because they do not represent the variability of motion and possible changes might be masked (Sinclair et al. 2013). The evaluation of fast cutting manoeuvres is of high interest for sports practitioners due to the high risk of anterior cruciate ligament injuries during this task, which is repeated up to 100 times per match. Unfortunately, laboratory conditions cannot mimic external factors such as opponents or tactics and internal factors such as fatigue (Lee et al. 2013). Due to the complex setup, motion capture is expensive and only a limited number of people have access.

To overcome these limitations, wearable sensors, e.g. inertial sensors, are used more and more frequently. These sensors incorporate accelerometers to measure the linear acceleration in three dimensions, gyroscopes to measure the angular rate in three dimensions and magnetometer to measure the magnetic field strength. These sensors are cheap and enable measurements in an unrestricted field. Data can be collected and send to another device or into the cloud wirelessly. For this purpose, only an inertial sensor, a micro-controller and energy source are necessary as well as a WiFi or Bluetooth connection (Fig. 3).

Inertial sensors are already frequently used as every smart phone contains an inertial sensor. The magnetometer is used as a compass for improved navigation, the gyroscope enables the rotation of the screen when the phone is turned, and different

Fig. 3 One example of an inertial sensor used for motion analysis. The sensor consists of a WiFi-shield, a micro-controller, an inertial sensor shield, and a lithium-ion-battery. The sensor weighs about 35 g

applications make use of the accelerometer and gyroscope, e.g. to determine the number of steps of a user. This can also be achieved by fitness trackers. These are available as wrist bands or watches from different suppliers. Based on the algorithm that is used within the trackers' software the accuracy varies. Most devices provide information such as the number of steps, calories burnt, activity levels, or sleep quality. Some devices also evaluate spatio-temporal parameters of walking or running gait such as step frequency, step length, or ground contact time.

In research, clinical or sports applications, inertial sensors are also used for more complex tasks. The first step is generally the fusion of the nine signals of the sensor using a Kalman filter to determine the orientation (Sabatini 2006). Due to the use of the magnetometer readings, the gyroscope data that tends to drift over time can be corrected. On the downside, magnetometers are dependent on a constant magnetic field, which is not the case in most laboratories or clinics (de Vries et al. 2009). Most commercial systems use the orientation of the sensor in space to estimate the orientation of the segment the sensor is attached to. Figure 4 displays a frequently used setup for inertial sensors attached to a human's body. Based on a setup like this, the lower limbs' joint angles can be determined.

For this calculation, the sensor-to-segment-assignment needs to be performed in a first step. Thereby, the system knows which sensor describes which segment of the body. This is necessary to determine which sensors need to be used for the calculation of a specific joint angle. If the sensor-to-segment-assignment is incorrect, e.g. right and left are swapped, no valid joint kinematic can be calculated. The second step that needs to be performed is the sensor-to-segment-alignment. During this step, the orientation of the sensor towards the segment is determined (Zimmermann et al. 2018). The joint angles can now be calculated based on an anatomical model that defines the joint axes and thereby the rotation axes. The correct determination of the rotation axes has huge influence on the calculated joint angles. Using classical optical motion capture systems, the anatomical model is based on marker positions. For inertial sensors, this is not possible. Hence, different

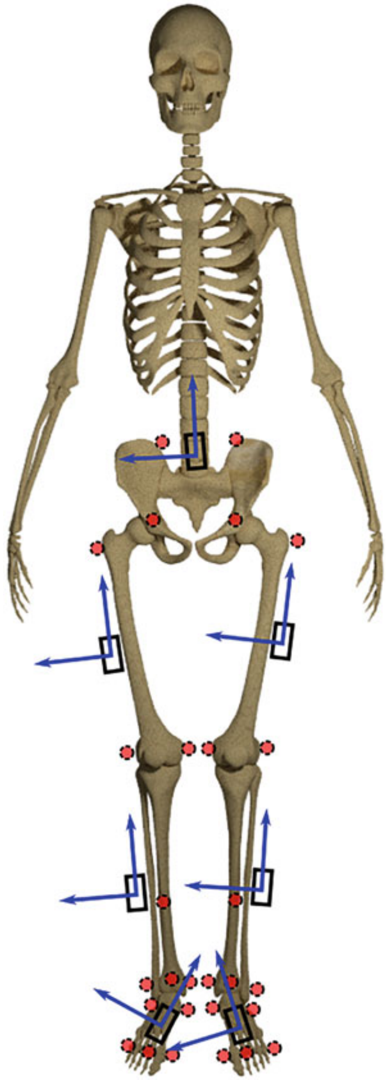


Fig. 4 Inertial sensors and markers attached to the human body. Markers are used by the optical motion capture system

calibration postures and movements have been proposed to determine segment coordinate systems and thereby the rotation axes. All proposed methods have the disadvantage that they are dependent on the execution of the position/movement by the user. If the user does not execute the calibration very precisely, the sensor-to-segment-alignment is faulty and causes flawed joint kinematics (Mundt et al. 2018a, b). To overcome this problem, research has been undertaken to become independent of magnetometers and calibration postures. Many of these attempts are based on machine learning approaches that are already frequently used for engineering applications.

Besides the joint kinematics, information on joint kinetics is at least as important, because this information gives insight into the load applied to the joints during motion. Kinematic parameters cannot be measured by inertial sensors directly, but machine learning approaches show that it is possible to estimate the kinetics of a joint based on its kinematics. Using previously collected datasets—either inertial sensor data or optical motion capture—machine learning models can be trained for this task. Paragraph 4 will detail different examples.

3.2 Machine Learning in Motion Analysis

Until today, high quality applications of machine learning to motion analysis are very limited, although the availability of open-source code for different platforms enables more research in this direction. However, the application of these algorithms without former knowledge in this domain causes problems in quality. Several recent publications show flaws in data handling (Halilaj et al. 2018). A common mistake is the attempt to develop a generalisable model based on a dataset consisting of data on few participants only. This is not possible as the variance of new unknown data cannot be covered by the trained model if the underlying data did not show equivalent variance. Another frequently made mistake is the evaluation of a models' performance on training data instead of on separate test data. The prediction accuracy of the training dataset will always be high (if the model was trained long enough and has a sufficient complexity), while this does not include any information on the model's performance on unseen data. Additionally, the necessity of adequate knowledge on the field of application should not be underestimated, because the algorithms are generally capable to find correlations between data that might physically be uncorrelated. Based on expert knowledge, appropriate input and output data can be chosen for the prediction. This is especially relevant for applications that are based on rather small datasets, such as those available in biomechanics.

The use of machine learning to predict joint kinematics based on inertial sensor data is very limited. Zimmermann et al. (2018) used artificial neural networks to automate the error-prone sensor-to-segment-assignment and sensor-to-segment-alignment. Based on this, the well-established methods to determine joint angles can be applied. Another research (Findlow et al. 2008; Goulermas et al. 2005, 2008; Mundt et al. 2020b, c) has already gone one step further and used artificial neural networks to directly predict joint angles based on inertial sensor data, thereby being independent on an anatomical model at all. However, this research was carried out more than 10 years ago. Based on the increased computing power nowadays, it can be expected that even better results can be achieved.

The determination of joint kinetics based on machine learning is a very active field of research. These applications are even more interesting as it is not only about improved accuracy to standard methods but more about simplifying measurement setups and thereby improving applicability. The determination of joint kinetics based on inertial sensors enables in-field measurements as no force plates are necessary and the time-consuming calculation of multi-body simulations becomes obsolete (Mundt et al. 2020a). Many applications focus on the analysis of a single joint or motion plane only (Aljaaf et al. 2016; Favre et al. 2012; Hahn 2007; Osateerakun et al. 2018) or are limited to a small sample size (Ardestani et al. 2014; Osateerakun et al. 2018). In sports applications machine learning is used to determine the ground reaction force or joint moments. Different motion has been analysed: running (Jie-han Ngoh et al. 2018), fast changes of direction (Johnson et al. 2018, 2019a, b, 2020; Mundt et al. 2019), lifting (Kipp et al. 2018) or jumps (Liu et al. 2009).

In summary, it has been shown that methods of Industry 4.0 are not limited to engineering applications but can support applications in daily life as well. First steps in this direction are promising, but until now, the availability of large biomechanical datasets is very limited, hence, many applications are still restricted, and the real potential of machine learning is not yet exploited.

4 Practical Examples

The following sections will detail three examples of artificial neural network applications to simplify measurement setups for motion analysis and avoid costly inverse dynamics calculations.

4.1 Prediction of Joint Moments Based on Joint Angles During Gait

As described in Sect. 3, the standard measurement setups to determine kinematic and kinetic motion parameters are very complex, which makes them expensive and hardly available to many people. Based on a sufficiently large dataset of joint kinematics and kinetics, artificial neural networks can be trained to predict joint moments based on joint angles. The calculation of joint angles is easier, faster, and less limiting than the determination of joint moments. The following example will evaluate the prediction accuracy of two different neural networks on this task. Besides a basic multilayer perceptron neural network (Fig. 5), a more sophisticated recurrent neural network will be used.

Due to its architecture, a recurrent neural network can learn time dependencies in the data. Therefore, in contrast to a multilayer perceptron network, it is not necessary to time-normalise input and output data, which enables applications in real time (Mundt et al. 2020b, c). However, the prediction accuracy of a recurrent network might be limited, because only the current time step of the data is visible for the network and it needs to remember all previous information, while a multilayer perceptron (MLP) has access to a complete motion sequence. Hence, this network does not learn any time dependencies, but also does not need to, to be able to achieve

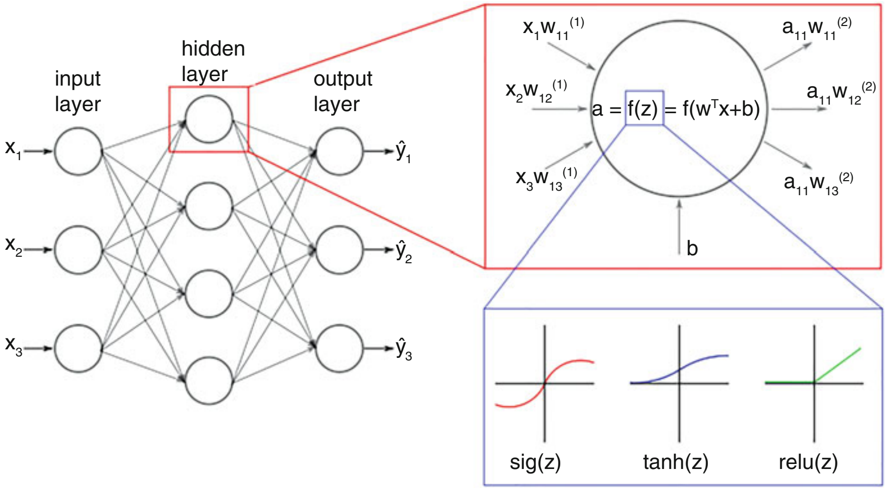


Fig. 5 Multilayer perceptron neural network

high accuracy. This makes the task for the recurrent network more complex. In this example, a long short-term memory (LSTM) neural network is used as one example of a recurrent neural network. These networks show the capability to remember longer time sequences (Koeppe et al. 2017).

A previously collected dataset of the German Sport University Cologne has been used, which comprises about 1500 steps of 85 participants. Each participant performed level walking trials. Joint angles and moments have been calculated based on optical motion capture data. Both angles and moments have been time-normalised to 100% stance phase and joint moments have been normalised to weight and height of the participant. Joint angles have been used as input data, while joint moments form the outputs. The dataset has been split manually into training, validation and test set to ensure that all data of one participant is part of one dataset only, hence, there is no data leakage. For both neural networks, the MLP and LSTM, a parameter study has been executed to find the optimum number of layers, neurones and hyperparameters. Figure 6 shows mean and standard deviation of measured and predicted hip, knee, and ankle joint moments in all motion planes for both networks. The predicted mean values of both networks are slightly different from the measured ones, but the progression in general can be captured well. The prediction accuracy of the LSTM is not as good as those of the MLP.

To quantify the prediction error, Fig. 7 shows the normalised root-mean-squared-error of single predicted joint moments. The violin plot shows that the accuracy of the MLP is higher than those of the LSTM. The mean error of the MLP is smaller than the one of the LSTM and the shape of the violin shows that more data is accumulated around smaller error values for the MLP. Interestingly, larger errors can be found in those motion planes with smaller ranges of motion (frontal and transverse). These planes already show larger variance in their input values, which might

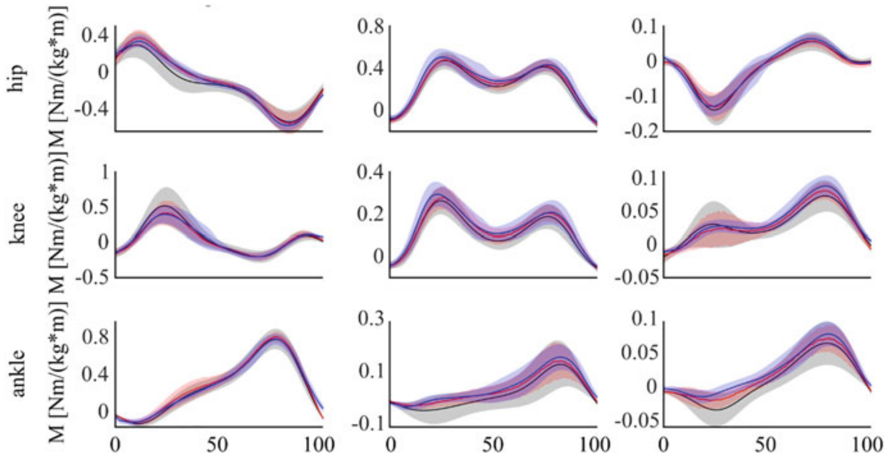


Fig. 6 Mean and standard deviation of measured (grey) and predicted joint moments using an MLP (red) and LSTM (blue) neural network

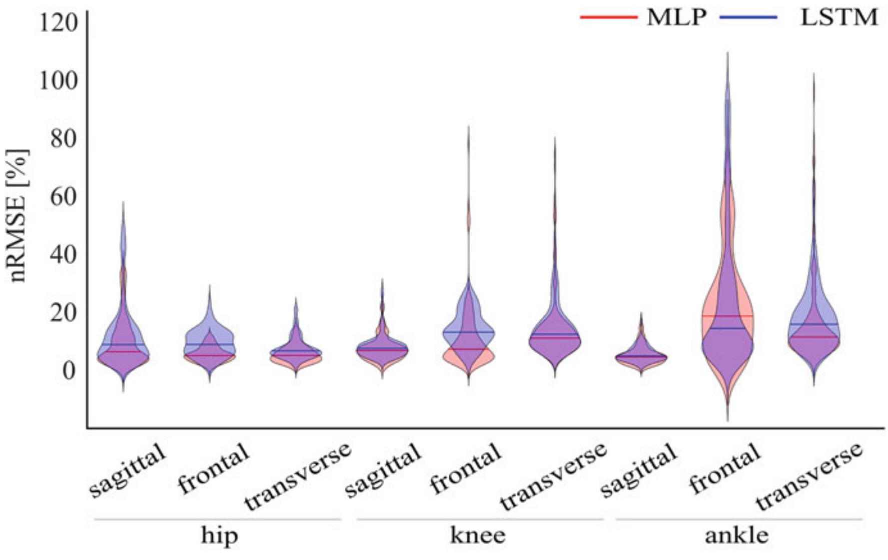


Fig. 7 Normalised root-mean-squared error of the prediction of joint moments using an MLP (red) and LSTM (blue) neural network for each joint and motion plane

explain the difficulty in predicting the associated outputs. This is most obvious for the knee joint moment.

Generally, the agreement between measured and predicted values is high, with a mean error of less than 20% in all motion planes. Probably a larger dataset will further improve the accuracy. These results indicate that it will be possible to predict joint moments based on joint angles that can be measured by inertial sensors.

Thereby, further insight into joint load in non-laboratory conditions can be granted. The usage of models and simulations might allow for the development of larger datasets to further improve neural network training. Following this training process, the neural network model can be applied with less computational effort. An LSTM can be used for real time applications, which are necessary for a feedback system that can be used for a predictive aging application. Based on time normalised data the LSTM shows a high accuracy (Mundt et al. 2020b). In other work, it was shown that complete motion sequences without time-normalisation can be used as input. Thereby, a further step towards real time application has been made (Mundt et al. 2020c).

4.2 Prediction of Ground Reaction Force and Joint Moments During Fast Changes of Direction

Fast changes of direction are highly relevant in all kinds of team sports. They are executed more than 100 times during a match. For this reason, the evaluation of this motion is of high interest for sport practitioners and athletes to improve performance on the one hand and understand injury mechanisms on the other hand (Johnson et al. 2018; David et al. 2017, 2018). Hence, it is important to know the ground reaction force and joint moments during these movements. Unfortunately, the use of force plates has several disadvantages in this setting as it does not allow for the player to move unrestrictedly. This causes deviations in the normal movement pattern. To overcome this, we aimed to determine the ground reaction force and joint moments of hip, knee, and ankle joint during a maximum effort 90° change of direction with the help of a multilayer perceptron (MLP) neural network. This kind of network was chosen as it promises the highest accuracy and using the proposed method, the analysis is still restricted to a laboratory, hence, real time applicability is not necessary yet. In this study, the use of different input parameters has been investigated: full-body marker trajectories, lower-body marker trajectories and lower limb joint angles. An optical motion capture system has been used to record the marker trajectories. In case of a successful prediction based on joint angles, the measurements could be undertaken with inertial sensors, too. The dataset was collected at the German Sport University Cologne and comprises fast changes of direction of 55 athletes. Each change of direction consists of two ground contacts: execution and depart contact. In total, 1050 ground contacts have been measured. The dataset was split into three parts: training, validation, and test set, ensuring that there is no data leakage, i.e. no data of a single participant in more than one of the datasets. To determine the optimum number of layers, neurones and hyperparameters, a parameter study has been executed.

Statistical analysis did not show any differences in the prediction accuracy of joint moments based on the different input parameters. The mean root-mean-squared-error was smaller than 20% for all joint moments and motion planes

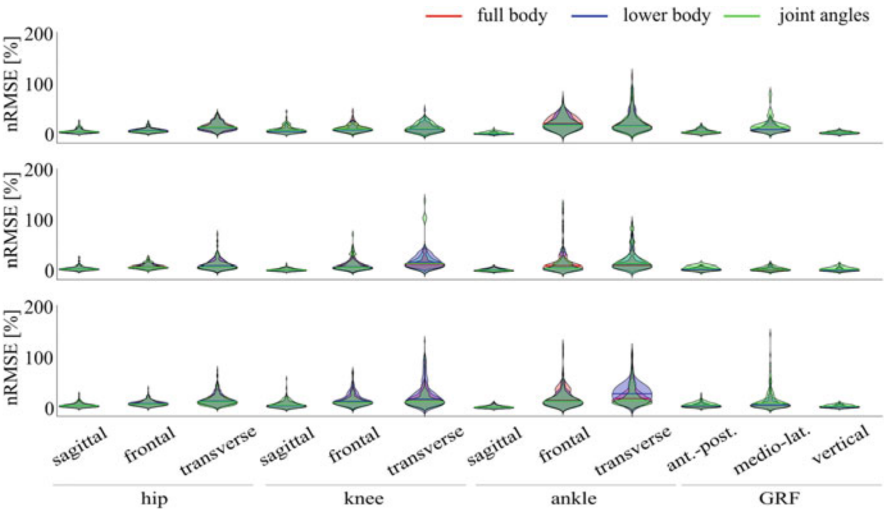


Fig. 8 Normalised root-mean-squared-error for the prediction of joint moments and ground reaction force based on different input parameters. Full body and lower body marker trajectories can only be measured using an optical motion capture system while joint angles can also be measured using inertial sensors

(Fig. 8). The prediction of the ground reaction force showed higher accuracy when using marker trajectories as inputs compared to joint angles. This can be explained by the closer physical relationship between marker trajectories and force compared to joint angles and force. For both joint moments and ground reaction force, several outliers in the prediction accuracy can be found. This might be attributed to the rather small dataset used for a motion that shows high variance. Another factor that might have influenced the outcome is the data pre-processing, especially filtering (Mundt et al. 2019).

4.3 Prediction of Ground Rection Force of Older People With and Without the Support of a Four-Wheeled Walker

Four-wheeled walkers are highly visible in our daily life. They are relevant for the mobility of many older and disabled people and enable them to live independently. However, the biomechanical consequences of using a four-wheeled walker have not been evaluated sufficiently yet. One reason for this is the difficulty of measuring kinetic parameters of older people's gait, because older people generally show a smaller step length. In a standard setup based on force plates, this regularly leads to double contacts on the force plate, which makes the trial unusable (Fig. 9).

To overcome this problem, we aimed to determine the ground reaction force of older people's gait with the help of artificial neural networks and inertial sensors

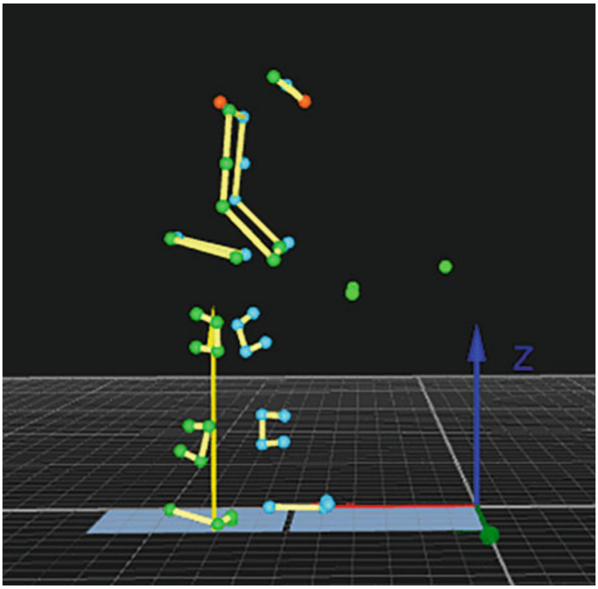


Fig. 9 Data collection for four wheeled walker supported gait. The participant shows a short step length, which causes double contacts on both force plates

placed on the participant's feet. A multilayer perceptron (MLP) neural network was used to model the correlation between the physically closely related acceleration and angular rate and ground reaction force. In cooperation with the Department of Geriatrics, RWTH Aachen University Hospital the gait of healthy older people (>65 years) and geriatric inpatients with and without the support of a four-wheeled walker was investigated. The collected data comprises 268 steps on a force plate by 28 participants. To augment the dataset, all inertial sensor data was mapped to the steps performed on the force plate, hypothesising a stable motion pattern. This processing step resulted in 1107 steps. Using a time warping algorithm, single steps could be extracted (Mundt et al. 2018a, b) and normalised in time to be used as input to an MLP network. To account for the small sample size, a five-fold cross-validation was performed ensuring that no data of a single participant is in more than one set of training, validation, or test set. It was found that the prediction accuracy is dependent on the participants in the test set, hence a larger dataset would be advantageous. The mean results show a high prediction accuracy for the vertical and anterior-posterior direction of the ground reaction force, but a low accuracy for the medio-lateral component (Fig. 10). The difficulty in predicting the medio-lateral component might be explained by the fact that this direction shows the lowest force and high variance.

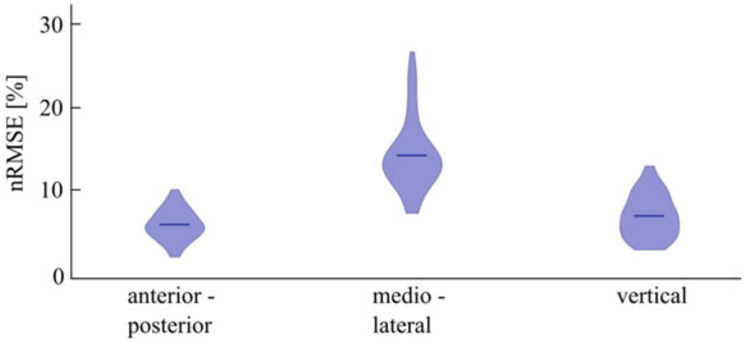


Fig. 10 Results of the ground reaction force prediction based on inertial sensor data

5 Summary and Outlook

The three practical examples on the use of artificial neural networks in biomechanics show the great potential of this method to simplify motion analysis. Costly setups and measurement technologies can be replaced by easy-to-use, cost-efficient technology if methods of Industry 4.0 are used in life science. These developments allow for motion analysis to become available for a larger range of people in daily life and clinical decision making. Therapy can be optimised by performing pre- and post-surgery motion analysis, rehab can be objectively evaluated and thereby improved. In future, it might be possible to integrate small measurement systems into feedback systems that can be used during daily life. These will help especially older people to maintain their independence.

References

Aljaaf AJ et al (2016) Evaluation of machine learning methods to predict knee loading from the movement of body segments. In: Proceedings of the International Joint Conference on Neural Networks, pp 5168–5173

Ambrose AF, Paul G, Hausdorff JM (2013) Risk factors for falls among older adults: a review of the literature. Maturitas 75(1):51–61

Ardestani MM et al (2014) Human lower extremity joint moment prediction: a wavelet neural network approach. Expert Syst Appl 41(9):4422–4433

Bradley SM, Hernandez CR (2011) Geriatric assistive devices. Am Fam Physician 84(4):405–411

Brandt A, Iwarsson S, Ståhl A (2003) Satisfaction with rollators among community-living users: a follow-up study. Disabil Rehabil 25(7):343–353

Damm P et al (2017) Comparison of in vivo measured loads in knee, hip and spinal implants during level walking. J Biomech 51:128–132

David S et al (2017) Identification and risk estimation of movement strategies during cutting maneuvers. J Sci Med Sport 20(12):1075–1080

David S, Mundt M, Komnik I, Potthast W (2018) Understanding cutting maneuvers – the mechanical consequence of preparatory strategies and foot strike pattern. Hum Mov Sci 62 (October):202–210

de Vries WHK, Veeger HEJ, Baten CTM, van der Helm FCT (2009) Magnetic distortion in motion labs, implications for validating inertial magnetic sensors. Gait Posture 29(4):535–541

Favre J, Hayoz M, Erhart-Hledik JC, Andriacchi TP (2012) A neural network model to predict knee adduction moment during walking based on ground reaction force and anthropometric measurements. J Biomech 45(4):692–698

Findlow AH et al (2008) Predicting lower limb joint kinematics using wearable motion sensors. Gait Posture 28(1):120–126

Goulermas JY et al (2005) Regression techniques for the prediction of lower limb kinematics. J Biomech Eng 127(6):1020

Goulermas JY et al (2008) An instance-based algorithm with auxiliary similarity information for the estimation of gait kinematics from wearable sensors. IEEE Trans Neural Netw 19(9):1574–1582

Hahn ME (2007) Feasibility of estimating isokinetic knee torque using a neural network model. J Biomech 40(5):1107–1114

Halilaj E et al (2018) Machine learning in human movement biomechanics: best practices, com-mon pitfalls, and new opportunities machine learning in human movement biomechanics: best practices, common pitfalls, and new opportunities. J Biomech

Jie-han Ngoh K, Gouwanda D, Gopalai AA, Yu C (2018) Estimation of vertical ground reaction force during running using neural network model and uniaxial accelerometer. J Biomech 76: 269–273

Johnson WR et al (2018) Predicting athlete ground reaction forces and moments from motion capture. Med Biol Eng Comput 56(10):1781–1792

Johnson WR, Alderson JA, Lloyd D, Mian AS (2019a) Predicting athlete ground reaction forces and moments from spatio-temporal driven CNN models. IEEE Trans Biomed Eng 66(3): 689–694

Johnson WR, Mian A, Lloyd DG, Alderson JA (2019b) On-field player workload exposure and knee injury risk monitoring via deep learning. J Biomech 93:185–193

Johnson WR, Mian A, Robinson MA, Verheul J, Lloyd DG, Alderson JA (2020) Multidimensional ground reaction forces and moments from wearable sensor accelerations via deep learning. IEEE Trans Biomed Eng:1–12

Kay RM et al (2000) The effect of preoperative gait analysis on orthopaedic decision making. Clin Orthop Relat Res 372:217–222

Kipp K, Giordanelli M, Geiser C (2018) Predicting net joint moments during a weightlifting exercise with a neural network model. J Biomech 74:225–229

Koeppe A, Bamer F, Hernandez Padilla CA, Markert B (2017) Neural network representation of a phase-field model for brittle fracture. Proc Appl Math Mech 17:253–254

Lee MJC et al (2013) Effects of different visual stimuli on postures and knee moments during sidestepping. Med Sci Sports Exerc 45(9):1740–1748

Liu Y et al (2009) Lower extremity joint torque predicted by using artificial neural network during vertical jump. J Biomech 42(7):906–911

Mundt M, Thomsen W, Bamer F, Markert B (2018a) Determination of gait parameters in real-world environment using low-cost inertial sensors. Proc Appl Math Mech 18:5–6

Mundt M, Thomsen W, David S et al (2018b) Assessment of the measurement accuracy of inertial sensors during different tasks of daily living. J Biomech

Mundt M, David S, Koeppe A, Bamer F, Markert B, Potthast W (2019) Intelligent prediction of kinetic parameters during cutting manoeuvres. Med Biol Eng Comput 57(8):1833–1841

Mundt M, Koeppe A, Bamer F, David S, Markert B (2020a) Artificial neural networks in motion analysis — applications of unsupervised and heuristic feature selection techniques. Sensors 20: 1–15

Mundt M, Koeppe A, David S, Bamer F, Potthast W, Markert B (2020b) Prediction of ground reaction force and joint moments based on optical motion capture data during gait. Med Eng Phys 86

Mundt M, Koeppe A, David S, Witter T, Bamer F, Potthast W, Markert B (2020c) Estimation of gait mechanics based on simulated and measured IMU data using an artificial neural network. Front Bioeng Biotechnol 8(February):41

Osateerakun P et al (2018) P 037 – Prediction of moments from movements without force platforms using artificial neural networks: a pilot test. Gait Posture 65(xxxx):299–300

Reeves ND, Bowling FL (2011) Conservative biomechanical strategies for knee osteoarthritis. Nat Rev Rheumatol 7(2):113–122

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(6088):533–536

Sabatini AM (2006) Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing. IEEE Trans Biomed Eng 53(7):1346–1356

Schülein S et al (2017) Instrumented gait analysis: a measure of gait improvement by a wheeled walker in hospitalized geriatric patients. J NeuroEng Rehabil 14(1):18

Schwartz MH, Trost JP, Wervey RA (2004) Measurement and management of errors in quantitative gait data. Gait Posture 20(2):196–203

Shahabpoor E, Pavic A (2017) Measurement of walking ground reactions in real-life environments: a systematic review of techniques and technologies. Sensors 17(9):2085

Sinclair J et al (2013) Three-dimensional kinematic comparison of treadmill and overground running. Sports Biomech 12(3):272–282

Wernick MN et al (2010) Drawing conclusions from medical images. IEEE Signal Process Mag (July):25–38

Zimmermann T, Taetz B, Bleser G (2018) IMU-to-segment assignment and orientation alignment for the lower body using deep learning. Sensors (Switzerland)

The Fourth Industrial Revolution: Information Security and Data Protection

Thomas Jäschke

Contents

1 Introduction

Due to the exponential progress in terms of technological developments regarding the Fourth Industrial Revolution or Internet of Things (IoT), many speculations are arising. However, the significance of information security and data protection is not part of any discussions because the legislative branch as well as companies and clients—the data subjects—already think of those terms as important. The significance of those terms plays an essential role in involved companies' existing and upcoming processes of change.

There are various promising approaches, although they require a serious change in awareness, in a similar manner to the Fourth Industrial Revolution, which is in full swing. Data receive an essential meaning on their way from the analog to the digital world and become a value and asset. Data evolve into an essential component of economic value added.

With an increasing complexity of technologies and thus rising possibilities to collect data masses, greediness is also growing. This greediness is not linked to cyber criminals alone. In that same context, industrial companies develop new business models and strategies for sales approaches. By using technological developments,

T. Jäschke (✉)
DATATREE AG, Düsseldorf, Germany
e-mail: thomas.jaeschke@datatree.ag

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_47

the data processor's responsibility increases. Study results confirm this development. In the context of a research, the Fraunhofer Institute examined fears that accompany the Fourth Industrial Revolution: "The most frequently raised concern regarding the Fourth Industrial Revolution is that data are not safe, business secrets get lost, and internal corporate matters are revealed to the competition" (Fraunhofer-Gesellschaft 2016).

The Fourth Industrial Revolution provides economic advantages—regardless of existing risks. This is another reason for companies to implement strategies and look for solutions in regard to efficient safety and security infrastructures.

2 Problems

The 2018 annual report of the Federal Office for Information Security (BSI) clarifies the current situation: Attacks on IT infrastructures in companies are no fringe phenomenon anymore, but part of the daily routine. During the previous two years, around 70% of German companies had to suffer a cyber-attack (Bundesamt für Sicherheit in der Informationstechnik, *Die Lage der IT-Sicherheit*). The extent of this situation can be comprehended only after considering the success rate of hacking incidents, which is around 50% (Bundesamt für Sicherheit in der Informationstechnik, *Die Lage der IT-Sicherheit*). In many of those cases, the attackers could access IT systems, influence their operating mode, or manipulate web presences. Of those cases, 50% had production and business downtimes as consequences (Bundesamt für Sicherheit in der Informationstechnik, *Die Lage der IT-Sicherheit*). Such attacks are made possible by an IT structure that is becoming more and more complex. It is irrelevant whether it is the IT infrastructure, or a software programmed in a complex manner. "On the Internet of Things, networked devices and assistants offer an easy target, which cyber criminals have actively used for years. Aside from attacks on the availability of devices, made-to-measure malware can also completely regulate the compromised systems" (Bundesamt für Sicherheit in der Informationstechnik 2018, *Mit Sicherheit*). The problem is intersectoral, thus critical infrastructures are as much affected as medium-sized businesses. The better a business is positioned in the market, the more likely it becomes a target for cyber-attacks (Bundesamt für Sicherheit in der Informationstechnik 2018, *Mit Sicherheit*). "The Fourth Industrial Revolution is based on a high interconnectedness of internal and corporate areas within the value chain. This network of multiple IT systems with different protections provides attackers with new possibilities to invade networks" (Pistorius 2020). Attacks using ransomware, phishing mails, but also attacks by means of infiltration of manipulated removable devices and conventional social engineering demonstrate the diversity of this spectrum: According to the current annual report of the Federal Office for Information Security "the number of new malware variations increased by 117.4 million" (Bundesamt für Sicherheit in der Informationstechnik 2020). Thus, there is no patent remedy to feel save concerning the area of information security. Illusions of a perfect security for systems and data need to be abandoned. Every security system can be conquered, if the measures used are strong enough (Fraunhofer-Gesellschaft 2016).

Humans will remain one of the biggest flaws. Thus, a change in awareness regarding information security and data protection is overdue, including necessary strategies, concepts, and measures. However, the awareness of employees needs to be addressed and established, too. This change is in full swing, such as the Fourth Industrial Revolution itself. Although a few years ago, many interviewees said that information security and data protection are annoying; by now, the perception prevails that information security is an essential issue of the respective business strategy regarding the ongoing digitalization. In its main features, it is a solid basis for making it accessible to the employees through a continuative operational approach (Jäschke and Domnik 2018). According to a study by the market research institution Prolytics, this issue needs to be improved to be applicable to the target group's working life (Prolytics 2019).

Amongst others, the manufacturing industry is concerned by that. In many cases, a deficient IT security is not an exception. In fact, the biggest flaw is mostly due to the system, but the employee's incompetence in operating plays an important role, as well.

"A fundamental problem of Industrial Control Systems (ICS: IT systems including networks) is the security culture, which does not exist so far, in contrast to the security culture of commercial IT. Process-oriented IT systems (e.g., firmware) are part of the systems, and have longer time frames than IT systems of commercial IT (up to 20 years). IT security mostly is not a primary goal of system manufacturers. Simultaneously, the operator often does not have any detailed knowledge about the used IT technologies" (Aterna and Roeb-Vollmer 2019).

2.1 Information Security

Due to a growing interconnectedness, possibilities of misuses increase, too. By now, whole production chains depend on a working IT system. Thus, the advancing digitalization requires an extension of the security infrastructure. However, this infrastructure may not simply concentrate on IT solutions. A global approach is needed, which explicitly integrates employees to a high degree. Security concepts, which allow for reacting to cyber-threats and other risks for IT systems, are required. Technical measures are only one of the aspects: "Password protection, virus scanners, firewalls, and data back-ups [are part of] the standard protection in German businesses. 91% work with encrypted network connections, 74% have electronical admission control, and 67% protocol every access. However, since malware is becoming more and more complex, those measures often are not enough anymore" (Pistorius 2020). Thus, organizational aspects integrated in the corporate culture should not be disregarded: They can help identifying and assessing risks, and name persons in charge with regard to a management system based on long-term considerations and to a sustainable improvement of business processes (Heidland 2019).

The digital interconnectedness does not only influence the business, which is directly concerned, but several other businesses must cope with negative consequences as well, although they have not been directly attacked. “IoT applications are used as gateways for attacks and data leaks: they provide a big (data) loot and through their interconnectedness open up possibilities of compromising other systems and applications. Other than that, IoT devices also pose completely new dangers: attacks on physical infrastructures could directly endanger people or even kill them – for example by paralyzing vital electrical power and water supplies or by manipulating medical devices or connected cars” (Schreier 2019).

In Germany, medium-sized businesses are especially relevant for attacks. The reasons are evident: In contrast to big corporations, they only have narrow resources but still are tightly linked to other businesses—thus, there are plenty of possibilities to cause damage on a whole different level.

“Three of four businesses – exactly 73 percent – with a number of employees between 100 and 500, have been victim to digital attacks over the last two years. The reasons are easy to explain: The attackers are interested in receiving certain detailed knowledge – so, they use medium-sized businesses’ systems as gateways to access data of big corporations” (Heidland 2019).

2.2 Data Protection

Complying with data protection allegedly is one of the biggest snares and challenges along the way to IoT. The Fourth Industrial Revolution is based on interconnectedness and data exchange, so relevant information accumulates to an unprecedented extent. Next to company secrets, personal data is also included, which is especially protected by law—the GDPR. The accountability is a new addition to protecting personal data: “The controller is accountable for the observance [. . .] and has to be able to prove the observance (accountability),” Art. 5 GDPR. A big challenge is combining possibilities and processes of innovative techniques with legal frameworks. In several cases, individual data allow a conclusion to be drawn about employees, clients, or consumers. The safety of data needs to be guaranteed by technical measures, but also by raising the employees’ awareness to a responsible handling of personal data. Similar to an information security management system, a data protection management system needs to be implemented. Ideally, this management system can be based on a central data and process system. By now, several software solutions are available on the market. However, the above-mentioned management systems rather are a bundling of processes than a software solution. The latter can function as an additional and facilitating tool.

3 Solutions

Although problems arising with the continuing digitalization in the industrial sector are manifold and serious, approaches and strategies to solve those problems already exist. A preferably all-encompassing system, which includes a security strategy amongst other things, is especially worth the attention. A very own information security management system (ISMS) and a data protection management system (DPMS) are some of the most effective measures for protection. Those management systems are suitable for bundling, planning, realization, and controlling of resources, which repeatedly examine processes in regard to their validity and their effectiveness. Thus, the digital process is embedded in a PDCA cycle and management systems. Big institutions and businesses are advised to use IT solutions for documentation, confirmability, and examination of the particular processes' effectiveness on a regular basis.

Furthermore, the implementation of those systems concerns processes that can be quite longsome and complex. In addition to that, the actual business processes are often slowed down.

Almost all measures of information security have one negative side effect: they limit the involved parties' freedom to act. Therefore, a consideration of higher risks, a weighing of security and usability, and an appropriate risk treatment by the management are of major importance (Jäschke, Domnik).

The earlier the general business strategy allows for the topic information security to be considered, the more effective the appropriate tools are. This concept should apply to all new processes and every project within the Fourth Industrial Revolution. Information security and data protection must be considered at an early stage of project management, because only with those matters recognized, "businesses can successfully participate in digitalization and are able to prevent damage from the very beginning," argued Arne Schönbohm, president of the BSI (Bundesamt für Sicherheit in der Informationstechnik 2018, *Mit Sicherheit*). Especially in terms of budget, conflicts arise frequently. Usually, the significance of a solid structure for information security is not appreciated: "Investments in security are similar to the fire department. We should not discard the fire department merely because there hasn't been a fire for the last two years," claims Michael von Röder, CEO of Sensorberg (Paymentandbanking 2018). Many people are aware of the fact that a case of emergency is much more expensive. However, it has to be considered anew from case to case. The tremendous chances of digitalization, which always have a positive impact on the financial statement, in some way must be bought by having an applied security strategy. "Networks of the Fourth Industrial Revolution need special protective measures, a sophisticated network technology, and effective test methods to uncover security flaws and to solidly solve those issues" (Fraunhofer-Gesellschaft 2016).

It is necessary that security is considered in every area of a business, although the realization of protective measures is different for every one of them. The management level is especially concerned. Success and failure of an ISMS or a DPMS can be mostly traced back to the management of an institution. The management is responsible for employing a CISO (Chief Information Security Officer), a data protection officer, or other persons in charge, and has the final say in dealing with major risks (Jäschke and Domnik 2018). With an applied security concept at all levels and with regard to all relevant structures, a business is prepared for hazards that accompany the Fourth Industrial Revolution and is able “to prevent itself from becoming an easy target and to ensure a high stability for newly developing infrastructures. Therefore, the matter of security has to be a solid component of all considerations regarding the Fourth Industrial Revolution. Only a carefully secured production system can cope with current attacks” (Bundesministerium für Wirtschaft und Energie 2016).

References

Aterna A, Roeb-Vollmer S (2019) “IT-Sicherheit im Zuge von Industrie 4.0” [“IT-Security in Regard to the Fourth Industrial Revolution”]. QZ-Online, https://www.qz-online.de/qualitaets-management/qm-basics/digitalisierung/industrie-4-0/artikel/it-sicherheit-im-zuge-von-industrie-4-0-1282001.html

Bundesamt für Sicherheit in der Informationstechnik (2018) Die Lage der IT-Sicherheit in Deutschland 2018 [The Situation of IT-Security in Germany 2018]. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2018.pdf?__blob=publicationFile&v=5

Bundesamt für Sicherheit in der Informationstechnik (2020) Die Lage der IT-Sicherheit in Deutschland 2020 [The Situation of IT-Security in Germany 2020]. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2020.pdf?__blob=publicationFile&v=2

Bundesministerium für Wirtschaft und Energie (2016) Leitfaden: IT-Security in der Industrie 4.0 – Handlungsfehler für Betreiber [Guideline: IT-Security in the Fourth Industrial Revolution – Procedural Errors for Operators]. https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/leitfaden-it-security-i40.pdf;jsessionid=E19D14F1C69F156FE4C9CD656FBA0F68?__blob=publicationFile&v=10

Fraunhofer-Gesellschaft (2016) Trends für Industrie 4.0. [Trends for the Fourth Industrial Revolution]. https://www.fraunhofer.de/content/dam/zv/de/Forschungsfelder/Produktion-Dienstleistung/Trends-fuer-Industrie-40.pdf

Heidland J (2019) “IT-Notfallplanung: Handlungsfähig auch im Ernstfall” [“IT Emergency Plan: Remaining Capable of Acting in Case of Emergency”]. Industrial Production. https://www.industrial-production.de/safety%2D%2D-security/handlungsfaehig%2D%2Dauch-im-ernstfall.htm?thema=datenschutz

Jäschke T, Domnik J (2018) “Einführung eines Managementsystems für Informationssicherheit” [Implementation of a Management System for Information Security”]. In: Jäschke T (ed) Datenschutz und Informationssicherheit im Gesundheitswesen [Data protection and information security in health care]. MWV, Berlin

Paymentandbanking (2018) FinTech-Podcast #167 -IoT & Industrie 4.0. Soundcloud, https://soundcloud.com/paymentandbanking/fintech-podcast-167-iot-industrie-40

Pistorius J (2020) Industrie 4.0 – Schlüsseltechnologien für die Produktion [The Fourth Industrial Revolution – Key Technologies for Production]. Springer Vieweg, Berlin, Print

Prolytics GmbH (2019) “Benchmarkstudie Datenschutz und Informationssicherheit” [“Benchmark Study: Data Protection and Information Security”]. MaBIS

Schreier J (2019) "Smart auch beim Datenschutz – Einige Handlungsempfehlungen" ["Being smart regarding data protection – some recommended actions"]. Industry of Things. https://www.industry-of-things.de/smart-auch-beim-datenschutz-einige-handlungsempfehlungen-a-789756/

Ethics of Digitalization in Industry

Arne Manzeschke and Alexander Brink

Contents

1 Introduction

As profound as the implications of the first Industrial Revolution were, we are now embarking on yet another transformation of our economy, based once again on innovation. The Industrial Revolution of the last two centuries – the *first* Industrial Revolution – was characterized by machines that extended, multiplied, and leveraged our *physical capabilities*. With these new machines, humans could manipulate objects for which our muscles alone were inadequate and carry out physical tasks at previously unachievable speeds. [. . .] The *second* industrial revolution, the one that is now in progress, is based on machines that extend, multiply, and leverage our *mental abilities*. The same controversies on social and

A. Manzeschke (✉)
Lutheran University of Applied Sciences - Evangelische Hochschule Nürnberg, Nürnberg, Germany
e-mail: arne.manzeschke@evhn.de

A. Brink
Bayreuth University, Bayreuth, Germany
e-mail: alexander.brink@uni-bayreuth.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_48

> economic impact are attending this second great wave of automation, only now a new and more profound question has emerged. Though we have always regarded our species as relatively mediocre in physical capacities, this has not been our view with regard to our mental capacity. The very name we have given ourselves, Homo sapiens, defines us as thinking people. [. . .] The spectre of machine intelligence competing even tangentially with that of its creator once again threatens our view of who we are. (Kurzweil 1990, pp. 7f.)

Ray Kurzweil, a distinguished mastermind of an all-round digitalized future (and, in his role as head of technical development at Google today, someone in a strategically prominent position), presented this description of the problem in the 1990s, when the **Internet of Things** and convincing **Artificial Intelligence** were far from reality. Whether one wants to follow his census of a first and second industrial revolution, or rather the census of four striking cuts in the process of industrialization on which this book is based (Industry 1.0: mechanical production with water and steam power, Industry 2.0: mass production with electrical energy, Industry 3.0: automation through electronics and IT, Industry 4.0: digitalization and networking of production) (Schwab 2015), plays only a minor role to questions of orientation. In addition to the social distortions of technical innovation, Kurzweil rightly points to a central problem: the image of the human being, which is changing because of technical developments. "How will we work?" is a less dramatic question than "Who will we be?" One might object, saying that the way we work does not have a considerable influence on how we understand ourselves. In any case, one has to admit that serious moral issues will be put to the test here. What does it mean to lead a self-determined life? What status do machines have when they become our "partners"? How powerful can machines become without endangering us? To reflect on such moral questions is the task of ethics, as the philosopher of technology Gernot Böhme put it: "A moral question in the area of ethics concerned with the formation of an individual mode of living is a question by which it is decided how a person regards himself or herself, and who that person is; a moral question in the field of the public discourse devoted to establishing social norms is a question by which a society regards itself and what it becomes. In each case these are questions in which matters become serious for the individual person or for the society." (Böhme 2001, p. 9)

The examination of moral questions over the course of industrialization is not new. In the 19th century, there were massive economic, ecological, and social effects which were connected to the mechanization of work, with the associated increase in efficiency contrasted with poorer working conditions and power concentrations. The rise of heavy industry began with mechanical production plants operated by water and steam power. Mass production was then introduced, with assembly line work in factory halls. The so-called "social question" was the height of this moral debate. With the process of digitalization, we are experiencing today, the technically feasible has increased enormously, and with it, the question of how we want to shape our future is gaining urgency (IEEE 2018).

These questions seem even more compelling as digitalization is disruptive (Christensen 1997; von Mutius 2017) and will change the world and the human image more than any technical innovation has done before. (Partially) autonomous cars, virtual assistants, smart home technologies, wearables, and new forms of treatment in medicine will determine our lives. While the Ford Model T customer was still a "buyer" of a product in exchange for money, nowadays customers often "pay" with data. The non-transparent collection, processing, manipulation of entire data profiles and increasing dependence and vulnerability of digital infrastructures raise pressing questions of privacy, security, self-determination, and justice. These moral questions cannot be solved by technical devices and procedures, no matter how intelligent they may be. They refer to a genuinely human domain: the demand and the ability to set goals for oneself and to represent and take responsibility for them and the means used to achieve them. The role of digitalization can only be a supporting one – it must not be raised to an "end in itself," but ultimately serves human beings and their determination of goals.

This description shows that the processing of moral questions in the field of digitalization can draw on long-established methodological and material resources, and that the ethical reflection of digitalization processes does not raise fundamentally new ethical questions. On the other hand, self-criticism must not fail to consider whether and in which way ethics is called upon to rethink its assumptions, concepts, and methods, insofar as digitalization shifts the coordinates. Thus, the description found in public representations, that a robotically embodied **Artificial Intelligence** "makes decisions," "performs actions" and "interacts socio-emotionally with humans" leads to the question of whether such robots should be granted a certain social, moral, or legal status. This would expand the circle of moral actors that we have so far restricted to humans. It is in this spirit that the European Parliament launched a draft law to regulate the action and responsibility dimension of so-called cyber physical systems, and has underlined the urgency of its initiative by saying:

> Humankind stands on the threshold of an era when ever more sophisticated robots, bots, androids and other manifestations of artificial intelligence ('AI') seem to be poised to unleash a new industrial revolution, which is likely to leave no stratum of society untouched. The development of robotics and artificial intelligence raises legal and ethical issues that require a prompt intervention at EU level. (EU 2015)

Beyond apocalyptic and euphoric attitudes towards digitalization (which here encompasses **Big Data**, algorithms, **machine learning**, robotics, **Artificial Intelligence** (AI)[1] and the **Internet of Things** (IoT)), we want to try to provide a sober presentation guided by normative orientation. Section 2 will turn to central aspects of a **philosophy of technology** in the field of tension between man, technology, and ethics, and will concern technology, because the world-mediating role of the digital is of particular importance for ethical and anthropological considerations. Characteristics of digitalization are developed in Sect. 3, and differentiated according to

[1] In this article, we will focus on weak AI, and not so much on strong AI or super intelligence. Weak AI is programmed to solve a specific problem. Through machine learning, it can optimize its calculations and actions (Brundage 2018). Strong AI acts independently in a complex context, taking many factors into account, including indiscrete factors, and is increasingly more powerful than humans. So far, there is no super-intelligence (Bostrom 2014; Grace et al. 2017).

individual, organizational, and social level. In Sect. 4, we will develop the main features of the ethics of digitalization. Starting with some basic explanations of morals and ethics and the various **spheres of ethics**, we will present three approaches to an ethics of digitalization: the condition of the possibility of moral questions, the importance of images of man for moral questions, and normative issues. Section 5 is devoted to the ethics of digitalization in industry and includes topics such as **platform economy** and "Work 4.0." In Sect. 6, we will summarize the results and give an outlook.

2 The Relationship Between Man, Technology, and Ethics

Technology is a cultural factor that has accompanied man throughout history, opening up a space of possibility as a habitat for a creature that is only moderately specialized in biological niches, and thus has a special influence, not only on any understanding of the world but also on man's self-image (Poser 2016, esp. p. 96 ff.). Here, technology does not only refer to certain **artifacts** for instructive use, but also to the processes to produce such **artifacts** and the form of reflection, with the help of which humans produce such **artifacts** and make themselves who they are. Palaeoanthropologically, the development of eye-hand coordination can be observed (Leroi-Gourhan 1980), which, in connection with language, led to increasingly complex mental performances, with which humans produced more and more complex technical worlds. The latter required new orientations which, as a way of life, always carried a normative question. This has, as Ricœur said, a three-digit relation: "Orientation towards the good life, with and for others, in just institutions." (Ricœur 2005, p. 252) All three components are equally important, and provide the coordinates of an ethics of digitalization, as we will develop them in this article. To consider them requires a form of reflection that transcends current social morals and the specific ethos of certain groups. Ethics aims at a tendency towards a universal norm about its validity and justification and must always be convincing in the respective contexts of action. This ethics is based on the view that man, as an individual (as well as a generic being, and not a machine), sets goals and thus makes decisions in relation to others and for the design of institutions.

Technology does not appear here as a "neutral thing," about whose "positive" use a society would only have to agree "after" its construction and "before" its use. It rather appears as designed by people, which as a project always tries to answer a problem, but also creates new problems due to its technical ambiguity. It is not least this ambiguity of technology and its undesirable side effects that contribute to widespread technology conflict (Renn 2013), which in turn lead to efforts for technology acceptance. These, often enough, only address secondary problems and neglect the main ones. The aim here should be to deal with a dilemma named by David Collingridge in an "ambivalence-proof," (i.e., informed, honest and responsible) manner: If a technique is not very well known or well-established, very little is known about its effects and unwanted side effects. Once a technique is widespread and established, it is difficult to control it or even take it back (Collingridge 1981, p. 19).

To the extent that the problems are located on the side of technology, talk of the "constraints" or the "arbitrariness" of technology becomes superficially plausible. An informed **ethics** or **philosophy of technology**, on the other hand, can critically and productively deconstruct the threatening and comfortable attitude of man towards its origins. The mediation of goals or ends and means is given, due to the specific **mediality** of technology (Hubig 2006; Fischer 2004, p. 103 ff.), and as such, reflexively is given to man.

The **mediality** of technology means that technical apparatuses and infrastructures can be seen as a "means to an end" (the hammer as strengthening and focusing the power to prepare objects, or the power station for energy supply), but that in this dyad they suppress the medial and reflective moment of their use of means: technology is a means to something for those who use it. Others can use the hammer or energy in a different way, and still others can turn a urinal into a work of art—as Marcel Duchamp did with his, in 1917.

> Therefore, the concept of technology is not a collective term for individual techniques, but represents a concept of reflection. Reflection can take place in different ways: as a reference determination by differentiating between technology and non-technology, as a functional interpretation by indicating (e.g., anthropological) functions of technology, by determining its place in contexts of action and cultures and by referring to reproducibility and regularity. (Grunwald 2013, p. 14)

The appropriation of the world through technology, its mediation between nature and society, extends beyond a simple instrumental understanding. It has its critical turning point where technology threatens to become "second nature." Technology as second nature means that man finds himself in a fundamentally technically formatted world, which then appears to be just as threatening and unmanageable as the first nature which he tried to ward off, if not to get rid of, using technology. **Ubiquitous computing** can be understood as such a comprehensive "technological texture" (Grunwald 2013, p. 16) – a point that will be reflected upon in more detail below.

The **philosophy of technology**—and with it, ethics—can tie technology, as the "art of the possible" (Hubig 2006, 2007, 2008), back to concrete forms of life and ethical deliberation, to negotiate how goals in a society are to be pursued with which means, and how responsibility is to be assumed. Here, it is important to consider the far-reaching implication of technology as a **mediality** that makes human life possible, as well as endangering, liberating, standardizing, and even hindering existence. Technology in general and digitalized technology in particular are not simply tools that humans use to be effective in this world. They do not only extend the repertoire of human action, but also communicate this world in a specific way (e.g., imaging procedures in medicine, augmented and virtual realities) and thus contribute to a dynamic understanding of the self and world: "Specific programs and logics are inscribed in technical artifacts, which make certain contents permanent and transportable, but also transform them, i.e., constantly modify their permanence." (Hubig 2015, p. 118)

Accordingly, a **philosophy of technology** must make visible the presence of technology in human life and consider its constitutive meaning, for all human actions, in its sense of possibility. To have possibilities means to epistemologically limit the variety of technically producible worlds and to give orientation by means of ethics, to what serves a good life and what harms it. This requires well-founded decisions which, from an ethical point of view, must consider the three relations: I, others, and institutions.

3 Characteristics of Digitization

The term digitization has its origin in the Latin digitus, the "finger" and the "toe," which also serves as a measure for length and for counting. Counting something on the fingers leads to calculating (computare). The computer is a calculator which carries out its calculations with the help of two digits: the numbers 0 and 1. For the computer to calculate the most diverse life-world processes, the necessary information must be converted to a data format of zeros and ones, i.e., digitized. The predecessors of the computer refer to interesting connections between time and number (Borst 1990). Its history proves to be by no means consistent and planned—which one would perhaps expect from this very machine, and the teaching of computer science on which it is based (Lévy 1995).

In line with Brennen and Kreis we understand "digitization as the material process of converting individual analogue streams of information into digital bits. In contrast, we refer to digitalization as the way in which many domains of social life are restructured around digital communication and media infrastructures." (Brennen and Kreiss 2014) Data is the common language for all digital devices and algorithms, as extensive sequences of commands for processing this data. They are produced (not least by social media) on an immense scale and form a constantly growing population of digitized processes and states. This complex is often described using the term "Big Data," although it is still not quite clear what the term actually means. It is characterized by up to seven "Vs": volume, velocity, variety, veracity, validity, value, and visibility (McNulty 2014). Sensors are an essential prerequisite for an ever-broader digitization of life-world phenomena, such as sounds, smells, vital parameters, mobility, consumer behavior, etc. It was estimated that by the year 2020, around 50 billion sensors will be installed and networked (Helbing et al. 2015). Digitization therefore means the networking of more and more digital end devices to form a large network of information and communication, in which with every tweet, every online order, every surfing process and any other digital articulation, further data is produced, collected, stored, processed, and used to control individual, organizational, and social processes (digitalization). On an individual level, the permanent data collection (monitoring) of a diabetic patient provides a much more accurate picture of their metabolism and the insulin requirement than sporadic measurements and average values allow (see Kucklick 2014, p. 7 ff.). In companies, digitalization leads to the optimization and automation of business processes via data management, e.g., completely digitized production processes, fully automated warehousing, and "the fundamental change in business models as a whole." (Hildebrandt and Landhäußer 2017, p. VI) Analogously, it is thought that more and more data will have to be collected and collated for the control of social processes to be able to control, on a large scale, in an ever more comprehensive and at the same time ever more refined manner (Zuboff 1988, 2019).

Digitization is thus based on an epistemological premise, which is that more data brings more insight. More knowledge then offers a better basis for the understanding and control of individual, organizational and social processes, i.e., for the shaping of the world and ultimately also for a good or successful life, as it is the subject of ethics. The equation "more data = more knowledge = better control" sounds plausible, and for many areas of practical life it also applies *cum grano salis*. However, this assumption comes up against some theoretical and practical problems (Pasquinelli 2017). From an epistemological perspective, one may argue whether empirical data can be the basis of knowledge or cognition at all. Even if one chooses an empirical approach, the question arises as to which data is necessary, and how much is sufficient for making valid decisions, which are the condition for the possibility of targeted action. This question is also of practical relevance, since in digitalized worlds, very heterogeneous data is in some cases brought together and makes the basis for decision-making.

Not all collected data has the same quality, and so errors can creep into the processing and control of processes, which can have far-reaching consequences, depending on the scale and type of processing. This reference is not intended as an argument against digitization, but as an indication of the corresponding duties of care. In addition, it must be considered that in self-learning algorithms (**machine learning** or **deep learning**) which process these data, it is less and less possible for humans to understand the operations and their results. Josef Weizenbaum therefore also speaks of technical systems that are "incomprehensible" for humans (Weizenbaum 2003, p. 58). He refers here to the theoretical considerations of Norbert Wiener, who pointed out this epistemologically and ethically fundamental problem in 1960:

> It is my thesis that machines can and do transcend some of the limitations of their designers, and that in doing so they may be both effective and dangerous. It may well be that in principle we cannot make any machine the elements of whose behavior we cannot comprehend sooner or later. This does not mean in any way that we shall be able to comprehend these elements in substantially less time than the time required for operation of the machine, or even within any given number of years or generations [. . .]. An intelligent understanding of their mode of performance may be delayed until long after the task which they have been set has been completed. This means that though machines are theoretically subject to human criticism, such criticism may be ineffective until long after it is relevant. (Wiener 1960, S. 1355)

4 The Main Features of the Ethics of Digitization

4.1 On the Distinction Between Ethics and Morals

Ethics and morals mean the same thing. In everyday language, "ethical-moral" is used synonymously. In scientific debates a distinction is made in part, but by no means consistently. Some call ethics what others call moral philosophy, others understand morality to be a systematized construct of ideas, and ethics to be the everyday actions of people. In the following, we will make this distinction: morality is a system of rules and norms for action orientation. It denotes the totality of accepted norms of conduct of a society, which have been stabilized by tradition, norms or goods which are actually valid in a group of people, or in an organization. Ethics is opposed to the reflection theory of morality. As such, it contributes to the objectification of moral (or moralizing) debates—by examining the genesis and validity of their intuitions and arguments in the "orientation towards the good life, with and for others, in just institutions." (Ricœur 2005, p. 252)

The question "What should I do?" is fundamental to people's everyday actions. It results from the fact that human being (as a generic being, as well as an individual) has to design life as a self and world relationship and, in doing so, has to design technically, culturally, and less than other living creatures. Such designs—such as the supply of food to one's own group, or the establishment of a high-tech strategy for Germany as a business location—are almost always associated with moral questions. Moral questions are those in which our humanity, as an individual and as a generic being, is at stake. Three elements come together in the moral question: first the person (group, organization, or society) who experiences this question as addressed to them, second the situation as an excerpt of the world which concerns this person, organization, or society in a moral way, and third the moral demand (Løgstrup 1989) which this person, organization, or society experiences in this situation. Only in the interplay of these three elements can one speak of a moral question, in the fullest sense. Moral is a demand, because it demands in a very fundamental way a response to the other, who demands recognition and justice from "me" in what "I" do, and which affects them at the same time. If the development of **Artificial Intelligence** and robotics requires a "prompt [...] intervention" in the sense of the above legislative initiative from the European Parliament, the moral demand is obviously very strongly experienced here, and the answer to it does not only concern the present generation here in Europe, but also coming generations all over the world. It is therefore a moral question of enormous scope, which—similarly to climate change—should not be ignored but answered very urgently and carefully.

Moral questions are not formulated first by professional ethicists and approached from the outside. Rather, they are already contained in the respective situations and are sometimes experienced more, sometimes less explicitly, or articulated in terms such as self-determination, justice, human dignity, or freedom. The task of ethics is to critically reflect the moral attitudes or moral judgements that exist in such situations to bring the genesis and validity of the actions based on them to political negotiation and—ideally—to promote decisions based on justice and responsibility.

4.2 Digitalization in Spheres of Ethics

In line with the differentiation of our society and its scientific, cultural, and technical spheres, the ethical reflection of moral questions has also become more differentiated and produced a broad register of specialized ("applied") ethics (Stoecker et al. 2011; Nida-Rümelin 2005). These are characterized by a mediation between the specific subjects and the fundamental moral questions. So-called "**spheres of ethics**" have established themselves, which deal with questions of moral orientation in their own subject areas, fields of action or professional fields. Despite all the differentiation in "hyphen ethics" that already exists and of which more is to be expected, along with the potential encapsulation in internal discourses associated with it, their common concern is developing as prescriptive ethics "as far as possible principles of high general validity in the form of supernorms and general rules [. . .] but then above all a specialized panorama of norms and rules for exemplary subject areas." (Krämer 1992, p. 373)

In this, there is an unmistakable further development of the ethical instrument. Accordingly, for the field of digital technologies, separate discourses with specific questions and concepts have developed. At first, **computer ethics** established itself as its own part of **ethics of technology** (Moor 1985; Johnson 2001). Later, the fields of **machine ethics** (Anderson and Anderson 2011), **robot ethics** (Lin et al. 2017), **information ethics** (Floridi 2013, 2015) and an even more precisely developed **algorithm ethics** (Jaume-Palasí and Spielkam 2017) were developed.

It would be obvious to regard digitization as a topic for a new **sphere of ethics** to be developed, a kind of "ethics of digitalization" or "digital ethics" (Otto and Gräf 2018). It could also be conceived as a specialized form of **ethics of technology** because it involves special technical **artifacts**. If, on the other hand, one considers data as the starting point and central element of development and realizes that they are the "capital of the 21st century" (Mayer-Schönberger and Ramge 2017) or the new means of payment, the new currency in exchange for goods and services, then one could just as well conceive of digitalization as a topic of **business ethics**. This would be supported by the fact that new markets, innovations, and business models such as the **platform economy** are at the heart of digitalization. If one goes a step further and recalls the considerations of Langdon Winner (1980, p. 128 f.), that technical infrastructures have the character of legislative acts, then **political ethics** would also come into question as the custodian of the topic.

This sectoral ethical approach encounters certain difficulties. We therefore advocate a multi-dimensional approach that integrates different area-specific approaches—just as the process of digitalization not only links the most diverse end devices and actors with one another, but in this way also undermines the earlier separation of social spheres (e.g., science, business, law) and the categorical distinction between private and public. Accordingly, we pursue an integrative approach here, which can integrate the special knowledge of **spheres of ethics** into a broader approach to the problem. By "problematization," we mean the mode of access to social transformations and the associated methods of processing, as outlined by Foucault (1997), and elaborated further by Rabinow (2003) in the fields of anthropology and ethics.

Both technical progress and ethical-anthropological reflection within this development give rise to uncertainty and call for the clearest possible orientation. Ethics as an orienting science must not rush to meet these demands, but it can reveal the conditions of the possibility for a discourse oriented on moral questions, can stimulate corresponding debates and deliberations, and accompany them with its own expertise, and can ethically empower the participants, while also reminding them of their responsibility. For this purpose, sorting these questions is useful. In the context of the debate on digitalization, robots and **Artificial Intelligence**, the question of whether these developments are actually associated with "new" moral questions is repeatedly raised, or whether they are not merely a continuation of what we already know from previous debates on the use of technology in production and its potential for innovation and new jobs—which have already been dealt with in a plausible way. Such questions are categorically different from questions about whether **Artificial Intelligence** can and should actually make decisions, or whether virtual reality and power support by exoskeletons are a consequent business "enhancement" of the worker to realize the necessary efficiency gains in the economic competition. Therefore, the different categories of moral questions must be distinguished first, and then approaches for their treatment must be presented.

4.3 Access to the Ethics of Digitization

If we consider the breadth of the ethical problems associated with digitalization and the sometimes enormous depth of detail of concrete ethical questions, a selection is inevitable here. As shown in Sect. 2, the ethical approach we advocate combines elements of an Aristotelian ethics of good living with deontological moments of responsibility (Bayertz 1995), which must be defined more precisely at the respective levels of action and in relation to the subjects of responsibility.

In the following, we will develop three categories of ethical problems that should help to summarize, as precisely as possible, the many questions that arise in relation to the digitalization of industry, thus making it possible to deal with them.

- What is the condition of the possibility to address moral questions in the field of digitization?
- Which implicit (or explicit) image of man is associated with the use of a certain technology, and how does this affect moral questions?
- Which norms, goods or values are aimed at a direct orientation of action for the use of a certain technology?

4.3.1 On the Condition of the Possibility of Moral Questions

In their technically or culturally mediated experience of the world, this world is only incompletely accessible to humans, and always in the form of symbolic representations. This means that the way in which humans perceive "their" world is not direct but always mediated, whether by technical means that guide or even improve this perception, or by the ideas (concepts) of what and how this world should be (Cassirer 2006). Technology allows us to see many things in this world more precisely or at all (microscope, telescope, etc.), but it does not make this world more direct, or close the gaps in the self-presentation of humans in the world but reproduces this "blind spot." It guides our experience.

A technology that adapts to our respective usage habits and, if necessary, also compensates human performance fluctuations is perceived too little or not at all—a characteristic that we usually expect from good technology. On the other hand, the resistance of the world is lost through such adaptive assistance technology. Technical assistance may guide human experience, with the result that people in this assistive environment perceive less and less what determines them. However, this technical subversion could be a danger for the self-determination of humans (Wiegerling 2012), which distinguishes them as moral subjects and finds its fundamental expression in the predicate of dignity (Bieri 2013).

Since, in the words of Immanuel Kant, "all our cognition begins with experience" (Kant 2005, p. 136), the loss of concrete experience of use through optimized technology proves to be problematic on its inconspicuousness and self-evidence. The loss of concrete experience may result in a loss of reflection and judgement. This affects both our ability to recognize and cope with the challenges posed by technology and our ability to shape the conditions and practices of use. It is the critical judgement of humans which is required in the respective **socio-technical arrangements**, and which must not be undermined by them. By **socio-technical arrangements**, we mean the interrelationship between man and machine in a concrete situation. It is therefore not a matter of questioning individual technical devices or systems of their potential or limits, but rather of always bearing in mind that only from a concrete interaction and benefit situation—which can also be subversive to the imagination of the designer or supplier—can morally relevant questions arise. This applies to the respective action, its consequences, and to the human being's self and world relationship, which changes through interaction.

Digitalization means that decisions are increasingly being outsourced to smart machines, and that the human being is relieved of this task. Beyond the question of what this means for our earlier concepts of decision and responsibility, the first thing to be identified is technical. These are machines that prepare decisions based on fed-in data and certain calculation rules (algorithms), or even carry them out completely. This applies to topics such as medical treatment, job interviews, the granting of a loan, the purchase of securities, the assessment of the recidivism rate of criminals, and the elimination of suspected terrorists via drones. The relief of the people has its good points where, for example in medical studies, a larger amount of data can be processed in a shorter time, and the decision can be put on a more secure basis. At the same time, and this once again illustrates the ambivalence of technology, this method of making decisions algorithmically will push people increasingly into the background. They will be less and less able to oppose mechanically calculated recommendations and decisions. **Artificial Intelligence**, which calculates with incoming data and a constantly growing database on the net, will be able to "beat" them in more and more fields, as it is already the case in chess and poker. In the medium term, this could be a self-fulfilling prophecy. The smart machines (or **Artificial Intelligences**) calculate certain probabilities for a given situation. These probabilities become the basis for a decision, which then enters the pool of data used for further calculations as a new date. At this point, the critical question from an ethics reflecting the conditions of decisions is first: how must the technical systems be designed so that they do not generate self-fulfilling prophecies in their feedback dynamics? Second, it must be asked how abuse based on this dynamic can be (largely) ruled out. Protocols about the surfing behavior of persons or groups ("click numbers") usually have an influence on the advertising revenue of the website provider. Therefore, higher "traffic" is generated automatically (e.g., by bots), which is also used to finance pages with questionable content. If one delves deeper here aspects arise, for example, from discrimination phenomena in the algorithms, biases, questions of the justification of mechanical decision-makers, social control or the legality of decisions and their legal challengeability (Jaume-Palasí and Spielkam 2017).

If, at a certain point, technical systems that are used for decision support or for decision making are no longer comprehensible to humans, this represents a problem from an ethical point of view, in that a critical evaluation of the machine components of decisions for the system-bound decision and the resulting action by humans is no longer possible. They no longer have an overview of the decision and can therefore no longer be regarded in the full sense as the responsible actors. Is a human being in such cases therefore no longer a moral actor, but has delegated this status to a machine? In certain situations, it may be quite useful (and even functionally better) to delegate elements of decision making or the entire decision to machines. Automated examination and decision making could also be sufficient for the examination of certain legal or insurance contracts. The decisive factor is probably that people in critical situations are still able to master the necessary "maneuvers" (such as the pilot in a plane or the driver in an autonomous motor vehicle), or that people are available as an examining authority and contact person for mere mechanical decisions. These two requirements, first, to keep the corresponding human competences still available, even if in most cases machines execute the tasks, and, secondly, to have humans as the "last chance" for checking decisions, are therefore decisive preconditions for **socio-technical arrangements** to be able to deal with moral questions at all. It will therefore be important to keep open the possibility that people set goals, choose means, make decisions, and take responsibility: this seems indispensable from an ethical perspective. Therefore, it is necessary to examine exactly if and when human makes decisions on machines, in which form, and can they still check them, despite any difficulties. Christoph Hubig has proposed forms of parallel communication (2008).

In a weaker form, the automated production of decisions is linked to the question of how and to what extent technical systems should relieve people, provided that the demands of stress are an impetus for learning, critical self-reflection and maturing. The compensation or substitution of human work by technology is accompanied by a relief of possibly heavy, dangerous, or even stupid activities. Assistance in more and more areas of life (Biniok and Lettkemann 2017) inevitably leads to the loss of once-learnt competences, which we do not know today whether we will consider unnecessary or whether we will mourn as irretrievable losses. An ethics of digitalization will consider these aspects, from the construction of digital devices and infrastructures to their distribution and disposal and inform the relevant actors.

4.3.2 The Importance of Images of Man for Moral Questions

Some moral questions relate to an implicit (or explicit) conception of human associated with the use of a particular technology. Constructions of technology are always based on a conception of human: why they produce, use, or should use this technology. This phenomenon is not new either: in the 16th century, mechanistic ideas became part of the self-description of man (Meyer-Drawe 1996, esp. p. 55 ff.), and the transformation of a bodily self-reference into an analytically distancing understanding of the body as a machine took its course (Kamper 2001; Harrasser 2013). In the wake of these developments, the **data licensing** of bodies and their data-based optimization in sports, at work, in nutrition or health can also be understood. Here, too, it is not a matter of rejecting or approving such developments per se, but rather of discussing, in the sense of Boehme's "serious questions," which technical developments we design in such a way that they allow us individually and socially to unfold our own humanity and not, technically framed and tracked, to subject it to foreign expectations and design fantasies. Such expectations or fantasies become concrete, e.g., in certain images of age (Aner et al. 2007) or more generally in images of a performance society and its constant, technical optimization, even surpassing the human condition (Spreen et al. 2018).

In this area, too, it is not a matter of establishing a certain binding image of human, but rather of bringing the implications of the respective images to ethical reflection and—despite all the change that is also historically recognizable in the way humans see themselves—to consider questions of the human condition as fundamental to ethics. It could be pointedly said that in every speech about human there has always been a "thinking hidden to human himself" (Hogrebe 2013), and that its effects unfold. It is the task of ethics to make this implicitness in all its provisional and incomplete nature conscious, and thus reflect on its importance for ethical analysis and upcoming decisions.

4.3.3 Normative Issues

Ethics that deal with digitization in the industry must not only deal with the problems and risks associated with it, but must also consider the potential and opportunities, although this must not amount to a simple offsetting of "losses" and "profits." Problems remain problems, even if they can be balanced out by positives in the total sum.

As far as the field of normative questions is concerned, we must now move from more general theoretical considerations to concrete **socio-technical arrangements**. Here, we will deal with questions in the concrete interplay of humans and machines, which goods or standards are in conflict, and how such a conflict (e.g., between freedom and security), which cannot be decided a priori, can be ethically evaluated and decided.

5 Ethics of Digitalization in Industry

Ethics, as we develop them here, are to be understood as an integrative approach which combines the special knowledge of various **spheres of ethics** in their reference to phenomena of digitalization and thus, on human, to thematize selfhood and living together with a view to a good life, in just institutions. All three relations are affected by the upheavals of digitalization. The speech about the second industrial revolution quoted at the beginning of this article has strong effects on the form of work and its meaning in human life in general, the implications for social life, and for the distribution of value creation and thus the conditions of a just living together.

This can be experienced daily through the **platform economies** that penetrate more and more areas of life (Shapiro and Varian 1998; Parker et al. 2017; BMWi 2017). Deliveries of everyday and extraordinary needs are conveyed, as well as mobility (e.g., Uber), communication (social media such as Facebook and Instagram) and moments of social coexistence (via neighborhood networks or partnership agencies). The same applies to health and social services, and logistics and goods production. An important aspect of the **platform economy** are the so-called network effects: the more people use the platform, the more useful it becomes for the users, and the more profitable and efficient it becomes for the platform operator.

At present, it can be seen that certain platforms have achieved quasi-monopoly status. From economic theory, it is known that monopolies are counterproductive for the market and competition and should be avoided for reasons of political blackmailing. Ethically, the topic is no less problematic because it raises questions of justice and power. Another problem the **platform economy** and its networks have is that negative effects have a much greater extent and speed than in earlier interactive structures. On top of this, **platform economies** produce so-called rebound effects (Eichhorst and Spermann 2016): By buying goods online, individual car traffic during shopping should be reduced, and thus also emissions. It has been shown, however, that general transport traffic to individual customers has instead increased, and more than overcompensated for any reductions.

On the data side, it should be noted that such platforms gain their functionality by "training" the self-learning algorithms with corresponding data sets. The more users, the more data, the more accurate the algorithm becomes, and the more powerful and convenient the service. This self-reinforcing network effect uses the semi-voluntary and usually free provision of data by the users. This leads to several problems. The voluntary nature of data provision is rather precarious. Conditions would have to be created that would put the user as a "data donor" in a negotiating position with the service providers and data recipients that could be called fair. Corresponding approaches are called "ART of AI": the Accountability, Responsibility, and Transparency of Artificial Intelligence (IEEE 2018). However, these are developed for the US American market, and do not exactly match the European situation.

Thus, asymmetrical competitive situations result from the different handling of freedom rights and the possibilities for disposal of one's own data on the market of digital products. For example, to train automated vehicles for their use, they require large quantities of images as training data. In Europe, however, images of people taken automatically and en masse, which make it possible to recognize and pay attention to people in traffic, require the consent of the person in the image (GDPR 2018), which is not so easy to obtain when there are tens of thousands of cases. In other countries, where this right does not exist or is not consistently respected, it is easier and faster to generate training data for **Artificial Intelligence** in cars. Companies in these countries thus have a clear competitive advantage over companies in Europe that comply with the **GDPR** (and other legislation).

At present, there is a need in Germany and Europe—and here, governmental aid is called for—to provide companies and other organizations with appropriate training data, which is both data protection compliant and sufficiently informative, to ensure that competition with the large companies of the **platform economy** can be accepted and appear realistic. In the case of **platform economies**, a regulatory basis in a global perspective must first be created to be able to ask and negotiate certain moral questions.

If one looks at the different types of data that are collected, processed, and aggregated to metadata in the course of digitization, they must be clearly distinguished with regard to their ethical sensitivity. An item date in the warehouse for a component in production has a different sensitivity than the cardiac monitoring of a patient: both differ once again from the protocol data used to send communication content. If one looks at data from an ethical perspective, one must always ask which data should be protected and for what purpose. It is obvious that personal data should be protected more carefully than other data, and here health-related data should once again be given greater importance than communication or mobility data. The **General Data Protection Regulation** (GDPR 2018) has resulted in much stricter regulations, which force companies to handle personal data according to relatively clear rules. However, compliance with this body of regulations cannot dispense companies, states, and individuals from discussing the moral implications of data collection, processing, storage, and use. This means that we must think about the deletion of data and offer transparent, responsible, and reliable procedures to the data provider. On the other hand, these forms of data and the way they are processed are often a form of business capital, which is why the required transparency can contradict operational confidentiality. Here, companies and state regulatory bodies are required to ensure a differentiated handling of personal protective goods (such as privacy, physical integrity, and property) differentiated according to data types and legal interests, in consideration of entrepreneurial freedom, innovation, intellectual property, etc. (Eifert and Hoffmann 2009; Hoffmann et al. 2015).

Looking at **automation** as a central moment of digitalization in industry, some aspects are particularly worthy of consideration from an ethical point of view. For example, the **automation** of work has always been associated with a displacement of human work by technology. This in itself is nothing new, but the enormous speed with which jobs are being destroyed here—and often easily replaced by others—demands completely different reaction times from companies, politicians and the individual than was previously the case with transformations in the world of work. If one realizes that the transformation from coal-mining districts to other living and working environments can easily take 40 or 50 years, then it quickly becomes clear that when workers are released from the transport industry (because of automated driving), the insurance industry (because of the algorithmization of the checking and decision-making processes) or from call centers (because of the use of a simple AI), many more people will be affected over a much shorter period of time, and the social efforts will have to be correspondingly greater.

One of the key questions is whether and to what extent digitalization will destroy jobs, and to what extent new jobs will be created and which qualifications will be required. The figures on jobs vary widely (Frey and Osborne 2013; Dauth et al. 2017). It is obvious that a purely utilitarian assessment—there are more winners than losers—does not go far enough here, as it can create or worsen dramatic imbalances (digital and social divides). Frey and Osborne argue that the digitalization of the world of work will attract highly qualified employees, while simple jobs will be eliminated by digitalization. In addition, it will have to be seen what effect the removal of temporal and spatial boundaries from work using digital information and communication technologies has on the private lives of employees (Busch-Heizmann et al. 2018).

The redistribution of work raises questions of justice, which are expressed in terms of who can still take over which job after digitalization. The rule of thumb is that robots equipped with AI take over the work for which they can be technically designed, and which is worth the investment. There will be work for which it is not worth building a robot—such work will still have to be done by humans. But this work will neither be very attractive nor well paid ("down-skill" effect). On the other hand, there will be work that should not be done by a robot because of its complexity, its specific forms of interaction, or perhaps because of its representative status. This will then presumably be paid much better ("up-skill" effect). These developments require not only corresponding socio-political cushioning, but also further considerations as to how humans can keep their dignity under these conditions. The historical and ethical lessons learned from the observation of past transformations (such as the first industrial revolution in the 19th century, or the mass unemployment of the 1920/30s) should not be carelessly gambled away here. The design of Work 4.0 is not only a challenge for individuals or companies, but also a social challenge (Pelikan and Rehm 2018; BMAS 2017; Botthof and Hartmann 2015).

Our understanding of digitalization refers to the discounting of future positive and negative effects. In the sustainability discussion, which also considers upstream and downstream production and thus complete value chains and even integrates them into entire economic cycles, the term "footprint" was first brought into discussion. This concerned the prevention of so-called negative external effects (e.g., the prevention of CO2 emissions, or water pollution). Later, the positive effects were added under the term "handprint" (UNESCO 2018). For example, the positive effects of a value chain could be understood as the promotion of a social project through a loan or the improvement of sustainable mobility through a self-propelled car. The ambivalence of the digitalization process can thus be illustrated using a handprint-footprint logic. Ethically speaking, the debate should be expanded and accentuated by insights from the sustainability discourse (Vogt 2009).

The fact that ethics considers the good life of all, and that the digitalization of industry is not limited to the internal production processes within the company but includes upstream and downstream value chains, means that the monistic objective function (as is known in economics, with its profit orientation) dissolves in favor of a pluralistic objective function (moral, ecological, public welfare balance). The evaluation as "good" or "successful" can thus no longer be solved only economically, mathematically, or algorithmically, but requires a moral, social, ecological evaluation, and flexible determination of the current situation.

> Whether our ethical practices are Western (Aristotelian, Kantian), Eastern (Shinto, Confucianise), African (Ubuntu), or from a different tradition, by creating autonomous and intelligent systems that explicitly honor inalienable human rights and the beneficial values of their users, we can prioritize the increase of human well-being [sic!] as our metric for progress in the algorithmic age. Measuring and honoring the potential of holistic economic prosperity should become more important than pursuing one-dimensional goals like productivity increase or GDP growth. (IEEE 2018, p. 2)

6 Conclusion

This article was an introductory presentation of the form and content of normative (especially moral) questions and their ethical treatment. The field of digitalization shows that a classical area-specific access, as is practiced in other fields of applied ethics, does not apply here. Instead, multidisciplinary access (such as **business, technical and media ethics**) is required across several levels of the scope and responsibility of the various actors. Based on an expanded understanding of technology (**artifacts**, procedures, form of reflection), the specific **mediality** of technology was thematized, which can open up this world and the actions in it, but can also determine the world and its actions, and thus close it. What is required of us humans is therefore an ethical reflection that does not try to resolve the tension between potential and danger one-sidedly, but to shape it responsibly. This applies to digitalization to a particularly high degree because it affects all areas of social life at great speed and causes rehearsed differences (such as those between public and private, original and copy, production and consumption) to oscillate and demands new answers from us. New answers can (and should) connect to old orders and practices, but they must not become an obstacle to really understanding the new and accepting it as a challenge to be shaped. Thus, the question of how to make people's lives successful in individual, social, and institutional terms is so central. To discover this question in the various facets of digitalization, to perceive and appreciate it in its differentiated **socio-technical arrangements**, requires (in addition to all active shaping) a moment of reflection. In this sense, ethics is also a form of interruption of our industrial and productive routines in order to ask once again—and probably again and again—what kind of people we want to be, and what form of living together we will take as a basis for humanity. The choice that has been decided upon in this question is not an arbitrary one. Primarily, it is based on a freedom that is given to us and thus also abandoned. Our task is to give shape to the human condition as a form of life committed to freedom. Digitalization in its various processes and with its dynamic structures is a medium of freedom as well as a threat to it. It is up to us to make digitalization not only an entrepreneurial but also a civic task and to give it a humane face—especially in the robots that are now appearing.

References

Anderson M, Anderson SL (eds) (2011) Machine ethics. Cambridge University Press, Cambridge

Aner K, Karl F, Rosenmayr L (eds) (2007) Die neuen Alten – Retter des Sozialen? Springer, Wiesbaden

Bayertz K (1995) Verantwortung. Prinzip oder Problem? Wissenschaftliche Buchgesellschaft, Darmstadt

Bieri P (2013) Eine Art zu leben. Über die Vielfalt menschlicher Würde. Carl Hanser, München

Biniok P, Lettkemann E (2017) Assistive Gesellschaft – Multidisziplinäre Erkundungen zur Sozialform "Assistenz". Springer, Berlin

BMAS (2017) Weissbuch – Arbeiten 4.0, Reihe: Arbeit weiterdenken, bud Potsdam

BMWi (2017) Weissbuch Digitale Plattformen. Digitale Ordnungspolitik für Wachstum. Innovation, Wettbewerb und Teilhabe, Spree Druck, Berlin

Böhme G (2001) Ethics in context. The Art of Dealing with Serious Questions. Polity Press, Cambridge in association with Blackwell Publishers Ltd., Oxford

Borst A (1990) Computus. Zeit und Zahl in der Geschichte Europas. Wagenbach, Berlin

Bostrom N (2014) Superintelligence. Paths, dangers, strategies. Oxford University Press, Oxford

Botthof A, Hartmann EA (eds) (2015) Zukunft der Arbeit in Industrie 4.0. Springer Vieweg, Berlin

Brennen S, Kreiss D (2014) Culture digitally. September 8. http://culturedigitally.org/2014/09/digitalization-and-digitization/. Accessed 1 Jan 2022

Brundage M (2018) The malicious use of artificial intelligence: forecasting, prevention, and mitigation. https://www.repository.cam.ac.uk/bitstream/handle/1810/275332/1802.07228.pdf?se-quence=1&isAllowed=y. Accessed 1 Jan 2022

Busch-Heizmann A, Entgelmeier I, Rinke T (2018) Digitalisierung und Entgrenzung. Welche personenbezogenen Merkmale beeinflussen die Gestaltung von Berufs- und Privatleben unter dem erwerbsbezogenen Einsatz in IuK-Technologien und wie lassen sich die Zusammenhänge überprüfen? Hans-Böckler-Stiftung, Düsseldorf

Cassirer E (2006) An essay on man. An introduction to a philosophy of human culture. Gesammelte Werke Hamburger Ausgabe, Vol. 23. Meiner, Hamburg

Christensen CM (1997) The innovators dilemma: when new technologies cause great firms to fail. Harvard Business School Publishing, Boston

Collingridge D (1981) The social control of technology. The Open University Press, Milton Keynes

Dauth W, Findeisen S, Südekum J, Wößner N (2017) German robots – the impact of industrial robots on workers, IAB-discussion paper 3

Eichhorst W, Spermann A (2016) Sharing Economy: Mehr Chancen als Risiken. Wirtschaftsdienst 96(6):433–439

Eifert M, Hoffmann W (2009) Innovationsverantwortung. Duncker & Humblot, Berlin

EU (2015) European Parliament on rules for robotics, 2015/2103(INL) – 16/02/2017 Text adopted by Parliament, single reading. http://www.europarl.europa.eu/oeil/popups/summary.do?id=1477231&t=d&l=en. Accessed 1 Jan 2022

Fischer P (2004) Philosophie der Technik. Eine Einführung. Wilhelm Fink, München

Floridi L (2013) The ethics of information. University Press, Oxford

Floridi L (2015) The fourth revolution. How the Infosphere is Reshaping Human Reality. University Press, Oxford

Foucault M (1997) Polemics, politics, and problematizations. In: Rabinow P (ed) Ethics: subjectivity and the truth. The essential work of Foucault 1954–1984, Bd 1. The New Press, New York, pp 111–119

Frey CB, Osborne MA (2013) The future of employment: how susceptible are jobs to computerisation? https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf. Accessed 1 Jan 2022

GDPR (2018). https://gdpr-info.eu/. Accessed 1 Jan 2022

Grace K, Salvatier J, Dafoe A, Zhang B, Evans O (2017) When will AI exceed human performance? Evidence from AI experts, arXiv preprint server. https://arxiv.org/abs/1705.08807. Accessed 1 Jan 2022

Grunwald A (Eds) (2013) Handbuch Technikethik. J. B. Metzler, Stuttgart/Weimar

Harrasser K (2013) Körper 2.0. Über die technische Erweiterbarkeit des Menschen. transcript, Bielefeld

Helbing D, Frey BS, Gigerenzer G, Hafen E, Hagner M, Hofstetter Y, van den Hoven J, Zicari RV, Zwitter A (2015) Digitale Demokratie statt Datendiktatur. www.spektrum.de/t/das-digital-manifest. Accessed 1 Jan 2022

Hildebrandt A, Landhäußer W (Eds) (2017) CSR und Digitalisierung. Der digitale Wandel als Chance und Herausforderung für Wirtschaft und Gesellschaft. Springer Gabler, Berlin

Hoffmann A, Schulz T, Zirfas J, Hoffmann H, Roßnagel A, Leimeister JM (2015) Legal compatibility as a characteristic of sociotechnical systems – goals and standardized requirements. In: Business & Information Systems Engineering (BISE), pp 1–11. https://doi.org/10.1007/s12599-015-0373-5. Accessed 1 Jan 2022

Hogrebe W (2013) Der implizite Mensch. Akademie, Berlin

Hubig C (2006) Die Kunst des Möglichen, Vol. 1: Technikphilosophie als Reflexion der Medialität. transcript, Bielefeld

Hubig C (2007) Die Kunst des Möglichen, Vol. 2: Ethik der Technik als provisorische Moral. transcript, Bielefeld

Hubig C (2008) Mensch-Maschine-Interaktion in hybriden Systemen. In: Hubig C, Koslowski P (eds) Maschinen, die unsere Brüder werden. Mensch-Maschine-Interaktion in hybriden Systemen. Wilhelm Fink, München, pp 9–17

Hubig C (2015) Die Kunst des Möglichen. Vol. 3: Macht der Technik. transcript, Bielefeld

IEEE | Institute of Electrical and Electronics Engineers (2018) Ethically aligned design. A vision for prioritizing human well-being with autonomous and intelligent systems. https://standards.ieee.org/industry-connections/ec/ead-v1.html. Accessed 1 Jan 2022
Jaume-Palasí L, Spielkam M (2017) Ethik und algorithmische Prozesse zur Entscheidungsfindung oder -vorbereitung. In: AlgorithmWatch Arbeitspapier Nr. 4/2017. https://algorithmwatch.org/wp-content/uploads/2017/06/AlgorithmWatch_Arbeitspapier_4_Ethik_und_Algorithmen.pdf. Accessed 1 Jan 2022
Johnson DG (2001) Computer ethics, 3rd edn. Prentice Hall, Upper Saddle River
Kamper D (2001) Körper. In: Barck K, Fontius M, Schlenstedt D, Steinwachs B, Wolfzettel F (eds) Ästhetische Grundbegriffe, vol 3. J. B. Metzler, Stuttgart, pp 426–450
Kant I (2005) Critique of pure reason. The Cambridge Edition of the Works of Immanuel Kant. Cambridge University Press, Cambridge
Krämer H (1992) Integrative Ethik. Suhrkamp, Frankfurt
Kucklick C (2014) Die granulare Gesellschaft. Wie das Digitale unsere Wirklichkeit auflöst, Ullstein, Berlin
Kurzweil R (1990) The age of intelligent machines. MIT-Press, Boston
Leroi-Gourhan A (1980) Hand und Wort. Die Entwicklung von Technik, Sprache und Kunst. Suhrkamp, Frankfurt
Lévy P (1995) Die Erfindung des Computers. In: Serres M (ed) Elemente einer Geschichte der Wissenschaften. Suhrkamp, Frankfurt, pp 905–944
Lin P, Jenkins R, Abney K (eds) (2017) Robot ethics 2.0. From autonomous cars to artificial intelligence. University Press, Oxford
Løgstrup K (1989) Die ethische Forderung. Mohr Siebeck, Tübingen
Mayer-Schönberger V, Ramge T (2017) Das Digital. Markt, Wertschöpfung und Gerechtigkeit im Datenkapitalismus. Econ, Berlin
McNulty E (2014) Understanding Big Data: the Seven Vs. http://dataconomy.com/2014/05/seven-vs-big-data/. Accessed 1 Jan 2022
Meyer-Drawe K (1996) Menschen im Spiegel ihrer Maschinen. Wilhelm Fink, München
Moor JH (1985) What is computer ethics? Metaphilosophy 16(4):266–275
Nida-Rümelin J (2005) Angewandte Ethik: Die Bereichsethik und ihre theoretische Fundierung. Ein Handbuch. Kröner, Stuttgart
Otto P, Gräf E (eds) (2018) 3TH1CS. Die Ethik der digitalen Zeit. Bundeszentrale für politische Bildung, Bonn
Parker GP, van Alstyne MW, Choudary SP (eds) (2017) Platform revolution: how networked markets are transforming and how to make them work for you. Norton & Company, New York
Pasquinelli M (2017) Arcana Mathematica Imperii: the evolution of western computational norms. In: Hlavajova M et al (eds) Former West. Art and the Contemporary after 1989. MIT Press, Cambridge
Pelikan R, Rehm J (eds) (2018) Arbeit im Alltag 4.0 – Wie Digitalisierung ethisch zu lernen ist. Beiträge aus betrieblicher, arbeitsmarktpolitischer und theologischer Sicht. LIT, Münster
Poser H (2016) Homo Creator. Technik als philosophische Herausforderung. Springer VS, Wiesbaden
Rabinow P (2003) Anthropos today. Reflections on modern equipment. Princeton University Press, Princeton/Oxford
Renn O (2013) Technikkonflikte. In: Grunwald A (ed) Handbuch Technikethik. J. B. Metzler, Stuttgart/Weimar, pp 72–76
Ricœur P (2005) Ethik und Moral (1990). In: Ricoeur P (Eds) Vom Text zur Person. Hermeneutische Aufsätze (1979–1999). Meiner, Hamburg
Schwab K (2015) The Fourth Industrial Revolution. Foreign Affairs, Dec. 12, https://www.foreignaffairs.com/articles/2015-12-12/fourth-industrial-revolution. Accessed 1 Jan 2022
Shapiro C, Varian H (1998) Information rules: a strategic guide to the network economy. Harvard University Press, Boston

Spreen D, Flessner B, Hurka HM, Rüster J (2018) Kritik des Transhumanismus. Über eine Ideologie der Optimierungsgesellschaft. transcript, Bielefeld

Stoecker R, Neuhäuser C, Raters M-L (2011) Handbuch Angewandte Ethik. J. B. Metzler, Stuttgart/Weimar

UNESCO (2018) Centre for Environment Education. https://en.unesco.org/greencitizens/stories/centre-environment-education. Accessed 1 Jan 2022

Vogt M (2009) Prinzip Nachhaltigkeit. Ein Entwurf aus theologisch-ethischer Perspektive. oekom, München

von Mutius B (2017) Disruptive Thinking. Das Denken, das der Zukunft gewachsen ist. Gabal, Offenbach

Weizenbaum J (2003) "... der Mensch, nicht die Maschine, ist das Maß" Ethische Aspekte der Informationstechnologie. In: Niederlag W, Lemke HU, Bondolfi A, Rienhoff O (eds) Hochtechnologiemedizin im Spannungsfeld zwischen Ökonomie, Politik, Recht und Ethik. Health Academy 1(2005):58–63

Wiegerling K (2012) Zum Wandel des Verhältnisses von Leib und Lebenswelt in intelligenten Umgebungen. In: Fischer P, Luckner A, Ramming U (eds) Die Reflexion des Möglichen. Zur Dialektik von Erkennen, Handeln und Werten. LIT, Münster, pp 225–238

Wiener N (1960) Some moral and technical consequences of automation. Science, New Series 131(3410):1355–1358

Winner L (1980) Do artifacts have politics? Daedalus 109(1):121–136

Zuboff S (1988) In the age of the smart machine. The future of work and power. Basic Books, New York

Zuboff S (2019) The age of surveillance capitalism: the fight for a human future at the new frontier of power. Public Affairs, New York

Part VII
Management and Working Environment

Management for Digitalization and Industry 4.0

Julia Arlinghaus and Oliver Antons

Contents

J. Arlinghaus (✉)
Otto-von-Guericke University Magdeburg, Magdeburg, Germany
e-mail: julia.arlinghaus@iff.fraunhofer.de

O. Antons
Chair of Management Science, RWTH Aachen University, Aachen, Germany
e-mail: antons@oms.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_49

1 New Management for Digitalization and the Fourth Industrial Revolution

1.1 Changes in the Current Competitive Environment

Global trends are changing how and what European companies manufacture. Progressive globalization and constant and immediate availability of relevant information through the Internet are not only changing the structure of global value networks but also compelling companies to operate under increasingly intense time, quality, innovation and cost pressure. Hence, the associated complexity of products, value networks and the planning and control processes is consequently growing.

European companies are thus facing numerous seemingly conflicting challenges that their respective management has to reconcile. First, companies must establish suitable planning, control and risk management structures to respond to the growing complexity of global networks. At the same time, companies are steadily striving to optimize and streamline value-added and administrative processes to reduce costs. Second, customers are demanding such an increasingly greater degree of personalization of products and services that lot sizes in procurement, manufacturing and even distribution are steadily decreasing as a result. At the same time, environmental and sustainability demands on companies are growing constantly. Third, ensuring profitability and competitiveness given high labor costs in Europe while fulfilling social responsibilities shaped by government and society, such as the reduction of poverty when outsourcing both globally and locally.

This situation has made the vision subsumed under *Industry 4.0* a beacon of hope for European business. It promises innovative products, services and business models, increased versatility and robustness of production, logistics and transportation systems, and protection of innovation, value creation and jobs in Europe (Bitkom 2014). This vision has long since grown beyond its origin as a digitalized smart factory to encompass virtually every domain of business, including logistics, transportation, energy, facilities and products (Kagermann 2017). A multitude of studies published in recent years describes the changes associated with digitalization and automation and enumerate their potential capabilities. These range from internal capabilities, such as cutting costs and boosting productivity, to approaches to improved customer retention down to increased sales through expanded and novel

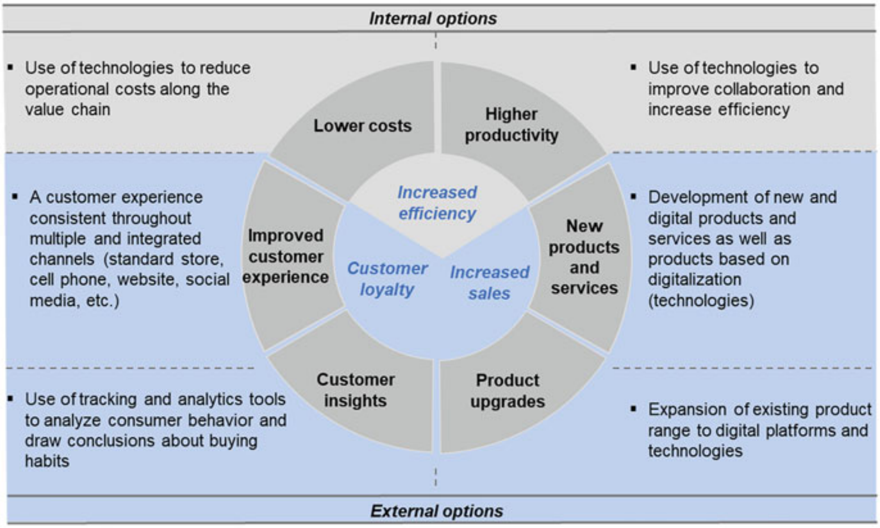


Fig. 1 Internal and external capabilities of digitalization (based on Strategy & Transformation Consulting 2018)

products and services (Fig. 1). This promises to shorten time-to-market (i.e., the time from the beginning of product development until market launch) by as much as 50%, to cut maintenance and repair expenditures by as much as 40%, to reduce machine downtime by as much as 50%, to boost manufacturing employees' productivity by as much as 45%, and to increase the accuracy of forecast customer demand by 85% (McKinsey Global Institute 2015). This explains the multitude of companies already pursuing concrete **digitalization projects**. A current study by the KfW Bankengruppe (2018) reveals that over half of the companies in Germany have firm plans to complete digitalization projects in the next two years and final decisions are pending in another 23% of the companies.

1.2 Necessary Adjustments in Management

Realizing the aforementioned potentials of the Fourth Industrial Revolution through corresponding digitalization projects is confronting European companies' management with new challenges. Along with answering a multitude of detailed questions about using new technologies, it is essential to decide which business functions will run supported by computers or even fully autonomously in the future and how they must be planned, controlled, and monitored. Furthermore, management must refine its corporate culture so that employees and processes are prepared for the transformation into a self-managing and digital company. Only when management systems, organizational structures, education and training systems as well as innovation, communication and human resource management have been adapted to the growing agility, decentralization and technological direction will European companies be able to withstand global competition in the long term too.

2 Digitalization Strategy as Part of Future Corporate Strategies

2.1 Difficulties Launching Digitalization

The diverse capabilities associated with digitalization and automation are an important building block for business and government dealing with currently fierce competition. At the same time, both business and government are often confronted by the question of "where to begin?" On the one hand, the desire to be technology and innovation leader often exists in conjunction with the concept of "Industry 4.0." On the other hand, projects that take current challenges and problems as their starting point can be implemented significantly more easily. Small and medium-sized businesses in particular often do not have the financial and human resources needed to tackle digitalization and automation that does not arise out of a specific challenge.

2.2 Complex Industry 4.0 Market Environment

The complex Industry 4.0 market environment is making the decision situation even more confusing for many companies. Deciding when, in which units and how a company ought to approach digitalization and Industry 4.0 requires analyzing and assessing the complex interplay between established manufacturers, large digital companies, Industry 4.0 technology start-ups and niche providers in order not to become unduly **dependent** on either particularly small or particularly large and powerful partners in the medium term.

Figure 2 provides an overview of the current Industry 4.0 market environment. Tech giants (e.g., *Google*, *Amazon,* and *IBM*) and manufacturers are vigorously striving to acquire smaller technology providers. *Daimler*, for instance, invested in the delivery robot start-up *Starship Technologies* in 2016, thus transforming from a pure automaker into more of a transportation provider. *Amazon* acquired wireless router manufacturer *Eero* in 2019 to expand its dominance in the smart home market achieved by its *Amazon Echo* devices. Tech giants and traditional manufacturers are additionally entering close collaborative partnerships of mutual interest. *VW* and *Google*, for instance, announced a collaborative partnership in quantum computer development in 2017. At the same time, established manufacturers' innovativeness is being diminished by niche providers and Industry 4.0 start-ups that are taking

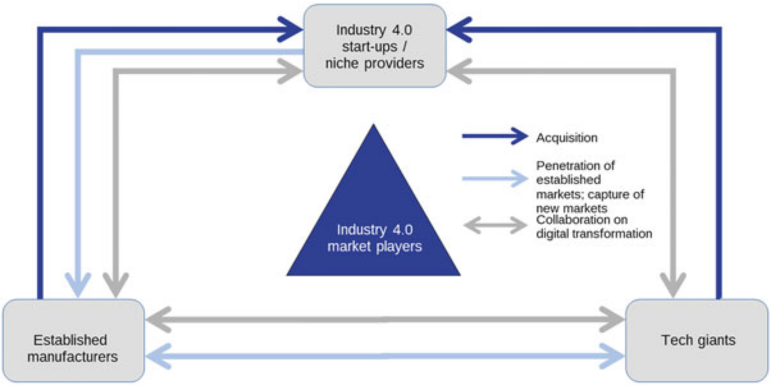


Fig. 2 Industry 4.0 market actors (based on Bechtold et al. 2014)

smaller market shares away from them and occupying emerging market niches. Experts expect that *Tesla* will continue to dominate the German electric vehicle market, while established industry players, such as *BMW* or *VW*, will play secondary roles in the electric transportation sector at best. The *Shapeways* company founded in 2007 offers a comprehensive 3D printing service from development to production down to order fulfillment. Moreover, tech giants are also increasingly competing with established manufacturers. *Google* is already so far advanced with its development of self-driving cars that its subsidiary *Waymo* in Phoenix, Arizona is offering the first commercial robotic taxi service. *Amazon*, on the other hand, has announced its intention to establish a novel service for on-demand 3D printing.

A company's **positioning** is fundamentally important in this challenging market situation and requires precise understanding of one's own starting situation and the role digitalization can play for the company internally and externally. Established manufacturers consequently ought to continue pushing their collaboration and acquisition efforts and systematically develop their own digitalization expertise further to be able to survive in the long term.

2.3 Digitalization Strategy as a New Element of Corporate Strategy

The direct link between harnessing the efficiency and sales capabilities ensuing from digital assistance and the digitalization of products, business models and processes necessitates rapid and systematic adjustment of management systems. Consequently, **digitalization strategy** must become an integral part of future corporate strategies (Schröder et al. 2015). The digitalization strategy defines how digitalization and automation are ingrained in a company organizationally and culturally to produce smart goods and services and process efficiency as well as to redesign customer

Fig. 3 Key issues of a digitalization strategy

experiences and customer interfaces (Gläss and Leukert 2016). The digitalization strategy thus specifies *how* a company profits from automation and digitalization when providing customer services and *what* products and services a company offers based on which business models. Moreover, the strategy prioritizes the units in which digitalization and automation ought to be expedited in the company and defines a substantive road map and schedule based on this. Ultimately, it defines how the change process is ingrained in the organization and how employees ought to be trained for its implementation.

Digitalization activities in a company can be assigned to three categories: (1) *digital process excellence and supply chains*, (2) *organization, culture, and values*, and (3) *innovative customer experiences and interfaces* (Fig. 3). *Digital process excellence and supply chains* encompass every effort to digitalize and automate internal and cross-company processes by employing IT systems and new technologies to boost processes' efficiency, quality, flexibility and robustness (Sect. 3). The changes surrounding *organization, culture and values* are needed to provide executives and employees the general conditions required for successful digitalization (Sect. 4). Finally, *customer experience and interfaces* are being changed by data-driven analyses of customers' buying habits and communication modes (Sect. 5). In combination, these modifications are spawning new and digital products and thus innovative business models (Sect. 6).

2.4 From the Digital Vision to the Digital Company

The development and implementation of a digitalization strategy is a years-long process. Exact knowledge of one's own market situation and a concrete vision of how digitalization will change processes, products and customer experience in one's own company constitute the starting situation. This process is often called digital transformation. Although the term suggests a one-time change, it stands for continuous adaptation and ongoing development that employees and executives have to complete as they change processes, organizational structures and products, etc. to profit from digitalization's capabilities (Schuh et al. 2017). It is also employed to denote the reorganization of processes, organizational structures, corporate values and skills, which facilitate changes in goods and services, the range of services, and customer interaction.

3 Digital Process Excellence and Digital Value Chains

3.1 Digitalization as the Foundation of Data-Driven Management Decisions

The so-called third digital revolution already increased productivity substantially by supporting business processes with computers. Many processes in companies are nevertheless still performed wholly manually, neither supported by computers nor dominated by connected physical and virtual worlds. The term *digitalization* originally denoted the conversion of analog information into digital information for storage and processing by high-performance IT systems to support employees and processes in a company optimally. The term also covers the linking of digital data with physical items in conjunction with output on analog devices to be read by humans or machines with a time lag. This makes it possible to refine processes continuously, thus achieving more transparency and improved forecasting capability, flexibility and efficiency. The linkage of digital data with physical items and the systematic acquisition, aggregation, analysis and interpretation of business data with external information form the foundation of **data-driven management decisions** and ultimately the basis for the development of new products, services and business models.

3.2 Stages of Process Digitalization

Many companies' prime motivations for occupying themselves with Industry 4.0 and digitalization are the major cost cutting capabilities and the opportunities to boost productivity and quality. Various stages that build upon each other must be completed when transitioning from pure computerization to autonomous systems and digital management consonant with the vision of Industry 4.0 (Fig. 4).

The integration of computer systems in machines and process flows to boost efficiency, especially in repetitive processes, termed **computerization,** establishes the foundation of process digitalization and is usually ascribed to the Third Industrial

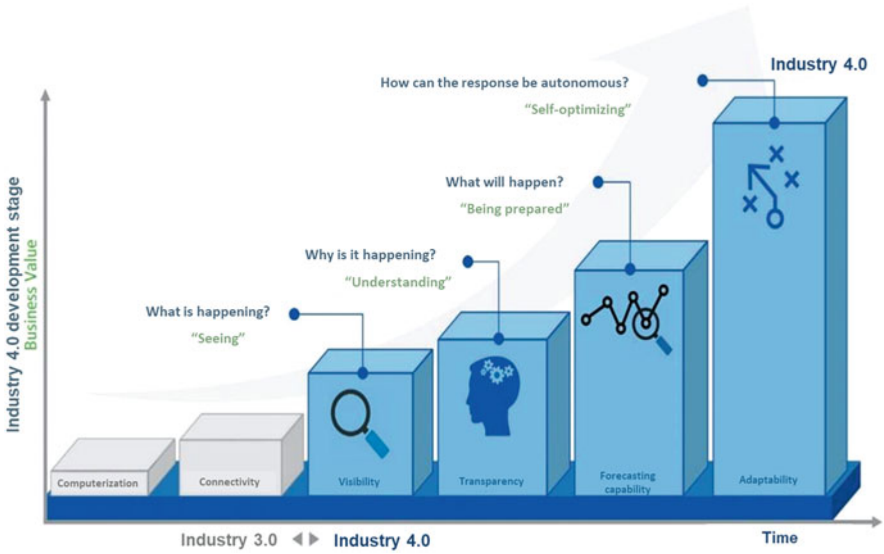


Fig. 4 The stages of process digitalization and digital management based on it (Schuh et al. 2017)

Revolution. A CNC milling machine exemplifies this stage well: Whereas the machine that machines workpieces runs automatically thanks to computerization, the model file specifying the stages of machining is transferred to the CNC milling machine manually. The CNC milling machine is unconnected—it is a stand-alone machine. Connecting the first-stage processes and objects is called **connectivity**. A multitude of developments in new physical communications structures, e.g., 5G, and logical refinements, e.g., IPv6, constitute the technical foundation for all actors' joint communication in the digitalized process: the connectivity of manufacturing facilities, IT networks and employees (Schuh et al. 2017).

This connectivity makes it possible to map every process in a company digitally and to generate a so-called **digital footprint**, i.e., a virtual and accurate copy of the real processes. The digital footprint thus renders relevant information throughout business operations, e.g., product life cycle data, machine and job order data, and manufacturing, logistics and material stream data, **visible**. A sufficiently detailed digital footprint facilitates the development of improved process understanding as well as targeted optimization.

Linking and analyzing the data underlying the digital footprint enables making the actual processes and potential disruptions **transparent**. This permits supporting, improving, and expediting decisions because causal relationships are fully understood. Cloud-based software systems make it possible to manage fields, for instance. Combining satellite data with sensor-based crop information enables optimizing and applying fertilizer quantities environmentally soundly.

When the digital footprint projects transparent events onto the future, probabilities of occurrence can be estimated for different conditions. This enables companies to forecast impending events and take appropriate actions whenever necessary. *Predictive maintenance* is based on continuous analysis of machine parameters and other measurable factors that affect malfunctions. The causal relationships identified are used here to forecast different components' future failure modes and take individual maintenance actions. **Forecasting capability** thus enables companies to reduce downtimes and adjust maintenance processes to actual machine utilization.

The knowledge garnered from the transparency and forecasts generated can be continuously adapted to develop an adaptable, self-optimizing system. **Adaptability** enables companies to adapt manufacturing, among other things, to changed general conditions in the business environment autonomously and in real time. Such autonomy ranges from autonomous internal logistics to extensive autonomous supply chain operations in which a machine in the digitalized process reorders needed parts without human assistance (ten Hompel and Henke 2017). Self-optimizing manufacturing is not necessarily the goal of *digital transformation*, however. The magnitude of transformation remains contingent on the specific company's requirements, the particular processes and structures, and a comprehensive cost-benefit analysis (Gassmann and Sutter 2016).

3.3 Customer Orientation and Integrated Processes Characterize Digital Value Chains

Digitalization and automation present new capabilities to organize cross-company processes and global supply chains efficiently. Just as digital products and business models, digital supply chains are characterized by customer centricity and customization as well as rigorous cross-company integration of value-added processes (Accenture 2016b).

Platform foundation: Cloud-based platform systems that connect the different members of a value network and thus data from the widest variety of sources, such as mobile endpoint devices, different points of sale, and internal and external data warehouses, and provide them to all actors in the network in real time are the foundation of digital value networks.

Planning based on customers and greater segmentation: Digital supply chains are planned and controlled based on customer demand. Available customer data are used to identify trends and patterns in demand, thus dividing customers into increasingly numerous and smaller **segments**. This makes it possible to provide optimal products and suitable delivery services to these segments corresponding with specific demand patterns and the demands on sales channels and service levels. Switching from traditional make-to-stock strategies to customer-driven manufacturing and delivery not only improves demand forecasts and thus inventory and capacity decisions for individual segments but also expedites and makes supply chains demand-responsive.

Communication in real time: Social platforms facilitate cross-company and cross-organizational communication and collaboration among colleagues and stakeholders. Qualitative and quantitative information is exchanged in real time. The parties can, for instance, discuss reports and scenarios and plan and control sale and manufacturing processes in real time. The great advantage is that every discussion and exchange of information, which often take place in emails, thus making them difficult to reconstruct later, is saved permanently and with every change.

Integration through global network planners and control towers: Greater integration and rigorous end-to-end orientation also make value networks more effective. Digital supply chains will be planned beyond the company confines in the future. Moreover, the new role of *cross-company and cross-organizational network planner* will be created. Planners will be given full transparency of all operations along all value stages and full responsibility for the complete value chain's performance. Many of the planning and control functions in a digital supply chain, now organized separately in different companies and business units, will be centralized in a so-called integrated *control tower* to produce the requisite transparency. This will enable cross-organizational decision-making and steady integration of manufacturing, procurement and customer service in planning and control units. This will establish transparency about current stocks in inventory and in transit, customer priorities, etc. and simultaneously facilitate rapid responses to specific events (e.g., a transport vehicle accident) along the entire value chain. Comprehensive analysis of the potential causes of disruptions in process flows ultimately make cross-company **risk management** and thus continuous improvement of processes to **optimize** costs, inventories, quality, customer value and capacity utilization possible.

4 Digitalization, Industry 4.0 and the Role of Organization, Culture, and Values

4.1 Digitalization Is Changing the World of Work

Digitalization and automation are changing the work environment and ways of collaborating in companies. The ability to respond to volatile market conditions rapidly and flexibly necessitates agile processes in lieu of long-term planning (Gehrckens 2016). While repetitive processes are increasingly being supported by IT or fully automated, creativity and problem-solving skills are gaining importance. Digitalization also increases the transparency of operations and outputs. In combination with new forms and means of communication, team members can collaborate flexibly from different locations. Younger employees are particularly requesting such **flexible work arrangements** (e.g., telework) as well as greater freedom, more participation, and a balance between life and work (Arnold et al. 2015).

4.2 Competences and Further Training

Companies require qualified professionals for new forms of work and the use of new technologies. Insufficient expertise is preventing small and medium-sized businesses in particular from profiting from digitalization. The growing dynamic in the development of new products and services requires employees in teams from different organizational units to develop solutions collaboratively and autonomously. Interdisciplinary skills in the use of technologies as well as professional and methodological skills, such as creativity, interdisciplinary and process-driven thinking, and social skills for communication and collaborative partnerships must be fostered. This makes training a major success factor for companies, which therefore ought to be ingrained in corporate strategy and culture in such a form that continuous investments are made in training, and it is made integral to professional life.

Companies are thus facing the task of identifying training demand early and accurately and incorporating it in human resource (development) planning. To do this, companies need a comprehensive overview of their employees' knowledge and skills profiles to ascertain the respective need for training. Digitalization also offers new options for implementing training. Innovative and **interactive learning formats**, such as management games and simulations, also reach groups of people that are difficult to motivate for training with standard classroom programs. Learning on the job, i.e., "on demand," additionally permits learning without social pressure.

4.3 New Forms of Work Require New Management Concepts and Organizational Forms

These new forms of work require new management concepts too. A sense of responsibility and creativity cannot be forced out of people by pressure, though. Instead, they are based on the principles of motivation, encouragement, and empowerment (Oestereich 2015). The more a team is self-organized and the more its members are distributed geographically, the more it becomes management's job to establish a common framework in which employees can work independently and creatively. For executives, this means a substantial shift away from more traditional values such as constancy and perfection to creativity and agility (Schweer and Seidemann 2015). Executives are the major knowledge facilitators and decision makers in traditional organizational and management structures. When solving complex problems, they **facilitate** their employees' problem-solving processes instead of giving instructions and acting as key knowledge facilitator themselves. Executive also become coaches and service providers for their teams. This means that they share power and involve employees in management and strategic decisions more. Integrative culture makes employees contributors, thus making it a **competitive edge** (Stoffel 2015).

To act more rapidly and more flexibly in the future, decision-making skills must be linked to the individuals with the situationally relevant knowledge and requisite skills—and not just to the formal position of executive (Grund 2015). At the same time, knowledge and skills must be recombined repeatedly so that networks of collaborating actors supersede traditional organizational silos.

New forms of work and changed demands on employees require fewer hierarchical structures. The term "agile structures" is often used here. The term *agility* denotes a company's capability to respond to changes rapidly. An agile organizational structure enables a company respond to unforeseeable demands faster than its competitors can. Advanced internal and external communications systems are the key media for breaking up information silos and coordinating structures as well as working efficiently despite fewer instructions and greater decision latitude. Although agile organizational structures dispense with such principles as division of labor and differentiation, agile organizations are not fully nonhierarchical, either. Instead, the particular market situation governs the flow of information and instructions. Traditional linear structures will likely be found in certain business units in the future too (Rump et al. 2018).

First, the requisite conditions must be fulfilled to enable established companies with traditional structures to become agile. Rump et al. (2018) recommend working in parallel structures at first because, unlike start-ups that grow right into agile structures, this enables managing day-to-day business and operation side by side and focused on efficiency and excellence. The implementation of so-called innovation labs and think tanks and the development of one's own start-ups, for instance, are suitable for such structures. They can work on current strategic issues separately in agile and connected structures, arrive at more creative solutions faster, support the traditional business in the process, and receive support themselves. Approximately seventy-five companies in Germany are operating so-called innovation labs at this time (Capital 2018). *Lufthansa Innovation Hub*, for instance, works on its own innovation projects as well as on topics contributed by the company and on venture capital activities. *Daimler Digital Life Hub* intends to unite all digitalization activities in the company and announces company-wide *challenges*, for instance, in which employees submit proposals of their own, which colleagues can assess.

Many processes in administrative units in particular will require a clear structure and division of work in the future too. Active interface management is required for a functioning parallel structure, though, to avoid duplicated work, frustration among both structures' employees and potentially bad investments (Rump et al. 2018).

4.4 The Role of Culture and Values: Trust as a Key Value of Digital Transformation

Digital transformation means that old thought patterns and familiar processes as well as long-standing hierarchies and products are disappearing. Companies are evolving into service providers and problem-solving partners. This translates into radical life-changing events for many employees. Furthermore, the far-reaching digital support and automation of many processes cause fear. A radical change is taking place, especially among jobs that merely require low qualifications. Computers are increasingly performing such automatable jobs. So-called chatbots, for instance, could steadily replace traditional call centers in the future by providing customer assistance on a website or in a chat client following a predefined script.

Fear and insecurity cause major difficulties for any change. Proven processes and familiar organizational structures are no longer able to establish the necessary trust. As with any change, it is important that employees understand the problems their company is facing and the reasons changes are necessary. Above all, changes need a clearly defined framework so that employees support and accept them (Doppler and Lauterburg 2008). Since they especially present and convey a company's culture and values in the process, executives are the key actors whenever trust and a framework for new working conditions have to be established (Laudon 2017). Corporate culture thus plays a particularly important role whenever necessary changes are being implemented—the values ingrained in it often govern collaboration in the company consciously and unconsciously, provide a sense of direction, motivate, and create meaning (Lippold 2017).

4.5 Ethical and Legal Concerns

How extensively digitalization will change the world of work is still unclear. The spectrum already ranges from pure IT support of human decision-making through complete delegation of decision-making skills to computer programs. Computers that act autonomously as stockbrokers, for instance, have become commonplace. Jobs that require higher qualifications, such as the development of personalized insurance rates or the drafting of contracts, legal research, and pleadings in the legal advice sector, are also being replaced already.

The amount of changes practicable largely depends on hitherto unresolved ethical and legal issues too. One example is the issue of decision making in autonomous driving. This issue can be transferred to other smart and self-learning objects, such as machines and planning systems. The legal situation is still unclear precisely in situations in which a decision made by a smart object (e.g., in machine learning application) can no longer be clearly reconstructed later, not even by experts.

One example of poor decision-making in automated and digitalized applications is the project on automatic application document assessment recently halted by *Amazon*. An artificial intelligence was designed in this project to automate the assessment of application documents. The software developed turned out to favor male applicants. This action could be explained by the training data employed for the machine learning, namely historic application data from *Amazon's* early days when the majority of its employees were men.

Lawmakers' involvement is especially needed when such ethical and legal issues are being clarified to initiate social and political debate and establish **legal standards**. Government has established new standards through the General Data Protection Regulation (GDPR), among other things, which constitute a compromise between the processing and utilization of personal data, on the one hand, and citizens' digital sovereignty, on the other (Althammer 2018). The ethics committee has compiled initial outcomes of discussion of ethical and legal issues connected with the introduction and use of autonomous vehicles (Ethik-Kommission 2017). This has merely laid a foundation for legal standards, which lawmakers must finalize in the coming years.

5 Digitalization for Innovative Customer Experiences and Interfaces

5.1 Digitalization Is Changing the Customer Perspective

Digitalization is having a strong impact on customer loyalty. Customers are increasingly better informed, trying more out, and able to change from one provider to another faster and easier (switching economy). Customers are now finding information on websites, user rankings and tests on *YouTube*, opinions and rankings in online forums and on social networks instead of only in stores as in the past so that marketing must also factor in customers' growing orientation toward recommendations from these novel sources rather than centrally developed sales and marketing messages. As in corporate management, the importance of values for dealings with customers is also growing substantially. Products are no longer simply sold but penetrating deeply into the different spheres of customers' lives as smart, digitalized and connected products. *Smart medicin*e, for instance, is creating new capabilities to monitor illnesses such as skin cancer through small, connected sensors and is already enabling automated insulin administration among diabetics. *Smart homes*, on the other hand, connect different household appliances (e.g., lamps, consumer electronics, and smart washing machines and refrigerators). Users are thus not only notified of the end of a wash cycle but can automatically add detergent and food that are running low to their shopping lists, to name but one of already numerous applications.

A substantial amount of **trust** is needed for customers to allow this penetration in their private sphere. To this end, customer communication must change and start with customers' values. Companies are therefore striving to make customers the focus of their activities even more than before: Interfaces to customers are being improved and customers are being offered the perfect shopping experience. Intelligent analysis of available customer data is the key to this.

5.2 Digital Customer Experiences

Innovative customer experiences integrate digital and analog channels and ensure that customers get exactly what they want. Many companies are striving to create as dynamic, unimpeded and continuous a customer experience as possible that satisfies various customer groups' widely varying expectations and needs simultaneously to retain customers on their customer journey (Accenture 2016a). All offerings have to be coordinated, no matter whether a customer contacts the company by app, through an online shop or on site in a store. A current study, for instance, reveals that the customer experience is nearly just as important as quality and price. Sixteen percent of customers would even pay higher prices for a perfect customer experience (PWC 2018a, b).

5.3 Intelligent Analysis of Customer Data

The key to enhancing the customer experience digitally lies in the analysis of customer data and in their combination with other data sources. Sales information found on every receipt can be collected anonymized in large quantities and employed to generate revenue, for instance. If what buyers of a specific product typically purchase as well is known, for instance, targeted additional online and stationary offers can be made to customers. This makes it possible to optimize product lines in stationary retail as a **function** of location, day of the week and time of day. Sensors can be used to record how long customers have stood in line. Not only can they be provided additional information while they wait, but delays in the sales process can also be detected. Customer streams can be visualized by tracking foot traffic patterns in supermarkets or train stations by means of personal smartphones' WLAN fingerprints. Staffing requirements can be optimized and information on buying habits and **brand loyalty** can be acquired too. Whether a customer buys a croissant at baker A first and then coffee at baker B is visualized, for instance.

Skillful interpretation and aggregation of data into so-called *customer journeys* permits identifying where a customer leaves the sales process in online sales processes too—when the payment process takes too long, for instance. This makes it possible to take countermeasures. When combined with the systematic analysis of customer groups, discounts can be offered selectively or personal advisers or automated chatbots can make contact with customers, for instance.

Finally, digitalization allows establishing dynamic pricing in new industries. Analysis of large quantities of data and corresponding algorithms can now apply what consumers have long only been familiar with from gas stations and flight or hotel reservations to entirely new products. *Amazon* already adjusts its prices several times a day. Whereas prices were usually only changed based on simple rules in the past, e.g., when the competition did so, smaller businesses can now develop and implement their very own sales strategies with the help of dynamic pricing. Algorithms adjust prices so that a certain product is always kept deliverable until the last item is sold at the end of the season, for instance, or so that the highest profit margin is always obtained. Companies such as *IBM* or *Blue Yonder* now partly offer cloud-based solutions that calculate **dynamic price** based on purchase histories, sales figures or even the chosen access channel.

5.4 Personalized Services and Products

A current study reveals that customers want personalized offers and products. That is one of the reasons why a majority of customers still shops wherever retailers remember their purchase decisions and make them personalized offers. Such personalized offers can increasingly now be developed by self-learning algorithms and by combining of data from different sources. *EBay*, for instance, employs machine learning and smart algorithms to improve the relevance of search results as well as navigation. *Stitch Fix* offers a subscription service for so-called *curated shopping*. It strives to combine online retail with advising familiar from retail shopping. Apart from personal information, information from social media is also combined into a selection of five items of clothing in a box to give individual, personalized fashion recommendations. Information on returns is used to model customers' personal taste progressively better. This lowers the return rate over time and increases profitability (Möhring and Schmidt 2015). *Adidas* also employs this approach and assembles outfits with an automated system from the start-up *FindMine* without needing real stylists.

So-called mobility-on-demand offers transfer the idea of personalized service to the transportation sector. The actual demand for transportation becomes the starting point for designing transportation offers. Personal shuttle service can be ordered by app in Hamburg, for instance. Customers with similar routes are automatically grouped in ride pools with the aid of an algorithm, transported together, and brought to their individual destinations.

5.5 Customer Relationship Management

Data-driven knowledge about customers' personal needs also allows taking customer relationship management to a new level. Personalized and customized communication and cross-channel communication combine a company's online and offline offers. This means that all information is always immediately available, no matter which channel a customer selects. The customer is actively helped with changing the channel smoothly. A retailer can be referred to online, for instance, while providing directions and proposing appointments to reduce waits. Conversely, a retailer can use its connectivity with the online shop to ascertain demands from customers' histories and make them fitting **additional offers** (Correa and Alonso 2017).

Analysis of communication data is the foundation of so-called chatbots. They can answer customer inquiries on websites and in messengers at any time. Appropriate responses to customers' most frequent concerns in the sales process are compiled so that an automated chat can be established. Although recognition of emotions still causes enough problems that a majority of inquiries are still forwarded to support staff typically, customers can already be given a feeling of personal communication with a time lag, even online. *Dell*, for instance, provides the option of assistance from its virtual assistant *Ava* whenever questions arise when on their website. *Wetter Online* provides a WhatsApp chatbot service that provides information at a desired time of day and related to one's current location.

6 Digital Products and Business Models

6.1 Digitalization Is Changing Business Models

Digitalization is thus changing the relationships between companies and their customers decisively. Demand for digital offerings is rising in both the B2B and the B2C sector just as much as the number of Internet users and Internet-capable devices is growing worldwide, (Deloitte 2013; Roland Berger and BDI 2015). Despite the great capabilities ensuing for many companies from the use of new technologies and the digitalization of processes and products, the big winners of digitalization appear to be the companies that have digitalized their business models (BCG 2009). Among others, these include *Uber*, *Skype*, *Apple*, *Airbnb* and *Amazon*. Just as processes and products, many business models have been assembled over time, have evolved incrementally, and provide employees and executives a stable framework for day-to-day business so that the digitalization of business models is proving to be perhaps even more difficult than the digitalization of processes and products.

A business model describes the basic principle upon which an organization establishes and communicates values (Osterwalder and Pigneur 2010). It describes the business offerings, the value provided to customers, the design of the value chain, and thus the logic a company follows to make a profit too. It thus answers four key questions:

(1) Who is the target customer?—The business model describes which customer segments with which needs and wishes are met through which channels.
(2) What do we offer customers?—The business model defines how a company creates value for customers.
(3) How do we provide the service and how do we produce it?—The business model defines the design of the processes, resources, skills, and partners along the value chain, which are needed to fulfill the value proposition.
(4) How is value produced?—The business model describes the cost structure and revenue mechanics and thus how the company earns money (Gassmann et al. 2017).

Gassmann and Sutter (2016) define a digital business model as the integration of e-business and an Internet-based value proposition based on smart value chains. This means that processes and products are initially supported with IT with the aim of boosting efficiency and quality and cutting costs without changing the product itself.

Here, a smart value chain denotes the flexibly and efficiently designing internal and cross-company value-added processes by employing new technology, connectivity, and autonomous control. Digitalized products target service orientation and consumer convenience. New technologies and applications (e.g., social networks, new analysis methods and cloud solutions) enable products with software and/or connected with the Internet by sensors and actuators to provide additional services and to produce more indirect customer value. At the same time, companies can use the data generated to improve products and services further. In the factory environment, for instance, service contracts and assistance services drawing on information supplied through the cloud by pertinent sensors can in addition be offered to machines. Printers in the office environment have similarly been able to reorder consumables automatically for some time.

Fully digitalizing a business model fully requires digitally adapting and properly altering both the value proposition and the value chain. Digital business model innovation is only spoken of when the transformation is radical. High service and customer orientation is the key feature of digital business model innovation. The communication integrated in every phase of the customer relationship supplies sufficient data to assess demand for, quality of and effectiveness of a product fully and thus to test and optimize it quickly and easily. Digital business models are thus characterized by agility, leveraging of immaterial resources, use of digital channels, disclosure of intellectual property, and the addition of services to products.

6.2 Digital Business Models: Select Examples

Gassmann et al. (2017) demonstrate that 90% of all business model innovations are recombinations of 55 business model patterns. A large part of these business model patterns makes use of digital technologies or allows much easier implementation than conventional solutions. These include *pay per use models* in which customers only pay for actual use instead of a set amount. Rather than purchasing a machine or a printer anymore, customers pay for the number of hours run or printouts. Carsharing services such as *Car2Go* and many media offerings such as *YouTube movies* function based on this model.

Business models based on *platforms* are particularly successful at present. A platform's scope is the key success factor with an eye toward bringing providers and consumers together. Platforms such as *Expedia* or *Autoscout24* that make different providers' travel or vehicle offerings available centrally and standardized are particularly successful because these markets are very opaque to customers and their service simplifies searches and comparisons considerably. This is also attractive to providers that pay a fee to be listed.

The so-called *freemium* model provides users basic services for free and premium services for a fee. Users may get additional storage (*Dropbox*), additional information (*Xing*, *LinkedIn*) or music without ads (*Spotify*). The free offer attracts a vast number of customers while premium customers cross-finance the overall offering for the most part.

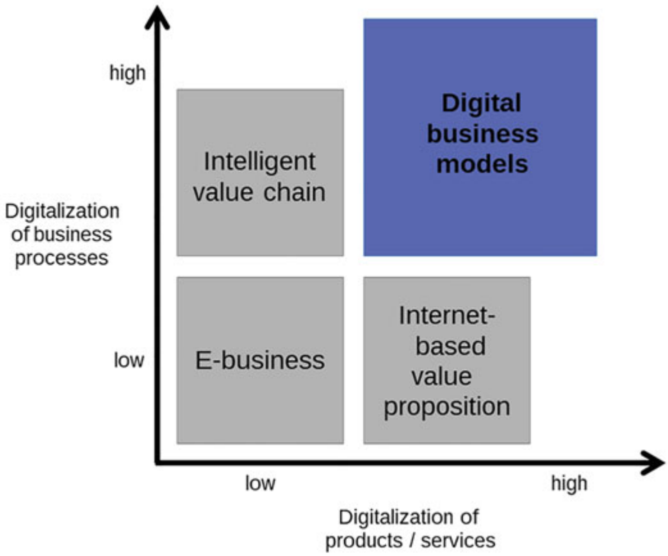


Fig. 5 Four forms of digitalization (Gassmann and Sutter 2016)

The *subscription business model* involves paying a periodic fee as for a print newspaper, for instance. Digital technologies enable transferring this model to a multitude of products and services, such as *Netflix* and *Parship*. Unlike *pay-per-use* models, revenue with this model is relatively stable and very predictable. In turn, customers are always guaranteed availability.

Digital business models will probably not supplant traditional models entirely. Rather, traditional and digital business models will exist side by side in the future (Fig. 5).

7 Conclusion and Outlook: Reality and Risks of Digitalization and the Fourth Industrial Revolution

7.1 Reality in German Business

The major media visibility of digitalization and Industry 4.0 often gives the impression that digital transformation is already far advanced in German business. **Flagship projects** showcased in the media appear to be industry standards. Although in reality, most companies—small and medium-sized as well as large—are removed from integrated digitalization and connectivity of processes, products, and business models in their day-to-day business. Fax orders, Excel spreadsheets and decisions based on “gut feelings” are still the order of the day in many companies’ reality.

Many companies are sensing a certain urgency and a need to deal with digitalization and Industry 4.0. The motivation for this often arises less out of concrete issues, however, especially in small and medium-sized businesses, than pressure diffusely generated by media and observation of the competition. The term "digital transformation" suggests that the necessary change process is a one-time adaptation with a clear beginning and end. On the contrary, it is a continuous process of adaptation that every company must organize and manage for itself specifically. This also means that not every process and not every product has to be digitalized. In the same vein, one specific technology does not automatically help boost customer value and process efficiency because the competition is using it. Rather, every company must reassess any digitalization project's general conditions and costs and benefits in the given situation.

7.2 Reality in Collaboration with People

People, i.e., employees and executives, remain crucially important to this continuous adaptation intended to increase digitalization and connectivity. Employees must understand why changes are necessary to support them. Since this is very difficult without an actual crisis, communication and trust are all the more important. Becoming more agile and flexible requires employees to turn from workers into shapers, to collaborate on designing change actively, and thus to be involved in strategic decisions. It is also essential to train employees to handle new technologies so that they use them effectively and do not wind up feeling uniformed, helpless and afraid of becoming expendable.

Digitalization means **radical change**. Executives must adapt their actions accordingly too. This in turn requires appropriately refining corporate culture and the values ingrained in it. The new forms of work require that the organizational structure provide employees and executives the general conditions appropriate for new participatory collaboration.

7.3 Digitalization Not an End in Itself. What, Though, Are the Costs of Not Digitalizing?

Since the necessary support for changes can only be obtained from employees when they understand the rationale, digitalization may not be an end in itself. At the same time, companies have to keep an eye on the opportunity costs of not digitalizing too. Even if they do have acceptable profit margins but cannot employ substantial financial and human resources, small and medium-sized businesses in particular have to scrutinize and anticipate how consumer demands are changing and how the competition is responding globally and modify their corporate strategy accordingly.

7.4 *Risks Are Frequently Underanalyzed*

Risks associated with digitalization are frequently underanalyzed, one hand, because the potential capabilities are enticing and, on the other hand, because the information base for assessing attendant risks is inadequate. This translates into new challenges, even for education and training. Engineers in manufacturing units, for instance, are highly trained to anticipate purely technical risks but unprepared to deal with malicious physical tampering and attacks such as *social engineering*. IT system and digital technology use and protection must therefore be integrated in many training programs to be able to assess the costs and benefits of digitalizing products, processes, supply chains, and business models fully.

References

Accenture (2016a) Die 4,6-Billionen-Euro Chance. Available at: https://www.accenture.com/de-de/insight-digitalisierung-steigerung-kundenerlebnis-ertragswachstum. Accessed 13 Mar 2019

Accenture (2016b) Supply chain for a new age. Available at: https://www.accenture.com/t20170206T210529__w__/id-en/_acnmedia/PDF-42/Accenture-Supply-Chain-For-A-New-Age-2.0.pdf. Accessed 13 Mar 2019

Althammer T (2018) Datenschutz und IT-Sicherheit in Zeiten der Digitalisierung. In: Digitaler Wandel in der Sozialwirtschaft. Nomos Verlagsgesellschaft mbH & Co. KG, pp 223–240

Arnold D, Steffes S, Wolter S (2015) Mobiles und entgrenztes Arbeiten

BCG (2009) Business Model Innovation: When the Game gets tough, Change. Available at: https://www.bcg.com/documents/file36456.pdf. Accessed 13 Mar 2019

BDI (2015) Die digitale Transformation der Industrie. Roland Berger, Munich

Bechtold J, Lauenstein C, Kern A, Bernhofer L (2014) Industry 4.0-The Capgemini Consulting View. Capgemnini Consulting, 31

Bitkom (2014) Industrie 4.0 – Volkswirtschaftliches Potenzial für Deutschland. Available at: https://www.produktionsarbeit.de/content/dam/produktionsarbeit/de/documents/Studie-Industrie-4-0-Volkswirtschaftliches-Potential-fuer-Deutschland.pdf. Accessed 13 Mar 2019

Capital (2018) Konzerne auf den Spuren von Startups 2018 Wie etablierte Unternehmen Acceleratoren, Company Builder, und Innovation Labs als Instrumente der digitalen Transformation nutzen. Available at: https://www.infront-consulting.com/relaunch/wp-content/uploads/2018/06/20180620_Online_final_komp.pdf. Accessed 13 Mar 2019

Correa F, Alonso R (2017) Development and implementation of a multichannel management support system for physical retail stores based on customer online data. Universidad de Chile

Deloitte (2013) Digitalisierung im Mittelstand. Studienserie Erfolgsfaktoren im Mittelstand, Roland Berger

Doppler K, Lauterburg C (2008) Change Management: den Unternehmenswandel gestalten. Campus Verlag

Ethik-Kommission Automatisiertes und Vernetztes Fahren (2017) Bericht Juni 2017. Available at: https://www.bmvi.de/SharedDocs/DE/Publikationen/G/bericht-der-ethik-kommission.html. Accessed 13 Mar 2019

Gassmann O, Sutter P (2016) Digitale Transformation im Unternehmen gestalten: Geschäftsmodelle Erfolgsfaktoren Fallstudien. Carl Hanser Verlag GmbH Co KG

Gassmann O, Frankenberger K, Csik M (2017) Geschäftsmodelle entwickeln: 55 innovative Konzepte mit dem St. Galler business model navigator. Carl Hanser Verlag GmbH Co KG

Gehrckens HM (2016) Agilität im Kontext der digitalen Transformation–Kernanforderung an die Organisation von morgen. In Digitale Transformation oder digitale Disruption im Handel. Springer Gabler, Wiesbaden, pp 79–108

Gläss R, Leukert B (2016) Handel 4.0: Die Digitalisierung des Handels-Strategien, Technologien, Transformation. Springer

Grund M (2015) Agile Softwareentwicklung als paradigmatisches Beispiel für eine neue Organisation von technischer Wissensarbeit. Das demokratische Unternehmen. Neue Arbeits-und Führungskulturen im Zeitalter digitaler Wirtschaft, 159–168

Kagermann H (2017) Chancen von Industrie 4.0 nutzen. In Handbuch Industrie 4.0 Vol. 4, Springer Vieweg, Berlin

KfW Bankengruppe (2018) Unternehmensbefragung 2018. Available at: https://www.kfw.de/PDF/Download-Center/Konzernthemen/Research/PDF-Dokumente-Unternehmensbefragung/Unternehmensbefragung-2018-%E2%80%93-Digitalisierung.pdf. Accessed 13 Mar 2019

Laudon S (2017) Wie die Digitalisierung die Führungskompetenz komplett neu definiert. In: HR-Exzellenz. Springer Gabler, Wiesbaden, pp 65–77

Lippold D (2017) Market-Oriented Business Management: Management in the Digital Age. Walter de Gruyter GmbH & Co KG

McKinsey Global Institute (2015) The internet of things: mapping the value beyond the hype. Available at: https://www.mckinsey.com/~/media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/the%20internet%20of%20things%20the%20value%20of%20digitizing%20the%20physical%20world/the-internet-of-things-mapping-the-value-beyond-the-hype.ashx. Accessed 13 Mar 2019

Möhring M, Schmidt R (2015) Daten-getriebene Unternehmensarchitekturen im E-Commerce für das präventive Retourenmanagement. INFORMATIK 2015

Oestereich B (2015) Praktiken und Prinzipien der Selbstorganisation: Führungsarbeit statt Führungskräfte. Das demokratische Unternehmen. Neue Arbeits-und Führungskultur im Zeitalter digitaler Wirtschaft, 231–245

Osterwalder A, Pigneur Y (2010) Business model generation: a handbook for visionaries, game changers, and challengers. John Wiley & Sons

PWC (2018a) The Future of Customer Experience Report. Available at: www.strategy-transformation.com/digitale-transformation-verstehen/. Accessed 13 Mar 2019

PWC (2018b) Kunden begeistern - vom Einkauf zum Erlebnis. Available at: https://www.pwc.de/de/handel-und-konsumguter/studie-retail-customer-experience.pdf. Accessed 13 Mar 2019

Rump J, Zapp D, Eilers S (2018) Erfolgsformel: Arbeiten 4.0 und Führung 4.0. Institut für Beschäftigung und Employability

Schröder C, Schlepphorst S, Kay R (2015) Bedeutung der Digitalisierung im Mittelstand. IfM-Materialien No. 244

Schuh G, Anderl R, Gausemeier J, ten Hompel M, Wahlster W (eds) (2017) Industrie 4.0 Maturity Index: Die digitale Transformation von Unternehmen gestalten. Herbert Utz Verlag

Schweer D, Seidemann S (2015) Die neue Macht–digitale Freiräume. Das demokratische Unternehmen. Neue Arbeits-und Führungskulturen im Zeitalter digitaler Wirtschaft, 129–134

Stoffel M (2015) Mitarbeiter führen Unternehmen–Demokratie und Agilität bei der Haufe-umantis AG. Das demokratische Unternehmen: Neue Arbeits-und Führungskulturen im Zeitalter digitaler Wirtschaft, pp 263–283

Strategy & Transformation Consulting (2018). Available at: http://www.strategy-transformation.com/digitale-transformation-verstehen/. Accessed 13 Mar 2019

ten Hompel M, Henke M (2017) Logistik 4.0–Ein Ausblick auf die Planung und das Management der zukünftigen Logistik vor dem Hintergrund der vierten industriellen Revolution. In Handbuch Industrie 4.0 Bd. 4. Springer Vieweg, Berlin, Heidelberg, pp 249–259

Digital Transformation of Companies

Heiko Kopf

Contents

1 Introduction

"Nothing is more constant than change." This quotation from the Greek philosopher Heraclitus accurately describes the environment which we live in today, both socially and professionally. However, due to digitalization, this change has taken on disruptive features that affect all areas of our lives. Especially the Internet plays a decisive role in this. The establishment of search engines such as Google has enabled new uses of this medium. The introduction of the smartphone has functionalized the Internet and thus forms the basis for completely new applications and business models. It is hard to believe that the first generation of smartphones was only brought on the market in 2007. In the last decade, we have become so dependent on these devices that it is hard to imagine what life would be like without them. As a result, digitalization has brought about profound changes in communication, which can be quickly visualized when looking for phone booths, for example. In addition to communication, new dependencies—but also opportunities—have emerged because of these technologies. Without claiming to be exhaustive, facts such as spontaneous

H. Kopf (✉)
Hamm-Lippstadt University of Applied Sciences, Hamm, Germany
e-mail: heiko.kopf@hshl.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_50

orientation in a foreign city, Wikipedia as a huge reference work, which the user always has at hand, and much more can be mentioned here. Smartphones have turned telephones into computers which, in addition to managing appointments, offer the possibility of taking photos at any time and, of course, of looking at old photos anytime and anywhere. Lastly, it should be mentioned that one can still use it to make phone calls, even if the younger generation in particular no longer considers this to be the essential aspect of a smartphone but prefers other ways of using the device for communication.

If one visualizes this technology and its effects alone, it quickly becomes clear that these changes inevitably have consequences for companies. They become obvious, for example, when looking at a city that is a tourist destination. Where just a few years ago people took photos of Trevi Fountain or other sights with their compact cameras, photos are now mainly taken with smartphones. It is easy to deduce what this means for the providers of compact cameras. If we stay with the subject of photos, we can observe a continuous decline in the number of branch-based photo shops. The background is varied and, in addition to the trend away from the compact camera described above, includes topics such as the increase in online purchases and thus new phenomena such as the theft of advice, that is, people getting free advice in specialty shops, while subsequently buying the product at a lower price on the Internet. In this respect, digitalization means much more than just the introduction of the Internet and the smartphone. It is rather a massive restructuring of social and individual behavior patterns, combined with the introduction of ever newer technologies and possibilities. This is happening at a rapid pace and with a complexity that society has never experienced before.

If one now takes the perspective of the company management, it quickly becomes clear that action is necessary, because non-action can lead to a reduction in competitiveness or even to a matter of economic viability. Even the size of a company does not provide any protection either, if one considers, for example, that the large mail-order company Quelle (established in 1927) was highly specialized in a market that is now booming in a digital format and is dominated by a start-up founded in 1994, amazon.com, Inc. If one looks at Quelle's corporate history, one will realize quickly that the company did not recognize the signs of the times and that it addressed the topic of digitalization too late. Yet there is also a different approach: one of Quelle's direct competitors, the Otto mail-order company, was quick to address digitalization strategies and is now the second largest online shop in Germany (€ 2.956 billion €), second only to amazon.com, Inc. (€ 8.816 billion) (Langenberg 2018).

Companies should therefore increasingly question the influence that digitalization can have on their own activities, or even better, how opportunities opened up by digitalization can be used for one's own company to provide for new unique selling propositions and/or serve new markets. However, it is like in school days: When you are shown the solution, it seems trivial and simple. The approach rather means to become aware of one's own company and the industries it supplies and then to identify suitable digitalization topics to find one's own way. Unfortunately, there are no universal solutions.

2 Digital Transformation As an Entrepreneurial Task

The term digital transformation is currently not yet sufficiently defined and standardized. For a detailed account of currently common definitions, please refer to existing literature (Schallmo et al. 2017, pp. 3–5). In the following, this term is understood as a process that a company (or an organization) pursues to actively incorporate the changes in environmental influences (e.g., from society, the customer, the market, or other elements) induced by digitalization into its own operational and strategic management.

Digital transformation is an actively designed transition of a company from its current status quo to a company that uses meaningful elements of digitalization in its processes, its business model, its products, its services or in its relationship management. Thus, digital transformation is an actively designed process that companies must actively initiate and shape themselves. Once this process has been initiated, it quickly becomes clear that it is a continuous task for the company, as the business environment is also constantly adapting. In large-scale enterprises, this continuity is often reflected in the fact that new board positions or organizational units close to the management board are established to deal with this issue. For this reason, the so-called Chief Digital Officer (CDO) has been in place for some time at the top management level of such companies, who is responsible for this process.

In the context of digital transformation, the primary concern is not the quantity of elements used, but rather the meaningfulness and interaction of these elements. If one looks at companies such as amazon.com, Inc., for example, the most diverse elements are orchestrated in a special way and systemically interconnected. This should not be the standard for all companies, but it reflects the actual task very well. This means that as part of this process, all activities and divisions of the company should be put to the test, right down to the "holy grail" of the company, the actual business model. This scope clearly illustrates the degree to which the company is affected. It is therefore not, as in the context of industrialization, about the workers in the production process and how they are affected, but about everyone involved in the company, from the production workers to the management. Sometimes it is also necessary to slightly leave the "German way" of perfection and to surrender to the venture of a new way. This does not mean that in the future products or services should be produced in a lower quality and that one surrenders to trash, but rather that one increases the willingness in the company to take risks and exposes oneself to the danger of failure to increase the potential of innovation. At this point, it is worth mentioning a quote from Robert Sutton of Stanford Business School: "As soon as companies try to reduce the number of flops, the innovation process usually comes to a standstill. The key to efficient innovation is to fail faster, not less frequently." (Sutton 2018) For many companies, the digital change in the world therefore also means a complex change in the company, ranging from the introduction of new technologies to business models and issues of innovation management to corporate culture. Touching on many issues during the process leads to a complexity in companies that often makes a structured start very difficult.

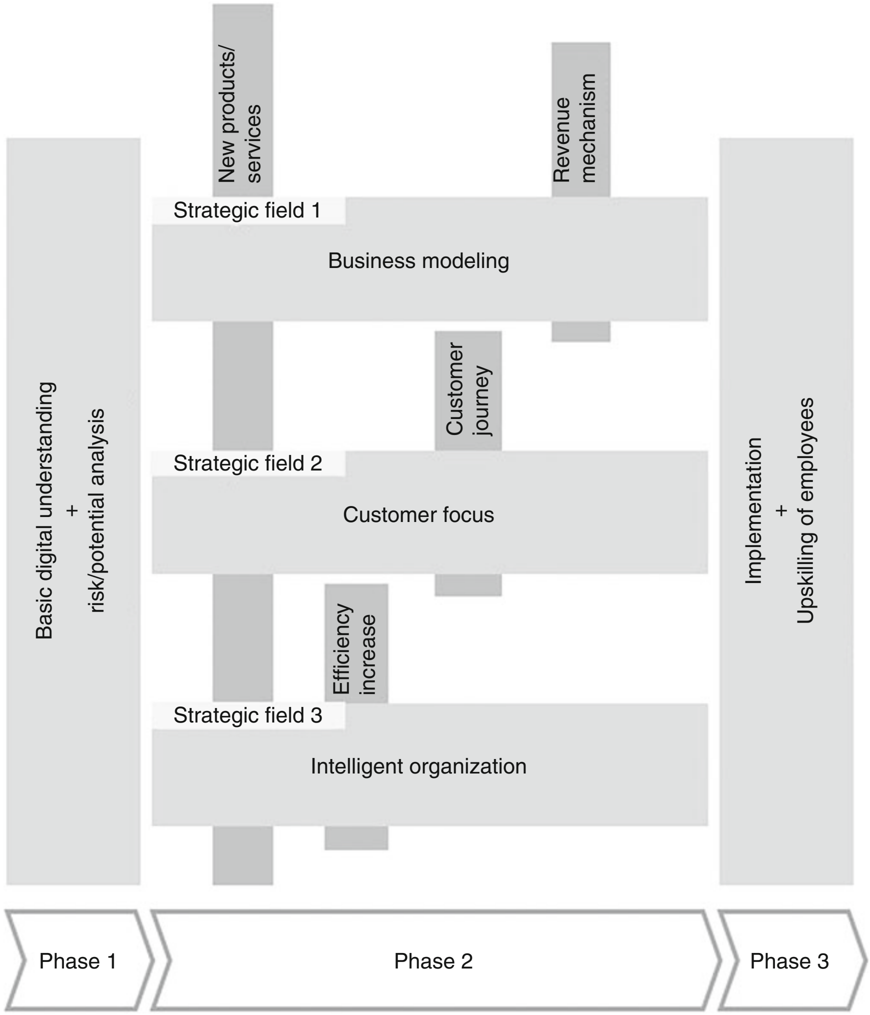


Fig. 1 Illustration of the digital transformation process

There are possibilities to segment the process so that a structured approach can be implemented. Figure 1 shows a possible three-phase procedure. The individual phases (designated Phase 1, Phase 2, and Phase 3 in the lower part of the diagram) and their respective measures are explained below.

Phase 1:
Since digital transformation is a both organizational and content-related process, the decision-makers who oversee this process and its structuring must also be competent to make decisions. To this end, the company, and in particular the people who structure the process, should acquire the necessary knowledge in Phase 1 to be able

to accompany the transformations with decision-making competence. For this purpose, the company should first acquire a basic digital understanding so that it can assess, albeit to a limited extent, the facts and interrelationships that occur in subsequent stages. This basic understanding is further applied in this phase when it is necessary to carry out a risk and potential analysis, often with external support. Comparable to the approach used in strategic corporate development, an analysis of the company's current situation with regard to its internal strengths and weaknesses, but also external threats and potential, takes center stage at this point—of course, this is done with a special focus on how these behave against the background of digitalization and related market changes and, obviously, what new influences affect the company (see Chapter 3).

In addition to gaining knowledge and raising the awareness of those involved, the aim of this phase is to take stock of the situation in a structured way to provide a basis for further action.

Phase 2:

Building on the first phase, three strategic fields should then be addressed, either individually or in combination:

- business modeling
- customer focus
- intelligent organization

Digitalization has at least influenced many ***business models*** (strategic field 1), if not created disruptive, innovative business models that have cannibalized the established ones. It is therefore necessary to look at one's own company with this focus and to reconsider the application of innovative business model mechanisms, such as revenue mechanisms.

A company must react to the massive changes in the customer's behavior, fears and needs and recognizes the "new" characteristics of its customer. The company should then optimize its ***customer focus*** (strategic field 2). Here, the customer journey, for example, plays an increasingly important role.

The ***company organization*** (strategic field 3) can and should be analyzed and evaluated against the background of digitalization. Looking at individual processes, means of digitalization can be used to increase efficiency, which can quickly lead to advantages on the cost side. Likewise, the use of new technologies can create new value added or reorganize the existing creation of value.

In addition to measures in the individual strategic fields, it can and will always happen that several strategic fields have to be considered for certain ideas. A new product and/or a new service (see Fig. 1) is given as an example. In such a case, implications would arise for the business model, the customer focus but also for the intelligent organization, which must be identified and also dealt with. The examples mentioned in the figure (new products, revenue mechanism, customer journey and efficiency increase) are to be understood as examples and must, of course, be worked out specifically for the company.

At the end of this phase, concrete objectives emerge which the company wants to achieve. In addition to this, measures are already being developed in this phase on how these objectives are to be achieved.

Phase 3:
To achieve the chosen objectives, the measures must now be operationalized and implemented in this phase. Since new interfaces may possibly arise, it is advisable to define future competences and responsibilities at this point. Furthermore, the employees concerned must be upskilled. Primarily, the awareness and understanding should be conveyed that these changes are necessary and important for the company’s competitiveness and are therefore also being dealt with in the interest of the workforce. In addition to this, competence-oriented upskilling must, of course, also take place so that the employees are able to understand their changed role and fill it both in terms of content and organization.

In the following chapters, individual topics, which are to be assigned to the respective phases, will be explained in detail. Of course, due to the complexity already mentioned, this cannot be done exhaustively in this framework.

3 Phase 1: Key Factors Influencing Companies

Digitalization can influence entrepreneurial activities through a wide range of channels. This influence can be partly indirect but also direct.

Social changes that can be induced by digitalization represent an indirect influencing factor. It is undisputed that the way in which society communicates has changed (for example, communication via social media). According to the latest ARD/ZDF online study (Koch et al. 2018), adults aged 14 years and older use the Internet 196 min per day, with 63 million people in Germany having Internet access. For companies, this results in new opportunities for reaching customers, but it also poses new risks. Obviously, digitalization is ascribing everyday significance to the smartphone, which must absolutely be taken into account in business considerations. The smartphone in its multifunctional applicability thus influences the behavior of individuals and thus of entire societies. In addition to the multiplication of dissemination speed of information, it has also led to it taking on increased functions that were previously performed in a different form. In addition to permanent connectivity (for communication and information procurement), it also opens up the possibility of remotely controlling devices (e.g., cameras) or carrying out procedures (e.g., banking) independent of one’s location, right through to transparency at any time (e.g., comparison of prices). This changes patterns of behavior, which companies need to consider. A need for corrective measures can arise for companies if, for example, the communication needs of customers and business partners change massively, or consumer behavior is affected. Here, so-called B2C companies often feel the need for direct adjustment, while companies with B2B business models tend to notice an indirect effect via their commercial customers.

New potential but also risks for companies arise from the ***enhancement*** of existing and the ***development of new technologies***. For example, companies can use new virtual environments, such as virtual reality, to present their products to the customer in a very realistic way (e.g., a digital inspection of the new bathroom at the tiler), which can lead to considerable added value. However, other companies, not previously identified as competitors, may also use technology for a purpose that has given competitors a competitive edge. In recent years, technologies have been developed to an industrial standard, which makes it necessary to consider to what extent they have a direct or indirect influence on one's own company.

At the same time, digitalization also induces ***changes in value chains***. For example, the so-called Internet of Things, the digital availability of physical products and services at seemingly any location, can break up value chains in established industries and bring about new arrangements of the individual elements and thus also a "redistribution" of the creation of value. Today, price comparison platforms play a role in the so-called customer journey, which was previously mainly determined and designed by the company itself. As a final consequence, companies have to buy services from these platforms and possibly reduce their own resources at this point or make special efforts to be able to develop independently of these platforms.

A substantial part of the dynamics of digitalization is induced by the start-up scene. Increasingly, new companies with ***new business models*** and a tremendous agility are entering the markets and succeeding in influencing them. When looking at the age of the largest listed companies in terms of their market capitalization (Hunter et al. 2018, p. 39), it can be noted that among the companies with the highest absolute market capitalization between 2009 and 2018 the top five are IT companies (Apple, amazon.com, Inc. Alphabet, Microsoft and Tencent). These companies have an average age (Apple (1976), amazon.com, Inc. (1994), Alphabet (1998), Microsoft (1975), Tencent (1998)) of 30 years. In addition, the top positions in the listing of the best venture capital deals of all time are dominated by companies with an IT focus (software or hardware) (Insights 2018). This type of financing enables leaps in growth, which, on the one hand, allow investors to achieve corresponding returns and, on the other hand, provide serious competition to existing companies active in the sectors, as if out of nowhere. In Germany, this could be observed not least on September 24, 2018, when Wirecard AG (a financial services company founded in 1999 and listed on the stock exchange since 2005) was newly listed on the DAX, and in exchange, Commerzbank AG, a founding member of the DAX, was removed from the listing. However, due to scandal, trading in wirecard AG shares was discontinued on November 2021.

In addition to these factors, digitalization has caused ***long-known mechanisms*** to no longer function in the way they used to. For example, for many years the principle applied that the size of a company can provide protection from competitors. Today, the size of a company and the scale of the assets it holds and operates no longer necessarily imply a protective function (exceptions confirm the rule, such as electricity network operators). Rather, especially those companies are at an advantage, which demonstrate a high level of agility and can react very quickly to changes in the market thanks to this adaptability. There is another aspect to this: future competitors often do not come from one's own industry or the familiar regional surroundings. By transferring mechanisms from outside the branch to new branches, companies active in the market encounter procedures and actions that they have not previously known from their branch. This makes it more difficult to keep an eye on competitors, as new mechanisms appear along with new competitors in the market. This effect is even aggravated since many business models are geared towards internationality and can quickly act internationally due to the Internet. Thus, existing customers seem to suddenly be approached by companies that were previously not visible to the current supplier in the familiar market.

The aforementioned and other factors thus have an impact on one's own company development, which leads to an increasing complexity of the environment. Furthermore, these factors affect various corporate organizational units (marketing, production, etc.). For many company managers, the result is a world of feelings between a massive "crossing of fingers," backed up by the hope that the train of digitalization will pass by one's own competitive environment, and the impression that one is in a boat, which seems to have increased leaks and one is in danger of losing control of the overall situation. Added to this is the difficulty that digital transformation is metaphorically always referred to as an "open-heart surgery." This means that the company must continue to function and earn money, even if it is currently undergoing restructuring. In Phase 2, three fundamental causes of influence (strategic fields) are presented below, which companies can analyze in the context of their actions in a company-specific manner.

4 Phase 2: Business Modeling (Strategic Field 1)

Digitalization has made business models possible, some of which have never existed before, or, alternatively, established business models were used in a new digital format to create new added value from them.

The term *business model* was established about 20 years ago and describes the logical relationships of how a company and its organizational units create added value for the customer and what revenues it generates in the process.

For a better presentation, the following analyses use the business model typology based on Gassmann (Gassmann et al. 2013, p. 6) (Fig. 2).

As can be seen in the figure, there are four essential elements (target group, offer, creation of product/service and revenue mechanism) which revolve around and interact with customer value at the center. In this context, a unique selling proposition can be created either in a single element or in a combination of several elements. In the following, the individual elements are explained very briefly to illustrate the system context. For further details, please refer to the available literature (Gassmann et al. 2013, pp. 5–10).

The customer's purchase decision should be the central aspect in every company. The customer will decide taking into account several parameters, including competitors. The ***customer value*** plays a central role in this. For this reason, the company

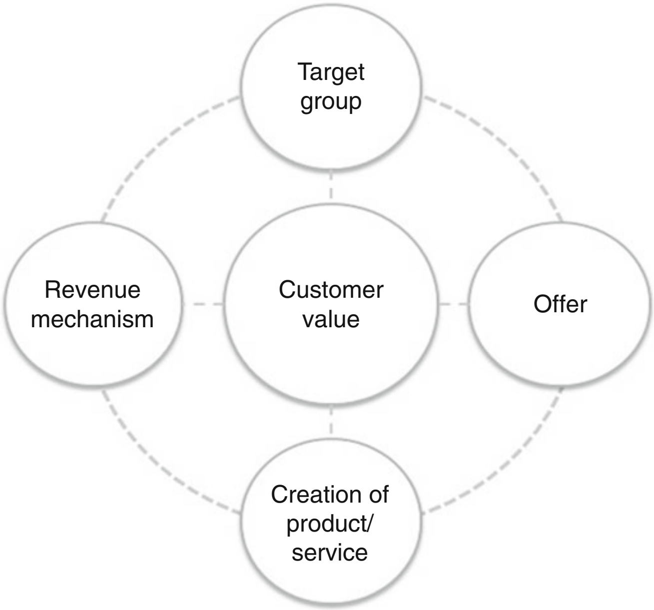


Fig. 2 Business model elements surrounding the central element of customer value

should approach the product/service from the customer's point of view. Here it is important to put oneself in the customer's shoes (What added value does the customer see in the product/service and/or what needs does it satisfy in his view). For example, it is advantageous to be able to solve an existing problem or satisfy a need of a specific customer group with the product/service. The customer value can include many different aspects—from a functional benefit to a gain in prestige that is expected from the product/service.

A good business model should address a clearly describable target group. This means that certain characteristics can be assigned to this ***target group***. Essentially, the first step is to distinguish between two different business models. On the one hand, business models with the company addressing end customers (consumers) as a target group. This is referred to as a B2C business model. On the other hand, other companies can also be the target customer. Such business models are designated by the acronym of B2B. The next step is to describe characteristics of the customer group. These can be of a completely different nature (e.g., in the case of end customers: demographic, behavioral, etc.).

The ***offer*** is the product or service that the company wants to provide. Within this element, the company must be clear about both the profile and the descriptiveness of the offer.

As regards the **creation of *product/service***, each company should have a clear idea of how it intends to provide its service or manufacture its product. To fulfil the value proposition made to the customer, processes and activities must be carried out.

Likewise, in this module one should also be aware of one's position in the existing (or newly generated) value chain. Taking these basic conditions into account, the proportion of value-adding processes that are performed within the company itself and the proportion that is to be supplied by third parties must be identified.

The essential issue of revenues is described within the element ***revenue mechanism***. A description of the company's so-called revenue mechanism is made. How are revenues generated with the products/services? Often the company has several revenue channels, such as product sales which are also associated with services.

By answering these five elements, the business model becomes tangible and discussable. Often this development takes place within the scope of a longer iterative process, as the formulation and modification of one element influences the other areas.

New business models particularly shaped by digitalization are characterized by the fact that they partly generate a new type of customer value, that they produce products/services in a new way or that they implement revenue mechanisms that could not be created before.

For example, e-commerce and m-commerce providers act in such a way that they manage to change the existing consumer behavior so that a ***new kind of value*** is created for the customer. Compared to the alternative of branch-based retailing, the greater value for the customer is that he does not have to consider shop opening hours, for example, and saves the effort of driving to the branch with the corresponding effort and expenses of driving or parking. In doing so, the customer is happy to accept constraints such as delayed delivery or not being able to inspect products directly.

New ways of product/service creation arise, for example, when the digital availability of products and services results in the transport of people being reorganized in a new way. The company UBER offers its customers transport by private persons. The customer can specify the place and time of pick-up, and the company organizes a vehicle (driven by a private person) to pick up the customer and take him to the specified destination. It quickly becomes apparent that this is actually the business model of a taxi company, which uses its own cars and employed drivers. The customer value is identical, but the way in which it is achieved is completely different, namely without having so-called assets (here own cars and employees) available.

Moreover, ***new revenue mechanisms*** can be implemented through digitalization. In particular, revenue models are emerging that aim at companies "pushing their way into" the so-called customer journey, i.e., the customer-specific shopping experience, and then generating revenues for their services. One example is the organization of price comparisons for customers. However, the service is not charged to the customer, but to the company that wants to sell its products/services to the customer.

The variety of these business models encouraged and facilitated by digitalization is large. However, a few business model mechanisms should be mentioned in brief, without claiming to prioritize any particular ones.

First, both from a current and prospective perspective, so-called ***data-driven business models*** should be mentioned. At their core, these business models use their own, acquired or known data and manage to turn this data into value in a new way. As a rule, mathematical formulas and methods, so-called algorithms, are used to extract desired information from the data. This makes it possible, for example, to make forecasts regarding potential borrowers as to whether they will pay amortization and interest rates in an exemplary manner. With the help of these forecasts, banks can assess their own risk. Analyzing the purchasing behavior of existing customers can be used to make purchase recommendations to new customers, as similarities in consumption can be identified. Likewise, such technologies can be used to make forecasts, identify congestion risks early on, and then direct traffic to avoid congestion.

Platform-based business models are so dominant that the term *platform economy* was derived. The basic principle of these business models is both simple and valuable. These companies operate platforms which, on the one side, bring together and also systematize a large number of providers with their products and services. On the other side of the platform are the customers who are looking for such products and services. The business model is now for the platform operator to bring the two groups together so that the customer purchases the product from a specific supplier. The platform operator receives a fee for brokering the purchase (e.g., private sales: ebay.de, search for craftsmen: myhammer.de). Data-driven modules often appear to complement this business model, as a platform is naturally supplied with a considerable amount of data from both sides (suppliers provide supply data and users provide demand profiles). In a next step, the platform operators can analyze this data and utilize it either for forecasts or targeted approaches.

If one now looks at the first two business models, one might think that digitalization only favors business models that focus on digital competence. However, it is striking that more hybrid business models are being developed. Here it is about upgrading physical products or conventional services with digital elements. These so-called digital ecosystems are becoming increasingly widespread. The idea behind this model is to use the digital elements to link the actual product, the actual service, with the user's smart device. The toothbrush manufacturer Braun, for example, offers an app for an electric toothbrush, with which user-specific toothbrushing habits (scope and quality) are entered into the user's calendar on his mobile phone. In addition, the user can even be actively prompted by his toothbrush—via a push message on his mobile phone—to clean his teeth. The ideas behind this system can be manifold. On the one hand, they serve customer loyalty by caring for and serving the customer in such a system so that the customer will continue to buy the right hardware in the future (in this specific example, revenue is generated on the hardware side). On the other hand, such systems can also be applied for new developments by using a data-driven element to learn more about the customer and be able to better judge him, to provide third parties with this information (in this case the revenue would be generated on the digital side) or to encourage the customer to continue consuming.

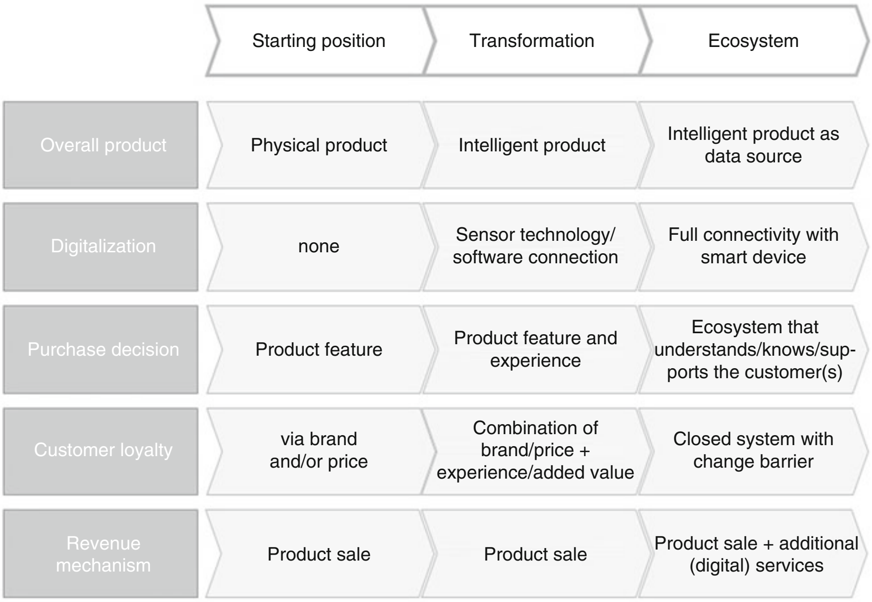


Fig. 3 Development phases of a digital ECO system

Example: Structure of Digital Ecosystem

Figure 3 roughly shows the phases towards a digital ecosystem. For simplification, the procedure is divided into a starting position and two digitalization stages. The heights of the steps used here do not correctly reflect reality, but are intended to provide a plausible representation of the procedure. As a notional example, a lawnmower is used here as a product.

Starting position: The company has a product on the market and offers it. The product does not yet have any digital features, and the purchase decision is made on the customer side based on the product quality (strategy of quality leadership) or the price (strategy of price leadership). If the company wants to keep the customer in the long term, the company's investment focus should be on either brand building, quality assurance or new production processes, depending on the primary purchase decision criterion. The company generates revenue through selling the physical product.

Lawnmower: The company sells a lawnmower that the user manually pushes through the garden.

Stage 1: Now, in a next step, the company begins to transform the product into a "smart product." This involves, for example, the integration of sensors into the product, which can be read out with the aid of a software platform. In addition to the integration of sensors, the company is also looking into the question of what data might be required and what information might interest the customer. This is intended to create a new experience on the customer side, which should lead to a positive purchase decision, and the company has created a distinguishing feature with regard to its overall product, which is marketable. This feature, the combination of product features and experience/added value, can be used to tie the customer more closely to the overall product.

Lawnmower: The lawnmower now has a sensor connection that allows the user to see what area he has mown, to what length the lawn has been cut and how often he has mowed. An app is available for this, which the user can download to his mobile phone.

Stage 2: Based on the experience gained in the earlier stage, further added values have been identified and integrated into the product. An increasing data-technical functionalization of the actual product is taking place, whereby the connectivity of the product is constantly increasing. As a result, connectivity and the customer experience are to be given higher priority in the purchase decision of the potential customer. The fact that the product is now being used to serve additional customer needs through additional features increases the customer's prospective dependence on the product. An ecosystem is created around the customer, which completely takes care of the user in certain areas. Such a system increases customer loyalty, since the customer feels a high barrier to switching, as by switching to another product he loses all the convenience that this system has provided him so far. In addition to this increase, other digital services (to the customer or to third parties) and products can be offered through such a system, thus creating new revenue opportunities.

Lawnmower: The lawnmower is now fed with other available data (e.g., weather data, lawn lengths before mowing, fertilization data, verticutting activities, etc.). This enables it to give explicit instructions on lawn care (notification: "Fertilization is necessary."). In combination with the company's own fertilizer, the amount of fertilizer could also be entered into the system using a code (additional revenues from selling another product). Likewise, the smart lawn thatcher of the company's own brand would automatically supply the necessary data. The lawnmower would inform the user when it is best to mow the lawn based on the user's calendar, the expected length of the lawn and weather data. Of course, the whole spectrum could be rounded off by autonomous devices.

By choosing the lawnmower as an example, a B2C business model has been used. To the same extent, however, this procedure can also be used for B2B business models, as additional features can also be integrated for corporate customers. An example of this is the maintenance of equipment, which can be developed from a routine reminder to predictive maintenance (without downtime) appropriate to the use.

To create such accompanying digital systems, companies must, of course, also contribute competences and skills to the development of products/services. To do this, the company must either develop itself or cooperate with competent partners.

Another recurring pattern is represented by business models that make individual adjustments for the customer. In literature, this is referred to as ***mass customization*** business models (Piller 2006). The basic principle here is that the customer (often the end customer) has the possibility to adapt and customize the product or services according to his special preferences. For such customization to take place, the offering company either configures its products from modules, or it is specific components of the product that can be enhanced with personal images or graphics. The modular application is used, for example, for shelf or cabinet systems which can be put together from individual modules using a configurator on the Internet (for example: Schrankwerk.de). The business model is also reflected in the personalization of food products, for example chocolate dragées, which can be adorned with personal images (e.g., mymms.de) and thus be given a personal touch.

On top of this, the current application of digital business models includes many long-standing business mechanisms that worked in the past and have now been converted into digital form or reinterpreted. There are elements such as the reinterpretation of old discount stamps (groupon.de), everlasting estate agent models (immobilienscout24.de) and digital dating agencies (parship.de). For further information on the influence of digital transformation on business models, please refer to existing literature (Gassmann and Sutter 2016, pp. 15–27).

5 Phase 2: New Customer Needs and Customer Focus (Strategic Field 2)

Apart from the emergence of new business models or reinterpretation of old ones, it is above all the customer who behaves differently compared to the past. This is a particular challenge especially for companies that operate a direct end customer business.

Of course, one cannot speak of the typecast customer, but when companies become aware of the identikit image of their "primary customer," they should always consider on which experience basis they assess the customer. In particular, it is all about the fact that the customer has been strengthened in certain characteristics through digitalization and the business models that have been created and modified as a result. Some of these aspects will be explained and presented here by using examples.

One of the main changes on the customer side is related to the customer's ***communication needs and behavior***. Whereas in the past, companies essentially communicated in the direction of the customer and the customer was the recipient of the message, this principle has not only softened, but now appears to be almost reverse. This means that the actual communication power has increasingly shifted from the company to the customer. For example, there are business models that see one of their core competencies in sorting and presenting customer opinions (tripadvisor.de). Likewise, social media inspire customers to express their opinions in a publicly perceptible way. These comments are so authentic for other potential customers that they often, without knowing the person and the reason, attach great importance to these comments when making their own decisions. Approximately two thirds of all German citizens no longer make a purchase decision without first studying customer ratings or experience reports (Miosga and Tropf 2017).

Consumption behavior has also changed. In addition to the increasing dominance of Internet purchases, the aim of consumption has also changed. While until a few years ago the customer saw himself primarily in a product society and his aim of consumption was rather oriented towards possession, his focus is now on the benefit rather than the actual possession. This has been particularly noticeable in the automotive industry in recent years, as many customers no longer attach the same importance to owning a car as they used to but are now looking for mobility solutions. Social trends are further intensifying this development, such as the need to make life as eventful and exciting as possible.

People's ***personal behavior*** has changed beyond consumer behavior. Because of permanent connectivity, the customer also expects prompt offers that are available at the place where he is. Moreover, the customer is less and less willing to act with patience. The mere issue of the provider's response time to customer queries can be a decisive factor in purchase decisions. In addition to this kind of impatience, there is also a sharpening of behavior that does not seem consistent at first glance. The consumer becomes increasingly lazy, thus behaving almost anti-proportionally to the provider. In addition to availability at all times and in all places, the consumer also expects a high degree of agility and willingness to solve problems from the optimal supplier.

In addition to these factors mentioned above, other general but also target-group-specific points can be identified. In the context of digital transformation, the company is faced with the task of identifying the customer's relevant characteristics. It is often an advantage to generate an "identikit picture" of the customer, which describes the targeted customer and his characteristics, behavior and needs. Based on this profile, two actions particularly should now be undertaken:

The ***centering of the business model*** on this customer should be reconstructed once again. Do the products/services that the company produces serve the customer optimally, or are there characteristics and needs that are not yet adequately satisfied? Is it perhaps possible to cater even more optimally to the interests already being addressed? Does the product or the customer value need to be focused even more strongly on certain aspects? Here, questions and interrelations of the business model might have to be discussed again. An example of the optimization of a business model with regard to a customer characteristic is the *Dash button* of amazon.com, Inc. The customer's need to place orders comfortably and without much effort was identified. The *Dash button* now targets exactly this characteristic. It is a product-specific button which, when operated, connects to the Internet, and orders the product depicted on the button from amazon.com, Inc. This means that when the customer is standing in front of the washing machine and the washing powder is almost used up, he no longer has to bother either to write a shopping list or to place an order via his computer or smartphone. This is done fully automatically by pressing the button. In return, the customer (possibly unconsciously) forgoes a price comparison.

The second central issue is the customer-oriented design of the so-called ***customer journey***. The term customer journey, which was coined in the marketing context, describes the customer's "journey" to his purchase decision and includes the phase of owning the product. This journey represents a cycle consisting of different stages. Various structural descriptions of this can be found in the literature, although it is sufficient at this point to mention that there are at least three phases:

- pre-purchase phase
- purchase phase
- post-purchase phase

In each of these phases, the customer should come into contact with the company that offers its products/services on the market. For each phase of the customer journey there are specific points of contact, the so-called touchpoints. On the one hand, there are physical touchpoints, such as radio, printed advertisements, and call centers, but on the other hand, there are also digital touchpoints such as e-mails, social media, chats, and apps. Specific touchpoints can be assigned to each phase of the customer journey. It is important to use the right tools for the right customer in the right phase of his customer journey. Through the development and commercialization of new technologies, this customer journey can also be extended to include other aspects such as virtual realities.

6 Phase 2: Intelligent Organization (Strategic Field 3)

The company organization has had to react to the increasing digitalization in various ways. By company organization we mean in this context all the necessary processes and structures that are necessary for commercial value creation. This includes, for example, the financing of the company, the property rights situation, or human resources development. Virtually none of the company's organizational units is or remains unaffected by digitalization.

For example, a situation arises in most companies that seems to be almost impossible to control with conventional methods. On the one hand, it is the fact that the company has to process a large amount of data per period (e.g., e-mails), but at the same time there is the requirement to shorten reaction times. From the management's point of view, this is a considerable balancing act: decisions must be made quickly, with a high need for discussion and, at the same time, increasingly complex processes, and in parallel, the company should also "reinvent itself" and be innovative to be able to secure its competitiveness in the future.

Especially "digital companies" such as Alphabet Inc. or amazon.com, Inc. seem to have found ways out of the dilemma, at least with regard to some points. For example, it is known that Google, at least initially, applied the well-known 20-% rule for its employees. This rule said that employees were allowed to use 20% of their working time (although critics claim that this rule is actually paradoxical, as employees do not record their working hours in the traditional sense) for things that Google could benefit from in future, but which are not part of their actual responsibilities (Page and Brin 2004). This should strengthen the ***innovation culture*** and employee satisfaction. Insiders report that Google products such as Adsense, Google News or Gmail have emerged from these activities. With increasing growth and the concentration on the restructuring of the company towards Alphabet Inc. this regulation faded from the spotlight until rumors started to spread that it might no longer be practiced. In 2016, Google set up an internal start-up incubator, Area 120, to help further structure these 20% activities. Employees can now apply for admission to the incubator with a business plan.

With regard to amazon.com, Inc., however, a different operational regulation is known, which Jeff Bezos partly explained in a market letter (Bezos 1997). For ***internal meetings***, for example, there is a stipulation that no PowerPoint slides or other presentation techniques may be used. Here the presentations should consist of six-page explanations presenting the agenda item. Meetings therefore begin with a joint, silent reading session during which the participants read through the submitted documents before discussing them. He thus ensures, on the one hand, that the participants are not influenced by presentation marketing and, on the other hand, that all participants have actually read the documents before discussing them. In addition, the burden of proof is reversed, which means that it is not the innovator who has to defend his idea, but the critics who have to explain and substantiate why the idea will not work.

Changes also occur due to the influence of technology. So-called Industry 4.0, which is described and discussed in detail elsewhere in this book, opens up new opportunities for commercial value creation.

In addition to this, randomly selected technologies should be mentioned which, through their use, can introduce new business models or original approaches and advantages into operational processes or create new types of customer relationships.

The availability of existing ***amounts of data*** and their precise analysis allows, for example, ***forecasts*** to be made from current data, which are the basis for optimization. These forecasts can be used to optimize delivery routes in terms of time, effort, or other characteristics. Various company goals can be given priority, such as internal cost reduction or service optimization with regard to waiting times for customers. Data may also be used to make ***forecasts***. This is basically not a new invention resulting from digitalization, since every ice-cream vendor in the past already controlled his production depending on the weather, based on experience. With the help of digital tools and existing data, it is now possible to draw conclusions from the noise generation of production pumps in refineries as to when these pumps are likely to fail. A failure would mean a production downtime, while an unnecessary and too early replacement would increase costs, so an early needs-based replacement is the ideal solution, which can be implemented with the help of these data analyses. A further application field of data could be the preparation of ***offers*** based on ***user characteristics***. For example, if an insurance company gets exact information about a customer's personal driving profile, the insurance company could assign this behavior to a risk class and then make a corresponding offer that reflects either the customer's willingness or aversion to take risks. Of course, a wide variety of other applications are also possible (Gentsch 2018, pp. 117–220).

The use of virtual technologies (virtual reality or augmented reality) can lead to different scenarios. Within the framework of ***virtual reality (VR)***, a virtual space is created in which the user can move and interact with the aid of control devices (e.g., joysticks) and an imaging element (mainly VR goggles). This space can now be assigned to different purposes. For example, complex training content can be conveyed in this space and damage-free interaction can be simulated. Likewise, several users can stay in a virtual space and interact with each other. Complex operational planning, for example of internal logistics, can also be tried and tested here before it is physically built and installed. In this way, costs due to bad planning can be avoided and first handling instructions can be given and trained in the virtual space parallel to the construction phase.

Augmented reality (AR), on the other hand, is a technology that allows the user to create a mix between digital content and reality using a smart device (such as a tablet or smartphone). With the help of the integrated camera, a real image is captured, which can then be enriched with digital content. Volkswagen, for example, has taken the first steps towards providing digital support for vehicle maintenance. For example, maintenance procedures are stored for the XL1 vehicle, which are then digitally superimposed on the real photo, so that a service technician can immediately see which inspection point he has to attend to within the scope of the desired check routine and where he can find it in reality.

Companies can also obtain new qualities of data by using ***drones***. The company doks.innovation GmbH, for example, has developed a technology with the help of which drones move fully independently in a warehouse and systematically record stored goods. This provides the company with up-to-date inventory data. As a result, the company can achieve high-quality processes that are based on the updated inventory being the planning basis.

In addition to the technologies already mentioned, and of course others that are not mentioned here, ***additive manufacturing***, also known as 3D printing, will play a key role in the future. Since there is a large number of technology platforms within this technology area, each of which differs in the type of materials used and in the implementation of the actual printing process, the range of applications for this technology is difficult to estimate (Ehrenberg-Silies et al. 2015, pp. 25–27). It ranges, for example, from the printing of prototypes or spare parts to the printing of semi-finished tools used for assembly purposes, through to industrial applications, since it is more economical, for example, to work with such a process when batch sizes are small.

7 Phase 3: Employee Upskilling

As shown in Fig. 1, the third phase is to include implementation measures and employee upskilling. In the following, particular attention will be paid to employee upskilling.

The employees play a, often *the*, decisive role in the transformation process and the subsequent implementation. For this reason, the company's workforce should always be considered part of the transformation process. However, it is recommended that such broad consideration is only made the focus of attention in the third phase, as otherwise the issue could possibly become a continuous obstacle in the deliberations (for example, through questions such as: Are we actually able to do this?). Depending on the corporate culture, it may also make sense to involve selected employees in the developments as early as the second phase. After the transformation strategy has been worked out in the earlier phase, two central activities must now take place with regard to the employees.

Since digital transformation is a special change project, the approach here should be one that has already proved successful in such projects. Employees want to know what to expect. Therefore, the responsible persons should approach the employees and inform them about the planned project. This should result in gaining the staff's understanding and the associated willingness to cooperate and, in addition, in keeping anxieties to a minimum. The next step would then be a competence-oriented development of the staff, so that they have the necessary competencies and skills to understand new processes and procedures and subsequently to be able to contribute in a value-creating way. In addition to this, management staff must also be developed in a method-oriented manner with regard to their new tasks. This is necessary not least against the background that the digital transformation of the company may also involve changes may require new management philosophies. On top of this, there are technical or new methodological skills that will be needed in the future. Thanks to the new competences at the staff and management level, the appropriate measures and activities can be developed that are necessary to bring the digital transformation to fruition (Disselkamp 2018, pp. 95–171). Of course, it is important here to find the right balance, appropriate to the corporate culture, between a top-down approach (the management levels define measures and activities) and a bottom-up approach (proposals for measures and activities are made by the employees).

References

Bezos JP (1997) Letter to shareholders. [Online] 1997. [Cited: 10 23, 2018.] https://www.sec.gov/Archives/edgar/data/1018724/000119312518121161/d456916dex991.htm

Disselkamp M (2018) Digital-Transformation-Management: Den digitalen Wandel erfolgreich umsetzen. Schäfer Poeschel, Stuttgart, p 2018

Ehrenberg-Silies S et al (2015) Additive Fertigungsverfahren/3D-Druck: Anwendungen und Potenziale. VDI/VDE-IT, Berlin, 2015. Study

Gassmann O, Sutter P (2016) Digitale Transformation im Unternehmen gestalten: Geschäftsmodelle, Erfolgsfaktoren, Handlungsanweisungen, Fallstudien. Carl Hanser, München, p 2016

Gassmann O, Frankenberger K, Csik M (2013) Geschäftsmodelle entwickeln: 55 innovative Konzepte mit dem St. Galler Business Model Navigator. Carl Hanser, München, p 2013

Gentsch P (2018) Künstliche Intelligenz für Sales, Marketing und Service: Mit AI und BOTS zu einem Algoritmic Business - Konzepte, Technologien und Best Prctices. Springer Gabler, Wiesbaden, p 2018

Hunter R et al (2018) Global Top 100 companies by market capitalisation. PriceWaterhouse Coopers International, London, p 2018

Insights CB (2018) From Alibaba to Zynga: 28 of the Best VC Bets of All Time and What We Can Learn from Them. [Online] 2018. [Cited: 12 4, 2018.] https://www.cbinsights.com/research/best-venture-capital-investments/.

Koch W et al (2018) ARD/ZDF-Onlinestudie. [Online] 2018. [Cited: 12 2, 2018.] http://www.ard-zdf-onlinestudie.de/

Langenberg C (2018) EHI Retail Institute. [Online] 2018. [Cited: 11 28, 2018.] https://www.ehi.org/de/top-100-umsatzstaerkste-onlineshops-in-deutschland/

Miosga J, Tropf TM (2017) bitkom e.V. Kundenbewertungen sind wichtige Kaufhilfe. [Online] 2017. [Cited: 10 28, 2018.] https://www.bitkom.org/Presse/Presseinformation/Kundenbewertungen-sind-wichtigste-Kaufhilfe.html

Page L, Brin S (2004) Founders' IPO Letter. [Online] 2004. [Cited: 12 4, 2018.] https://abc.xyz/investor/founders-letters/2004/ipo-letter.html

Piller T (2006) Mass Customization: Ein wettbewerbsstrategisches Konzept im Informationszeitalter. Universitäts-Verlag, Wiesbaden, p 2006

Schallmo D et al (2017) Digitale Transformation von Geschäftsmodellen: Grundlagen, Instrumente und Best Practices. Springer Gabler, Wiesbaden, p 2017

Sutton RI (2018) Der Querdenker-Faktor: Mit unkonventionellen Ideen zum Erfolg. Piper, München, p 2018

Framework for Managing Business Transformation

Gerhard Gudergan, Boris Alexander Feige, Denis Krechting, and Ruben Conrad

Contents

1 Challenges of Business Transformation

To stay competitive, the central challenge for many companies at present is to master the process of transformation in the sense of a fundamental redesign of central processes or even of the entire company. Digitization and the need to redefine and reposition oneself in a sustainable economy are just two examples of the main drivers of this transformation. In this context, the basic question of the fundamental necessity of a comprehensive transformation in the sense of a business transformation no longer even arises for most companies in the age of digital transformation. Instead, the focus is on the question of how companies can manage the complexity

G. Gudergan (✉) · B. A. Feige · D. Krechting · R. Conrad
FIR - Forschungsinstitut für Rationalisierung e.V., Aachen, Germany
e-mail: gerhard.gudergan@fir.rwth-aachen.de; info@fir.rwth-aachen.de; Denis.Krechting@fir.rwth-aachen.de; Ruben.Conrad@fir.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_51

associated with the scope of a transformation and the necessary changes in terms of the business strategy and on how to ensure the efficiency and success of the transformation. The challenge is to simultaneously design new structures and systems on the substantive level and also to break up established patterns of behavior. Müller-Stewens already summarizes the two essential tasks of a transformation into the two groups of strategy work and change work (cf. Müller-Stewens and Lechner 1999). Even if in the current discussion especially the structural aspects made possible by digital technologies or even the technological solution itself are in the foreground of the discussions about digital transformation, the change work itself is of immense importance. A transformation is therefore a highly demanding and complex process for all stakeholders. The successful management of this process requires a combination of different methods, which, in addition to strategy development and business model innovation as well as organizational design and methods of organizational realignment, also include instruments for managing the transformation process in particular (cf. Uhl and Gollenia 2012). These can be used to contribute to the adoption and establishment of new behavioral patterns. The behaviors of employees is of particular importance in transformation projects, as the difficulties of a corporate transformation are essentially characterized by uncertainties, conflicts or a lack of orientation for employees, which leads to resistance and inertia. With targeted management through selected instruments, these resistances can be counteracted, and motivation and commitment can be achieved among employees, which is an essential prerequisite for the success of a transformation (cf. Uhl and Gollenia 2012; Uhl and Ward 2012; Siepmann 2012; Mollbach and Bergstein 2012; Schuh 2006).

This article provides concrete added value for overcoming the challenges presented by providing a framework for dealing with the management tasks associated with a transformation. Furthermore, it provides orientation for the managers responsible for a transformation. It is based on extensive application experience and has meanwhile been tested in its practical use together with more than 90 experts and managers. This framework is complemented by an overview of the methods available for managing transformations and was also developed with experts and enriched by their expertise.

2 Definition of Business Transformation

A transformation describes a fundamental change in a company (cf. Lahrmann et al. 2013). Accordingly, a business transformation is best described as the realignment of all strategic and operational work of a company and the repositioning of key corporate activities (cf. Hanssen and Klaffke 2004). With concrete reference to business activity, business transformation is considered a “transformation, renewal and thus radical change in the business activity of an organization” (Schneider 2011). In principle, transformation affects employees, processes, and functions of an organization (cf. Schmid et al. 2013), which must be aligned and brought into line through proper transformation management (cf. Winter et al. 2012).

Business transformation must be distinguished from change management, whereby opinions differ widely on the definition and thus also the delimitation of change management (cf. Zimmermann 2015; Doppler 2004; Alisch 2010; Aldrich and Ruef 2006). Change management can be defined as the "ongoing adaptation of corporate strategies and structures to changing framework conditions" (Alisch 2010) and is therefore a change that is not accompanied by radical restructuring, which distinguishes it from corporate and business transformation. Thus, a transformation is always a change, but not all changes are also transformations (cf. Aldrich and Ruef 2006).

A further distinction is made between first-order and second-order change. While first-order change takes place in the form of small changes in the already existing frame of reference of a company, second-order change is described by the change in business logic as a significant change (cf. Schuh 2006) and thus corresponds to the definition of business transformation.

3 Framework for the Management of Business Transformation

3.1 Conceptual Foundation

One objective of this paper is to provide a management framework for dealing with the management tasks associated with a transformation and thus to offer orientation for the responsible transformation managers. The complexity associated with the management of transformations can thus be better mastered. The St. Gallen Management Concept represents a holistic management approach and, in the version presented by Knut Bleicher, offers a general and comprehensive dimensioning approach for management activities. It distinguishes between the three management levels of normative, strategic, and operational management, whose contributions to corporate development are divided respectively into activities, structures, and behavior (see Fig. 1). This approach makes it possible to bring strategic tasks into a meaningful order or to give them direction. Business transformation in the sense of a conscious change represents such a strategic task, as this change requires a fundamental reorientation of the company and, due to its long-term character, has a great influence on corporate policy as well as internal processes (cf. Bleicher 2004). Management processes encompass all fundamental tasks related to the design, control, and development of companies in the sense of purpose-oriented socio-technical systems (cf. Ulrich 1970). Management processes are divided into three levels. On the normative level, the focus is particularly on the ethical legitimization of entrepreneurial activities. This means that the company develops its own moral value system, taking into account the prevailing social value orientations. At the

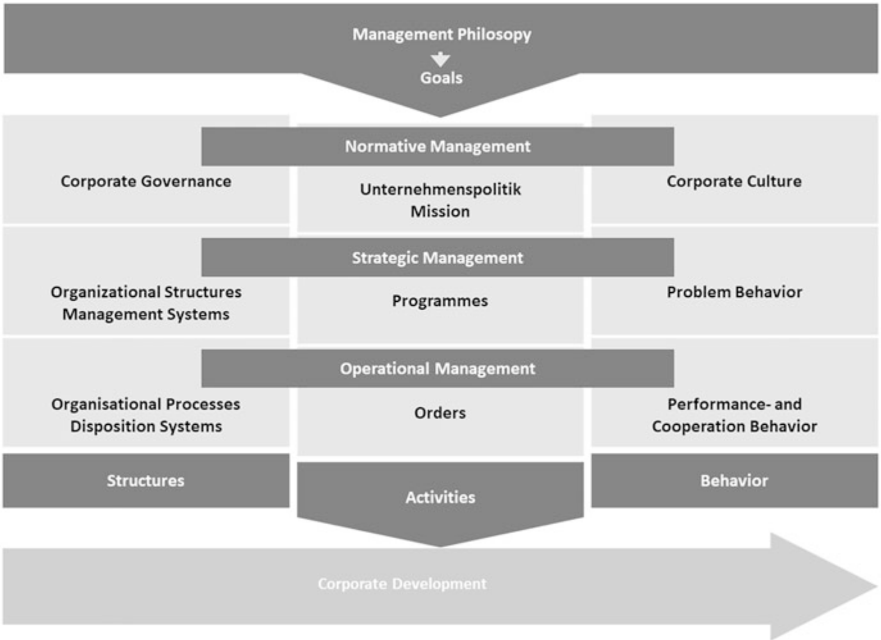


Fig. 1 Normative, strategic, and operative management in the St. Gallen management concept (own illustration based on Bleicher 2004)

strategic level, the focus is on securing a company's competitive, long-term future. In this strategic orientation, market signals and competition-relevant trends in the corporate environment in particular are included. Management processes at the operational level relate to tasks of immediate management of everyday business and in particular to efficiency in dealing with scarce resources (cf. Rüegg-Stürm 2003). The management concept according to Bleicher (cf. Bleicher 2004) serves as a superordinate structuring and explanatory pattern for the framework for the management of business transformation presented below.

The St. Gallen Management Model developed by Rüegg-Stürm build upon Bleicher's concept. Like the new St. Gallen Management Model developed by Rüegg-Stürm, the framework for the management of business transformation presented in this article understands a company as a system of processes. According to the St. Gallen Management Model, the processes of a company comprise the active and supporting value creation activities in the form of business processes and support processes as well as the management processes necessary for this, which are also referred to as management processes (cf. Rüegg-Stürm 2003).

3.2 Framework

In analogy to the new St. Gallen Management Model (cf. Rüegg-Stürm 2003), an external and an internal perspective are distinguished. The external perspective includes the categories environmental spheres, stakeholder groups and interaction topics that relate to the social and natural environment. The internal perspective includes the categories of structuring forces and resources as well as IT systems, processes and development modes, which refer to the internal view of the organization (Fig. 2).

3.2.1 External Factors

Environmental spheres can be understood as central contexts of a company with which a company constantly interacts. The environmental spheres have therefore to be analyzed in terms of their changes. Society, characterized by an increasing trend of digital transformation, represents the most comprehensive of these spheres.

Stakeholders have to be understood as organized or non-organized groups of people or institutions that are affected by the entrepreneurial value creation processes. The value contribution for the stakeholder groups is what establishes the purpose of a company in the first place. The interaction topics comprise the "objects" of the exchange relationships between the stakeholder groups and the company, around which the communication between the company and the stakeholder groups revolves. These are norms and values as well as concerns and interests. Values refer to fundamental views about a desirable life, norms build on them and refer to explicit laws and regulations. Interests refer to immediate self-interest, whereas concerns refer to general goals.

3.2.2 Internal Elements

As mentioned above, the framework for the management of business transformation considers a company as a system of processes in analogy to the St. Gallen Management Model according to Rüegg-Stürm 1998.

Management processes in themselves comprise all fundamental tasks that focus on the design, control, and development of companies in the sense of purpose-oriented socio-technical systems (cf. Ulrich 1970).

The areas of responsibility examined in more detail below detail the internal perspective shown in Fig. 3 and present it in the form of a structuring framework.

The Business Transformation Canvas is based on the chosen perspective of the St. Gallen Management Model, presents the management tasks and processes required for a business transformation in a systematic context and divides them into three areas. The first sub-area builds on the so-called development modes. In this

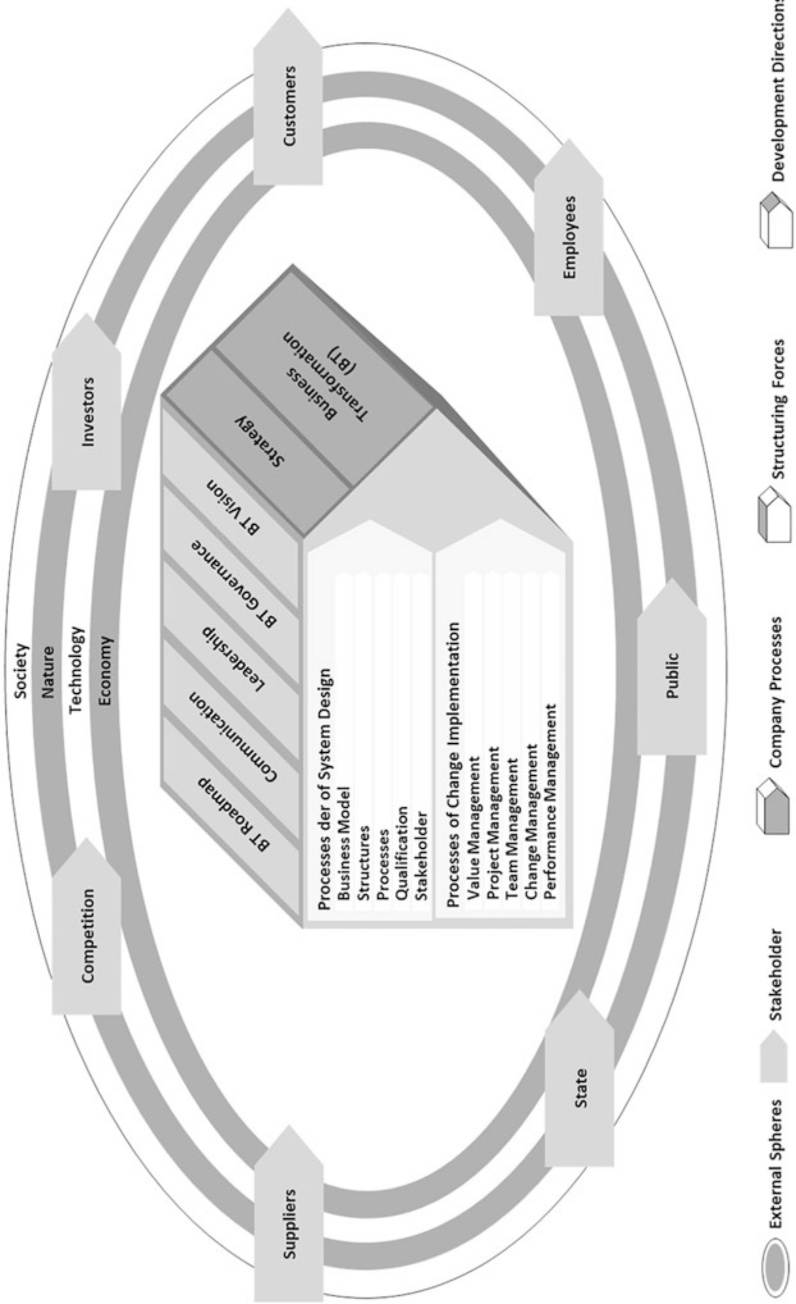


Fig. 2 Overarching structure of the regulatory framework for the management of business transformation (own illustration)

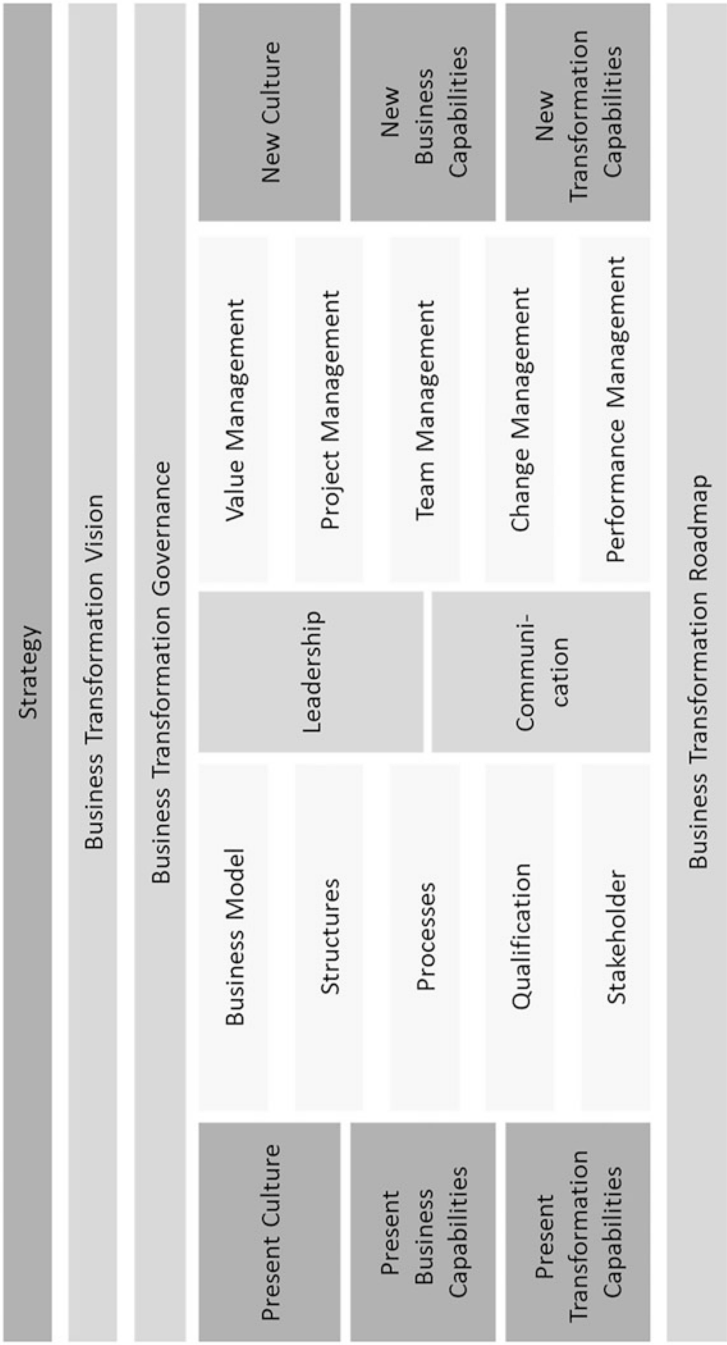


Fig. 3 Framework: Business Transformation Canvas (own illustration)

case, the development mode is the corporate strategy as well as the business transformation mode.

The second sub-area comprises the corporate processes necessary for business transformation, which, following Müller-Stewens and Lechner, can in turn be divided into the processes of designing future systems and processes of change in the narrower sense (Cf. Müller-Stewens and Lechner 1999).

The corporate processes that deal with the design of future systems include the development of business models, organizational structures, processes as well as the development of the necessary personnel qualifications and the future structure of roles and responsibilities. The approach presented here explicitly assumes that transformations develop out of and are derived from strategic considerations. Accordingly, the further sub-areas are designed based on a strategic decision and objective. The design and detailing of the business model is of particular importance in this context, because the business model integrates very diverse areas of design. A business model illustrates, for example, the design of the value creation and thus the company or provider perspective with which a customer benefit is generated. It also describes partner structures, core processes, sales channels, customer segments as well as the revenue logic and thus the way in which revenues are generated in the business model (cf. Osterwalder and Pigneur 2011). Through the concretization of the strategic goals and the model-like implementation, the business model can be considered as a link between strategy and business plan (cf. Müller-Stewens and Lechner 2005). In all areas of design, the digital transformation results in countless new forms and alternatives of realization. In addition, completely new value creation patterns are emerging, for example in the form of platforms (cf. Brauckmann 2015).

The business processes that are concerned with bringing about the change itself include the tasks of integrating and motivating roles and stakeholders, the tasks of project management, team management and change management, as well as the establishment of performance management for the transformation process.

The third sub-area comprises the holistic orientation and steering of the business transformation, as well as its structuring. This sub-area combines the development of an overarching vision, the establishment of a governance structure for the business transformation as well as the development and adoption of a comprehensive roadmap for the transformation by the management. These formative moments are held together by clear leadership and goal-oriented and consistent communication. In the new St. Gallen Management Model, these tasks are aggregated under the term structuring forces.

The three sub-areas listed are accompanied in the representation of the Business Transformation Canvas by the description of the current state of the company based on the currently existing competences and the current corporate culture on the left side of the representation. The right-hand side shows the competences required in the future and the desired corporate culture in a defined target state after the business transformation. This is because a prerequisite for the functioning of modern organizations is the guarantee of conditions for collectively uniform action. This is of paramount importance especially for a change of the company in the sense of the digital transformation. Collective convictions, values and norms help to bring about the necessary integration (cf. Schreyögg 1989).

4 Managing Business Transformation

As described above, complexity management in particular is of central importance in the management of business transformation. The framework of the Business Transformation Canvas serves to manage complexity in the management of a transformation. Some of the tasks have an explicit steering effect in the context of transformation (moments of order). These will be discussed in more detail below.

4.1 Managing by Vision

"The vision mentally advances a conceivable situation that could occur or be brought about in the future" (Bleicher 2011) and it is the origin of entrepreneurial activity (cf. ibid.). The vision is most often described with the word "image of the future" (cf. Menzenbach 2012; Bleicher 2011; Hahn 1994). At the beginning of a transformation, both a strategy and a vision are initially formulated (cf. Stiles and Uhl 2012). Especially against the background of the digital transformation, a vision must be described as clearly as possible, even if technological developments and possibilities are not yet foreseeable in detail (cf. Keicher and Bohn 2012; Uhl and Ward 2012). Ultimately, the vision in transformation projects is a picture of the future state of the later, transformed organization (cf. Jones and Recardo 2013). The vision answers questions such as "Where do we want the company to go, where will we be in three, five or seven years?" (Gassmann et al. 2013).

Kotter (1995) already made the importance of a vision clear 20 years ago in his much-cited work "Leading Change: Why Transformation Efforts Fail" in the eight steps for transforming organizations: "Creating a Vision," "Communicating the Vision," and "Empowering Others to Act on the Vision" are, according to him, three of the eight steps for successful transformation that are in the context of vision. Many other authors emphasize the relevance of a vision in transformation projects (cf. Uhl et al. 2013; Vom Brocke et al. 2013; Keicher and Bohn 2012; Goldbeck 2012; Stiles and Uhl 2012; Schuh 2006). In the 2012 Capgemini study, a clear vision is among the top three success factors of change processes between 2003 and 2012 (cf. Keicher and Bohn 2012). Lauer (2010) emphasizes that a company that wants to transform itself must know where it is heading. Finally, goals are derived from the vision and the vision is the basis for employee understanding, as it gives them orientation and helps them to align their own activities with the big picture. Thus, the vision serves the commitment of the employees and generally the conviction of stakeholders (cf. Uhl and Ward 2012; Goldbeck 2012).

4.2 Managing by Governance

As in all highly developed professions, managers also need standards of craftsmanship to be able to react professionally to the rapidly increasing complexity of all systems in the areas of society, economy, and politics (Malik 2008). One of these standards managers can use is the introduction and maintenance of a governance structure. This aims to introduce a combination of mechanisms to satisfy different stakeholders (shareholders, employees, financiers, suppliers, and others) (Goergen et al. 2005). The system works by clearly defining the rules and procedures for decision-making and by making activities, guidelines and chosen methods transparent (Lin 2011; Tricker 2009).

In view of the fact that a business transformation is an extremely complex project that affects almost all areas of a company, it is advisable to achieve the necessary transparency through a governance structure. In the best case, different management levels work towards a common goal whose control and coordination is clearly defined and described. The transformation manager's role is to act as a link between a network of individual functions and goals and the management board. The particular challenge here is to provide the necessary information for all sides.

In this context, the specification and communication of clear goals in transformation projects are of great importance (cf. Zimmermann 2015; Uhl and Hanslik 2012; Mascarenhas 2011). One reason for this is that tangible goals help to maintain employee motivation over a long-lasting transformation, since successes can be recorded step by step as each goal is reached (cf. Kotter 2006).

Furthermore, goals are the “basis of performance and reward” (Gassmann et al. 2013) and thus an integral part of a transformation. The formulation of goals gives employees orientation without restricting them in the way the goal is to be achieved. Especially in transformation projects, clear communication of goals is particularly important so that employees understand the meaning behind the changes, and it can be ensured that the goals are not only understood but also accepted (cf. Doppler and Lauterburg 2014). Only those who understand the goal are able to play a supportive role (cf. Cameron and Green 2012).

4.3 Managing by Leadership

Leadership is the “orientation of the actions of individuals and groups, mediated by interaction, towards the realization of predetermined goals” and involves “asymmetrical social relationships of superiority and subordination” (Alisch 2010). Leadership thus aims at the orientation of action and is thus one of the essential instruments of management (cf. Ochugudu and Ayatse 2013). The terms manager and leader differentiate two different leadership personalities (cf. Hettl 2013). A manager is described as extrinsically motivated, as a rational problem solver with little emotional attachment, who focuses on work processes and derives self-confidence from fulfilling role expectations. A leader, on the other hand, is described as intrinsically motivated, as an intuitive problem solver with high emotional attachment, who focuses on work content and acts with self-confidence regardless of role expectation (cf. Menzenbach 2012).

On leadership behavior in transformation projects, a distinction is made analogously between transactional and transformational leadership (cf. Uhl et al. 2013; Krüger 2012). Krüger (1994) calls the role of the supervisor with transactional leadership "efficient manager." The role of the manager with transformational leadership behavior "visionary leader."

Leadership behavior is of great importance for a successful digital transformation (cf. Uhl and Gollenia 2013; Lauer 2010), but there is no optimal leadership behavior (cf. von Kyaw and Claßen 2010; Olfert 2010). This is due to the fact that on the one hand the employees to be led have different character traits, each of whom prefers different leadership styles (cf. Franken 2010). Another essential point is the phase of transformation, which requires different leadership behavior. In the initial phase of a transformation, different resistances must be overcome by leadership than in a more advanced phase (cf. Bruch et al. 2012). In addition, the general increase in complexity due to the constantly and ever more rapidly changing framework conditions makes proper leadership more difficult (cf. Welk 2015); an example of this is currently the influence of the so-called "Generation Y," whose characteristics pose new challenges for leadership, as they are reluctant to commit themselves and are skeptical of hierarchies (cf. Welk 2015).

Winter et al. (2012) and Doppler and Lauterburg (2014) emphasize that social skills and emotional intelligence are much more important for transformation than IT know-how or analytical skills. These and other important prerequisites for good leadership in the Digital Transformation can be derived from the Capgemini 2012 study. More than half of the respondents agreed that a good leader must have the following skills in view of the current challenges: act as a role model, actively communicate as well as actively initiate and shape the changes, make clear decisions, and convince employees that the change holds positive benefits (cf. Keicher and Bohn 2012; von Kyaw and Claßen 2010). In addition, managers should be aware that their personal attitude towards change has an influence on staff attitudes and the success of the project (cf. Keicher and Bohn 2012).

In fact, leadership or leadership behavior is one of the key success factors in transformation projects (cf. Azhari et al. 2014; Uhl and Gollenia 2013; Gassmann et al. 2013): "Leadership can make – or break – a change initiative" (Jones and Recardo 2013). Azhari et al. (2014) emphasize that the change process cannot be delegated away from the leadership level, as transformation is not an isolated process. Instead, transformation requires leadership and new leadership tools. The key role that managers still play in transformation processes therefore places high demands on leaders (cf. Keicher and Bohn 2012). Top management has a particularly important role to play here, as goals cannot be met without their commitment (cf. Uhl and Hanslik 2012).

4.4 Managing by Communication

"Honesty is the best strategy" (Schuh 2006). Uhl and Gollenia express similar views on the role of communication in the digital transformation (2013). "Communication is not everything, but without communication everything is nothing"—this is how Lips comments on the topic of communication (Lips 2012). The three quotes indicate the importance of communication. In the change management study by Kienbaum (2011–2012), 98% of the top managers, executives and project leaders of change projects surveyed answered that communication is of high importance for the success of a change project in general and without specific reference to digital transformation (cf. Mollbach and Bergstein 2012). Tozer (2012) metaphorically describes communication as a lubricant for the drive of leadership (leadership engine).

Against the backdrop of the uncertainty currently associated with the discussion about digital transformation, the importance of communication is not rated so highly for no reason. Employees resist corporate transformations because they are uncertain and disoriented, afraid of new things (cf. Jones and Recardo 2013), or fear loss (cf. Gassmann et al. 2013).

Resistance in transformation processes can be reduced through an effective communication program (cf. Brown and Harvey 2006). Communication increases the motivation of employees, provided that the transformation project is communicated early and in detail with the employees (cf. Siepmann 2012). In summary, it is therefore not surprising that Uhl and Ward (2012) come to the following conclusion in their book on digital transformation: "A common lesson from many of the cases – even the successful ones – is that no amount of communication is ever enough!"

4.5 Managing by Roadmapping

Business transformation encompasses the restructuring of essential business activities. This is associated with a high degree of complexity, which must be actively planned and monitored by top management or the responsible transformation managers (Ackermann 2001). The roadmap serves as an important instrument for planning, implementing and controlling a successful business transformation. A structured roadmap can be used to structure complex issues in terms of time and logic (Cf. Phaal and Muller 2009). On the other hand, it provides the possibility of visualization and thus the basis for a common, implicitly articulated understanding of top management about the individual activities during the change process (Phaal et al. 2013). Specifically, a roadmap provides benefits in three different dimensions, which are briefly outlined below:

– The roadmap serves as a **planning and control** instrument for monitoring the planned objectives per time unit. Individual activities and milestones can be

presented transparently in terms of their respective input and result and subsequently evaluated.
- **Parallel or sequential activities** are shown and temporal as well as factual logical sequences of actions can be coordinated.
- **Necessary resources** can be assessed and efficiently allocated depending on the activity in terms of time, costs and qualifications.

The implementation of the business transformation based on the roadmap is preceded by its development. This so-called road mapping process involves different actors in the company and cannot be developed by top management or the respective transformation manager alone. Rather, it is important to use the necessary professional qualifications of the respective functional units and to integrate them as needed. For example, the company's communication department must be involved in the planning of communication measures for the content and timing of area and individual communication. The respective transformation manager or top management is responsible for the road mapping process. It must logically bring together the respective content-related and time-related planning strands and aggregate them into an overall roadmap. Typical mistakes that occur time and again include a homogeneous team composition, a lack of transparency about the interdependencies of individual activities and strands of action, and overly ambitious time planning.

References

Ackermann L (2001) The change leader's roadmap: how to navigate your organization's transformation practicing. John Wiley & Sons

Aldrich H, Ruef M (2006) Organizations evolving. SAGE, London

Alisch K (2010) Gabler Wirtschafts-Lexikon. Gabler, Wiesbaden

Azhari P, Faraby N, Rossmann A et al (2014) Digital Transformation Report. Hrsg.: Neuland GmbH & Co. KG

Bleicher K (2004) Das Konzept Integriertes Management - Visionen - Missionen - Programme. In: St. Galler Management-Konzept Band 1. 7., überarbeitete und erweiterte Aufl. Campus, Frankfurt am Main

Bleicher K (2011) Das Konzept Integriertes Management. Visionen - Missionen - Programme. Campus, Frankfurt am Main

Brauckmann O (2015) Neue Formen der Wertschöpfung. In: Brauchmann O (Hrsg.) Smart Production. Wertschöpfung durch Geschäftsmodelle. Springer, Berlin u. Heidelberg, pp 5–11

Brown DR, Harvey D (2006) An experimential approach to organization development. Pearson Education, New Jersey

Bruch H, Krummaker S, Vogel B (2012) (Hrsg) Leadership - best practices und trends. Gabler, Wiesbaden

Cameron E, Green M (2012) Making sense of change management. A complete guide to the models tools and techniques of organizational change. Kogan Page, London

Doppler K (2004) Managing change successfully. core questions, issues, and strategies. In: Boonstra JJ (ed) Dynamics of organizational change and learning. Wiley handbooks in the psychology of management in organizations. Wiley, Chichester, pp 115–132

Doppler K, Lauterburg C (2014) Change Management: Den Unternehmenswandel gestalten. Campus, Frankfurt am Main

Franken S (2010) Verhaltensorientierte Führung. Lernen und Diversity in Unternehmen. Gabler, Wiesbaden, Handeln
Gassmann O, Frankenberger K, Csik M (2013) Geschäftsmodelle Entwickeln. 55 innovative Konzepte mit dem St. Galler Business Model Navigator. Hanser, München
Goergen M, Martynova M, Renneboog L (2005) Corporate governance convergence: evidence from takeover regulation reforms in Europe. Oxford Rev Econ Policy 21:243–268
Goldbeck M (2012) Zielbild in der Integration - Prozess der Operationalisierung als Grundlage für das kulturelle Zusammenwachsen. In: Große Peclum K-H, Krebber M, Lips R (eds) Erfolgreiches Change Management in der Post Merger Integration. Fallstudie Commerzbank AG. Gabler, Wiesbaden, pp 145–160
Hahn D (1994) Untemehmungsziele im Wandel. In: Gomez P, Hahn D, Müller-Stewens G, Wunderer R (eds) Unternehmerischer Wandel. Konzepte zur organisatorischen Erneuerung; Knut Bleicher zum 65. Geburtstag. Gabler, Wiesbaden, pp 59–83
Hanssen D, Klaffke M (2004) Unternehmenstransformation. Roland Berger. http://www.rolandberger.com/media/pdf/rb_press/RB_feature_corporate_transformation_e_20040820.pdf. Stand: 21.05.2015
Hettl M (2013) Mitarbeiterfhrung mit dem LEAD-Navigator. Erfolgreich und wirksam führen. Gabler, Wiesbaden
Jones DJ, Recardo RJ (2013) Leading and implementing business change management. Making change stick in the contemporary organization. Routledge, London
Keicher I, Bohn U (2012) Digitale Revolution. Ist Change Management mutig genug für die Zukunft? Hrsg.: Capgemini Consulting 2012
Kotter JP (1995) Leading change. Why transformation efforts fail. Harv Bus Rev:59–67
Kotter JP (2006) Leading change. Why transformation efforts fail. Harv Bus Rev:1–10
Krüger W (1994) Transformations-Management. Grundlagen, Strategien, Anforderungen. In: Gomez P, Hahn D, Müller-Stewens G, Wunderer R (eds) Unternehmerischer Wandel. Konzepte zur organisatorischen Erneuerung; Knut Bleicher zum 65. Geburtstag. Gabler, Wiesbaden, pp 199–228
Krüger W (2012) Führungsstile für erfolgreichen Wandel. In: Bruch H, Krummaker S, Vogel B (eds) Leadership - Best Practices und Trends. Gabler, Wiesbaden, pp 99–114
Lahrmann G, Winter R, Uhl A (2013) Transformation management survey: current state of development and potential of transformation management in practice. In: Uhl A, Gollenia LA (eds) Business transformation essentials. Case studies and articles. Gower, Farnham, Surrey, England, pp 5–15
Lauer T (2010) Change Management. Grundlagen und Erfolgsfaktoren. Springer, Heidelberg
Lin TCW (2011) The corporate governance of iconic executives. Notre Dame Law Rev 87:351
Lips R (2012) Kommunikation in Change Prozessen. Möglichkeiten und Grenzen. In: Große Peclum K-H, Krebber M, Lips R (eds) Erfolgreiches Change Management in der Post Merger Integration. Fallstudie Commerzbank AG. Gabler, Wiesbaden, pp 268–285
Malik F (2008) Die richtige Corporate Governance: Mit wirksamer Unternehmensaufsicht Komplexität meistern. Campus Verlag, Frankfurt/Main
Mascarenhas O (2011) Business transformation strategies. The strategic leader as innovation manager. Response Books, Thousand Oaks
Menzenbach J (2012) Visionäre Unternehmensführung. Grundlagen, Erfolgsfaktoren, Perspektiven. Springer Gabler, Wiesbaden
Mollbach A, Bergstein J (2012) Change. Points of view. Change-Management-Studie 2011–2012. Kienbaum Management Consultants 2012
Müller-Stewens G, Lechner C (1999) Die Gestaltung unternehmerischer Einheiten: Der General Management Navigator TM als ein Konzept zur integrierten Strategie- und Wandelarbeit; Die Organisationsentwicklung, 18 Jahrgang, Nr. 2, S. 24–43
Müller-Stewens G, Lechner C (2005) Strategisches Management - wie strategische Initiativen zum Wandel führen; der St. Galler General Management Navigator - 3., aktualis. Aufl. Schäffer-Poeschel, Stuttgart

Ochugudu AI, Ayatse FA (2013) Leadership styles as recipes for transformation. Int J Bus Manage Rev 1(4):35–53
Olfert K (2010) Personalwirtschaft. Kiehl, Herne
Osterwalder A, Pigneur Y (2011) Business model generation: Ein Handbuch für Visionäre, Spielveränderer und Herausforderer. Campus Verlag
Phaal R, Muller G (2009) An architectural framework for roadmapping: towards visual strategy. Technol Forecast Soc Change 76:39–49
Phaal R, Farrukh C, Probert D (2013) Technology management and roadmapping at the firm level. In: Phaal R, Moehrle M, Isenmann M (Hrsg) Technology roadmapping for strategy and innovation. Charting the route to success. Springer, Berlin u. Heidelberg, S. 13–29
Rüegg-Stürm J (2003) Das neue St. Galler Management-Modell - Grundkategorien einer integrierten Managementlehre. 2. Aufl., Bern [u.a.]
Schmid A, Uhl A, Vom Brocke J (2013) Beyond project management. The challenges of managing large-scale, complex transformation programs. 360 Degrees Bus Transf J 8:1–77
Schneider R (2011) Business Transformation. Im Spiegel von Megatrends und Schlüsselkompetenzen, Brugg. Hrsg.: Stier, M. Wirtschaftsmagazin
Schreyögg G (1989) Zu den problematischen Konsequenzen starker Unternehmenskulturen. Zeitschrift für betriebswirtschaftliche Forschung 41(2):94–114
Schuh G (2006) Change management. Prozesse strategiekonform gestalten. Springer, Berlin
Siepmann J (2012) Führungskräftekommunikation. Ein Mix aus Information und Dialog. In: Große Peclum K-H, Krebber M, Lips R (Hrsg.) Erfolgreiches Change Management in der Post Merger Integration. Fallstudie Commerzbank AG. Gabler, Wiesbaden, S. 325–346
Stiles P, Uhl A (2012) Meta management. Connecting the parts of business transformation. 360 Deg Bus Transf J 3:24–29
Tozer J (2012) Leading through teaders. Driving strategy, execution and change. Kogan Page, London
Tricker B (2009) Essentials for Board Directors: An A to Z Guide. 2. Ausgabe. John Wiley & Sons
Uhl A, Gollenia LA (2012) A handbook of business transformation management methodology. Gower
Uhl A, Gollenia LA (2013) Business transformation management methodology. Gower, Farnham Surrey, UK
Uhl A, Hanslik O (2012) PRO3 at Allianz. A new dimension of customer centricity. 360 Deg Bus Transf J 5:51–61
Uhl A, Ward J (2012) Success and failure in transformation. Lessons from 13 case studies. 360 Deg Bus Transf J 3:30–37
Uhl A, Gollenia LA, Muench U (2013) The Transformational CIO. In: Uhl A, Gollenia LA (eds) Business Transformation Essentials. Case Studies and Articles. Gower, Farnham, Surrey, England, pp 27–37
Ulrich H (1970) Die Unternehmung als produktives soziales System. 2 Aufl. Paul Haupt Verlag, Bern
Vom Brocke J, Patry M, Schmiedel T (2013) How Hilti masters transformation. In: Uhl A, Gollenia LA (Hrsg.) Business transformation essentials. Case studies and articles. Gower, Farnham, Surrey, England, S. 203–213
von Kyaw F, Claßen, M (2010) Change Management Studie 2010. Business Transformation - Veränderungen erfolgreich gestalten. Capgemini Consulting, München. Hrsg
Welk S (2015) Die Bedeutung von Führung für die Bindung von Mitarbeitern. Ein Vergleich unterschiedlicher Führungsstile im Kontext der Generation Y. Springer Fachmedien Wiesbaden GmbH, Wiesbaden
Winter R, Townson S, Uhl A et al (2012) Enterprise architecture and transformation. The differences and the synergy potential of enterprise architecture and business transformation, vol 5. 36 Deg Bus Transf J, pp 23–31
Zimmermann G (2015) Erfolgreiche Gestaltung von Veränderungen. In: Zimmermann G (ed) Change Management in Versicherungsunternehmen. Die Zukunft der Assekuranz erfolgreich gestalten. Springer Gabler, Wiesbaden, pp 61–93

Work 4.0: Human-Centered Work Design in the Digital Age

Susanne Mütze-Niewöhner, Christina Mayer, Markus Harlacher, Niklas Steireif, and Verena Nitsch

Contents

S. Mütze-Niewöhner (✉) · C. Mayer · M. Harlacher · N. Steireif · V. Nitsch
Institute of Industrial Engineering and Ergonomics, RWTH Aachen University, Aachen, Germany
e-mail: s.muetze@iaw.rwth-aachen.de; c.mayer@iaw.rwth-aachen.de; m.harlacher@iaw.rwth-aachen.de; n.steireif@iaw.rwth-aachen.de; v.nitsch@iaw.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_52

1 Perspectives on Work in the Digital Age[1]

Where will we work? How, when, and with whom? Which jobs will change and in what way? Which jobs will be substituted and which will be eliminated without replacement? Will we still work at all in the future—in the sense of spending physical or mental effort to perform a service for which we receive payment in exchange—or will this be done by intelligent super-algorithms, robotic systems or artificially grown brains that control global production or added value systems?

Progress reports from technological research and development as well as numerous research activities in the fields of "Internet of Production," information and communication technologies, robotics, and artificial intelligence (AI) fuel the expectation that the vision of largely digitalized and automated product and service creation will become a reality in the medium term, at least in some sectors and areas of work. In the long term, automation, in particular that which is driven by AI technologies, might affect all aspects of work in some shape or form (see Bostrom 2014; Daheim and Wintermann 2016).

In addition, the studies cited below concerning the status and development prospects of work do not allow a reliable forecast of whether the corresponding expectations will actually be fulfilled, but they do support the thesis that people will still be significantly involved in the provision, design and control of work in the coming decades. Various models of work and job design approaches are currently being discussed and researched under buzzwords such as the "digitalized working world," "New Work" and "**Work 4.0**" (an analogy to "Industry 4.0," a term which refers to the fourth industrial revolution that is supposedly triggered by the widespread use of interconnected cyber-physical systems). These focus more on the person at work, the division of labor between people and technology and the organization of work. As we are still in digital transformation, the subject of "Work 4.0" cannot be fully grasped and research into the effects on our world of work and the people who work in it is by no means complete. This article can hardly provide definitive answers; however, it can pick up and discuss some perspectives surrounding development and design aspects that are particularly relevant for shaping the future of work from science human-centered perspective, especially in the context of "Industry 4.0."

1.1 Digitalization from the Perspective of German Employees

Digital technologies are already omnipresent in the world of work: According to a recent survey, more than 80% of employees use information or communication technologies at work, such as computers, the Internet, laptops, tablets or

[1]This chapter was written in the context of the project "TransWork" (FKZ: 02L15A162), funded by the Federal Ministry of Education and Research of Germany.

smartphones (BMAS 2016). As expected, the study cited shows clear differences between the occupational groups considered, with particularly high proportions among people who work in the field of business organization (99%) or provide knowledge-intensive services. However, it also backs up the fact that even for employees working in occupations that can be assigned to production technology (86%) or manufacturing (67%), the use of hardware and software is already part of everyday working life (BMAS 2016).

Comparable results concerning the **spread of digital technologies** are also provided by the DGB index “Gute Arbeit” from 2016 (Holler 2017): Only 18% of the employees surveyed stated that they had not been affected by digitalization in any way so far. In terms of sectors, penetration is particularly high in the information and communication sector (93%), followed by the “provision of freelance, technical and scientific services,” financial and insurance services and public administration (all between 80 and 89%). However, high proportions of digital work are also found in vehicle manufacturing (76%) and mechanical engineering (75%). The index also asked about the forms that the **digital work** took, whereby multiple answers were permitted. For employees in the manufacturing sector, the following picture emerged (Institut DGB-Index Gute Arbeit 2016): electronic communication via email, smartphone, social networks etc. (63%), work with supporting electronic devices, e.g., scanners, data glasses or diagnostic devices (54%), software-controlled work processes, e.g., route, production or scheduling planning (53%), work with computer-controlled machines or robots (43%) and Internet-based, distributed cooperative project work (33%). Holler (2017) found an overall concentration of digital work among higher-skilled workers.

Both studies point to the **risks and potentials** of technological innovations: As such, in the study commissioned by the Federal Ministry of Labour and Social Affairs (BMAS), 56% of the users who were surveyed perceived an increase in their own productivity, 32% perceived greater freedom of choice in the organization of their own work and 65% perceived a consolidation of work. In addition, more than half of the employees with a higher level of education complained about an amount of information that was difficult to manage due to modern means of communication (e-mail, mobile phone, Internet), while half of the respondents with a lower level of education reported at least a reduction of physical stress due to new technologies in the workplace. Almost 80% of the respondents saw a need to constantly develop their skills as a result of technological innovations. Overall, however, the work relief was perceived to be weaker than the experienced increase in demands (BMAS 2016). On subjective productivity gains following the introduction of new technologies, the results of the latest wave of this survey are also sobering: with the help of the newly introduced technologies, only around 46% of the employees surveyed stated that they are able to perform their tasks faster, around 39% reported to achieve better results and only 39% found their work less monotonous (BMAS 2020a).

The DGB Index 2016 also pointed to the risks of digital work, such as increased time-related pressures, more frequent disruptions and interruptions, increased work consolidation and an increase in the monitoring and control functions that were experienced (Institut DGB-Index Gute Arbeit 2016; Holler 2017).

1.2 Digitalization from the Perspective of German Companies

While from the employees' point of view work is already clearly permeated with digital technologies, not only in the service sector but also in manufacturing companies, surveys at company level concerning "**Industry 4.0 readiness**" show a heterogeneous picture: The spectrum ranges from companies that have not yet addressed the use of digital technologies and companies in which digital transformation is at least visionary or strategic in character to a few companies that are already using numerous digitalization technologies, especially in customer-facing areas such as marketing and sales (Arntz et al. 2016a; Lichtblau et al. 2015; Lerch et al. 2017; Zimmermann 2017). In a representative survey of companies on the topic of the "World of Work 4.0" conducted by the Institute for Employment Research (IAB) together with the Centre for European Economic Research (ZEW) in 2016, 63.5% of the production companies which were surveyed stated that they did not yet use modern digital technologies (Arntz et al. 2016a, 2017). However, the survey specifically asked about cyber-physical systems, smart factories, the Internet of Things, big data, cloud computing, online platforms and shop systems.

In their study on the status concerning **digitalization in production**, Lerch et al. (2017) also found a relatively low prevalence of digital technologies. In terms of readiness for Industry 4.0, the SMEs which were surveyed proved to be (still) more reticent than the large companies. In companies where digital technologies are already being used, this did not lead to a reduction in staff. With increasing Industry 4.0 readiness, there was significantly higher value creation per employee. The authors therefore emphasize the potential of digital technologies to generate productivity increases through more effective and efficient work processes (Lerch et al. 2017).

The different views of employees and companies concerning the status of digitalization in the manufacturing sector can certainly be attributed in large part to the fact that the status quo perceived by employees (which essentially describes the use of (isolated) computer-supported or software-supported work tools) is to be found at the lowest levels in the maturity models that have been defined in relation to Industry 4.0 (e.g., Jodlbauer et al. 2016; Lichtblau et al. 2015; Reuter et al. 2016). At present, it must be assumed that the operational reality in manufacturing companies, especially in small and medium-sized enterprises (SMEs), is still far from the vision of a so-called smart factory, in which highly automated, data-technically horizontally and vertically networked, communication-capable production systems (so-called "cyber-physical systems") manufacture smart, customer-specific products in varying quantities in intelligent, self-controlling and self-optimizing processes (concerning the visions of Industry 4.0, see e.g., Dispan 2017; Jacobs et al. 2017; Kagermann et al. 2013; Lichtblau et al. 2015; Reuter et al. 2016; Spath et al. 2013a, b; Vogel-Heuser et al. 2018; Wischmann et al. 2015).

This assumption is not only supported by the studies on Industry 4.0 readiness cited above, but also by economic studies that classify the relatively stable reluctance of SMEs to invest in innovations over the years as problematic in view of the continuously declining productivity growth in Germany (Bersch et al. 2018, 2019). Reasons for this reluctance can, for example, be an increase in productivity that cannot be proven or is difficult to prove, a negative cost/benefit balance or insurmountable contradictions and challenges (see e.g., Zimmermann 2017; Hirsch-Kreinsen and ten Hompel 2015, who point out potential barriers and obstacles to the implementation of digitalization strategies with regard to the "paradoxical character" of disruptive innovations). Arntz et al. (2016a) also point out that technical feasibility alone is not decisive as to whether the new digital technologies are actually used in companies. "In addition to the investment costs, the legal framework conditions and the company culture, the decisive factors include the associated expectations regarding the increase in labor productivity, the reduction of costs and the possibilities of being able to offer new products and services or acquire new customers." (Arntz et al. 2016a, p. 1, translated from German)

1.3 Digitalization from a Supra-Company Perspective

1.3.1 Digitalization and Employment

Triggered not least by the study by Frey and Osborne (2013), which found that 47% of US jobs were highly likely to be automated, numerous studies have also been conducted in and for Germany concerning the substitutability and/or the potential impact of increasing **digitalization and automation** on the **labor market** (Bonin et al. 2015; Cordes and Gehrke 2015; Dengler and Matthes 2015, 2018a, b; Düll et al. 2016; Kropp and Dengler 2019; Stettes 2018; Vogler-Ludwig et al. 2016). Due to differences in the data, models, evaluation routines and the assumptions that were used, the results are quite heterogeneous, especially with regard to the extent of the effects to be expected and their mutual influence (see critical discussions in Arntz et al. 2016b, 2020; Dengler and Matthes 2015, 2018a, b; Hirsch-Kreinsen 2016, 2018; Stettes 2018; for an international comparison based on the OECD Study of Adult Skills (PIAAC), see Nedelkoska and Quintini 2018).

Following a simple transfer of the approach taken by Frey and Osborne (2013), Bonin et al. (2015) first come to the conclusion that 42% of employees in Germany work in occupations with a high probability of automation. At the same time, however, they criticize the procedure and, like Dengler and Matthes (2015, 2018b), they argue for an activity-based approach to the analysis (Bonin et al. 2015). According to the authors, it is not the entire occupation that can be automated or substituted through the use of computers, but rather certain tasks or activities. Considering the **activity structures**, the proportion of employees working in occupations threatened by automation is reduced to 9% for the USA and 12% for Germany (Bonin et al. 2015).

Dengler and Matthes (2015, 2018a, b) use the so-called task-based approach of Autor et al. (2003) for their studies and analyze the occupations at the level of the activities to be performed. In 2013, "only" 15% of employees subject to social insurance contributions were affected by a high **risk of substitutability**. In the updated study (Basis: requirements matrix 2016), this proportion rose to 25% (Dengler and Matthes 2018a, b). The authors attribute the increase in particular to the fact that innovative technologies, such as mobile, collaborative robots, self-learning computer programs and some 3D printing and VR applications, have now reached market maturity and are thus potentially able to take over or replace human activities. With regard to Industry 4.0, it is noteworthy that the occupational segments "**manufacturing occupations**" (2013:73%, 2016:83%) and "occupations concerned with production technology" (2013:65%, 2016:70%) are still the most affected, but the increase from 2013 to 2016 was significantly weaker than, for example, that in the transport and logistics occupations, which recorded an increase of 20%. In industrial production, the finding that it is not only unskilled but also skilled occupations that are at greater risk of substitution than specialist and expert occupations (Dengler and Matthes 2015, 2018a, b) is certainly relevant, and possibly even explosive for industrial work in the future.

Although the introduction of new technologies, software systems and production processes since 2013 has led to the emergence of numerous new occupations, particularly those relating to the mastering of them, compliance with related laws and standards, or quality and process management, so far there have only been a few occupational profiles for which a successful adaptation to the digitalization-related changes in requirements has led to a reduction in their substitutability potential. Positive examples can be found among the IT service/scientific service professions as well as in the health sector. At the same time, however, completely **new occupational profiles** have emerged since 2013, such as the occupations of "data scientist" and "interface designer" (Dengler and Matthes 2018a, b).

On the potential impact of Economy 4.0 on the structure of occupational fields and requirements, reference is made to the scenario-based analyses (creeping digitalization vs. digital revolution) of the Federal Institute for Vocational Education and Training (BIBB), the Gesellschaft für Wirtschaftliche Strukturforschung (GWS) and the Institute for Employment Research (IAB) (e.g., Weber 2016; Wolter et al. 2016; Weber et al. 2017a, b).

Based on analyses by the Cologne Institute of the German Economy (IW Köln), Stettes (2018) does not expect negative employment effects, but emphasizes the design and development potentials. Similarly, based on their scenario analysis, Zika et al. (2018) only predict minimal effects of digitalization on the overall level of employment, instead, they expect significant shifts of jobs between sectors, occupations and requirement levels. A key prerequisite for exploiting these potentials is consistently seen in the provision and implementation of adequate **education and training** opportunities (e.g., Arnold et al. 2018; Düll et al. 2016; Weber 2017; Zika et al. 2018).

1.3.2 Development Scenarios for Production Work

From the perspective of industrial sociology, Hirsch-Kreinsen (2015) predicts a broad "**spectrum of diverging patterns in work organization**" (ibid., p. 93, translated from German) based on existing studies into the change of activity and qualification structures. This is limited by two poles: the "polarized organization" and the "swarm organization" (ibid., p. 93 et seqq., translated from German). The pattern of the polarized organization is characterized by two separate task areas: (i) simple, standardized monitoring and control tasks carried out by a few semi-skilled workers, and (ii) highly demanding anticipated activities and productivity management tasks with a large scope for action, the execution of which is the responsibility of a group of highly qualified experts and specialists (ibid.; cf. Hirsch-Kreinsen 2016). Hirsch-Kreinsen (2015, p. 94 et seq.; with reference to Neef and Burmeister 2005, among others) describes a swarm organization as a work organization in which simple activities are completely substituted by automation and the remaining anticipated and operative tasks (especially the solution of unpredictable problems) are not sharply defined and assigned, but are managed by an open, highly flexible team of very qualified, equal employees (engineers and skilled workers) in a self-organized and situation-dependent manner. While the strong separation of operational and anticipated levels in the first scenario leads to a polarization of qualifications, the combination in the swarm organization leads to an upgrading of qualifications (Hirsch-Kreinsen 2015, cf. Hirsch-Kreinsen 2016).

It can be assumed that not only operational activities but also indirect functions and control and management tasks will be affected by the implementation of Industry 4.0 innovations (Hirsch-Kreinsen 2014). Developments towards decentralized structures with flat hierarchies are expected combined with high **demands** in terms of flexibility and problem-solving skills as well as an increased merging of IT and production competences (Hirsch-Kreinsen 2014; Spath et al. 2013a, b; Uhlmann et al. 2013).

Three scenarios are differentiated in field studies concerning the identification of new requirements for the qualification of employees in the context of Industry 4.0 (Ahrens and Spöttl 2018, partly with additions based on Frenz et al. 2015):

1. **Tool scenario**: In this scenario, expert systems with the character of a tool are developed for qualified specialists, e.g., assistance systems for detailed production planning. Opportunities open up for demanding tasks that can be carried out by skilled workers with the contribution of their experience-based knowledge. Control and decision-making competences lie with the people. The qualification level of skilled workers is upgraded.
2. **Automation scenario**: The use of intelligent, self-controlling technologies is forced throughout. The scope of action for skilled workers is restricted, decisions are delegated to self-optimizing, highly automated systems. Increased complexity, decreased process transparency and the emergence of informational and functional distance increasingly prevent regulatory intervention by skilled workers. At the shop floor level, there are still simple monitoring tasks that can

be taken over by semiskilled workers. Troubleshooting can only be carried out by experts (e.g., for visual control, sensor technology, robot programming).

3. **Hybrid scenario**: The scenario describes close cooperation between people and machines. New forms of interaction and cooperation are developed for controlling and monitoring tasks, which lead to new demands on skilled workers. According to Ahrens and Spöttl (2018), the design of work organization gains particular importance in this scenario because it determines the type and quality of the requirements.

Frenz et al. (2015), among others, conducted six case studies in companies in the automotive supply industry, which they assign to the hybrid scenario. They analyzed tasks in detailed production planning that required the use of complex detailed planning systems. The systems that were used took over product data that was read out in real time via RFID chips and processed it further to give the skilled worker suggestions for designing the production process. Frenz et al. (2015) did not find any fundamental changes in the requirements, but rather an expansion through the assistance system.

1.3.3 Online Work Outside of Companies

Outside of company structures, the world of work is changing (more) rapidly. From the basic idea of sharing unused or little-used resources, the so-called sharing economy has developed, in which today, various actors with different motives participate in diverse business models. The sharing economy and platform economy are not only seen as drivers of service innovations, but they have also given rise to **new forms of (online) employment**, such as crowdwork.

Jeff Howe, who is considered the "discoverer" of so-called crowdsourcing, provides two definitions for the concept:

> The White Paper Version: Crowdsourcing is the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call.
>
> The Soundbyte Version: The application of Open Source principles to fields outside of software. (Howe 2006)

As different as these two definitions are, so are the reactions to this approach: While the second definition is associated with terms with rather positive connotations, such as independence, self-determination, equal participation (at least among Internet users), collaboration, creativity and innovation, the first definition provides points of attack for a thoroughly critical discussion about the repercussions on the hitherto stable structures of gainful employment, established labor & social law standards and protection mechanisms, and not least those criticisms surrounding the effects on the "**crowdworkers**" themselves working in the so-called platform economy.

Ver.di (United Services Trade Union) has been dealing intensively with these issues for several years and points out the sometimes precarious working conditions (ver.di 2015a, b). With reference to the guidelines for good digital work published by the Enquete Commission "Internet and Digital Society" of the German Bundestag from 2013, the trade unions have called for a need for reform and regulation to provide social security for digital work outside of company structures (see e.g., Schröder and Schwemmle 2014 as well as Eichhorst and Spermann 2015). In Germany, it is estimated that only "a few hundred thousand" people practice this form of online work, but the growth rates for this market are estimated at 30% internationally (Pongratz and Bormann 2017). According to Pongratz and Bormann (2017), online workers are the self-employed as well as employees and non-employed (students, unemployed, pensioners, etc.) with roughly equal shares in each group. According to Pongratz and Bormann (2017), the main problem areas of online work on Internet platforms are: refusal to pay, loss of rights (to contributed knowledge and results), price dumping (and consequently very low payment), information imbalance and data protection (see also Huws et al. 2018 and for an international view on the "gig-economy" cf. De Stefano 2016).

Leimeister et al. (2015a, b) provide a brief overview of the terminology, forms of work, roles, earning opportunities and the opportunities and risks of crowdsourcing and crowdwork. Schmidt (2017) also provides a distinction between the different forms of online work (cloudwork, crowdwork, gigwork).

So-called co-working concepts have also emerged in the context of the sharing economy. **Co-working** implies a new form of work organization in so-called co-working spaces, which provide at least professionally equipped, usually larger, open workspaces as well as areas for social interaction. Users are typically the self-employed, freelancers, participants in the so-called sharing economy and start-up companies who "pursue their individual projects and tasks together" (Bouncken and Reuschl 2016; Reuschl and Bouncken 2017). By bringing together employees from different companies or "freelancers" with different profiles and goals, co-working opens up new collaboration opportunities and fosters a sense of community within a shared space (Johns and Gratton 2013). Advantages are seen in the promotion of cooperative creativity and innovation, which is why co-working spaces are sometimes also set up within companies. Outside of companies, the risks of online work that have already been mentioned arise.

1.3.4 Digitalization and Demographic Developments

When assessing the opportunities and risks of increasing digitalization and automation, other trends have to be also considered, e.g., **demographic developments**, which are leading to a decline in the potential workforce in Germany and are resulting in an increased shortage of skilled workers for many companies (BMAS 2017, 2018; Burstedde and Risius 2017; Fuchs et al. 2019; Vogler-Ludwig et al. 2016). Other challenges relate to the risk of losing valuable process knowledge as a result of the departure of older employees, the increase in absenteeism as a result of inadequate working conditions, and the rise in the proportion of performance-changing workers. The diversity of the workforce in terms of disposition, constitution, competence and adaptation characteristics is increasing in many companies, not least as a result of globalization, migration and inclusion. Fields of action and concepts that are designed to meet these challenges have been known for a long time and range from the age-differentiated design of work, flexible, demographically robust forms of work and working time organization through to comprehensive approaches to life phase-oriented human resource management (cf. e.g., Gerlmaier et al. 2016; Hammermann and Stettes 2014; Kenny et al. 2008; Latos et al. 2017b; Schlick et al. 2017; Schlick et al. 2013; Schlick and Mütze-Niewöhner 2010; Schlick et al. 2009; Shephard 2000; Tempel and Ilmarinen 2013; WHO 1994). Some advocates of these approaches see digitalization as an opportunity to accelerate the still sluggish implementation in companies and to simultaneously drive forward digital transformation in the course of integrated strategies and interventions (see e.g., Hammermann and Stettes 2018).

Apt et al. (2018) also provide arguments in a similar direction: In their study concerning the use of digital assistance systems, they emphasize the potential of such systems for a **humanization of the world of work** that uses technological progress to achieve the target dimensions of health, participation and job quality.

However, the employee surveys cited at the beginning also make it clear that the digital technologies and work equipment that have been used to date do not yet meet the requirements for human-centered work design. For example, in the case of systems for controlling work processes and machines, poor operability and system malfunctions play a role (Müller-Thur et al. 2018). Such problems suggest that ergonomic design criteria have not been considered and that there are technical or functional deficiencies (Bridger 2018). However, the stress-triggering factors frequently reported in connection with the use of electronic information and communication technologies, such as time pressure, interruptions, multitasking and the dissolution of boundaries, point more strongly to work organization deficits that can arise in the course of digitalization, but which can, and should be specifically avoided or eliminated, just as in the case of ergonomic and technical deficits.

2 Human-Centered Design of Work 4.0 in Companies

In Germany, the term "work design" has a long tradition and there is a large multidisciplinary community of work scientists who deal with the human-centered analysis, evaluation and design of socio-technical work systems and/or individual subcomponents such as work processes, work tasks, workplaces or working conditions. In that sense, the scope of work design is somewhat broader than it might be associated with the traditional "job design" (cf. Hackman and Oldham 1976; Oldham and Fried 2016), although the corresponding theories and findings certainly form an important basis. Similarly, the notion of **work science** as it is used in Germany is inherently multidisciplinary and as such also includes concepts and methods which, in the English speaking world, may be attributed to research fields as diverse as Human Factors and Ergonomics, Industrial Engineering, Occupational Health and Safety or (applied) Organizational Psychology. While the validity of the various forecasts concerning the time periods and the extent of changes of work in the digital age is still disputed, there is agreement on at least two aspects: (1) Changes of work as a result of technological developments are already taking place and will likely increase; (2) There is **scope for decision-making and work design** that opens up the possibility of influencing the direction, course and outcome of these developments and this scope for decision-making and design should be used in terms of a **technologically and socially innovative, economic and humane design of work** (see studies cited above and Bauer et al. 2021; BMAS 2017, 2020b; Hartmann 2015; Hirsch-Kreinsen et al. 2018; Mütze-Niewöhner et al. 2021a, b; Wischmann and Hartmann 2018a, b; Wilson and Daugherty 2018).

Scopes for decision-making and design can be found both at the level of society as a whole and the national economy, as well as at the level of companies. Here, it can be broken down to the level of work activity and workplace design as well as the designing of environmental influences (see levels in the way work processes are viewed according to Luczak and Volpert 1987; Luczak et al. 2006a). To arrive at generally accepted design guidelines, framework specifications and, as the case may be, new or adapted regulations, the designing of Work 4.0 at the supra-company level requires an open discussion and negotiation process, especially between social partners and politics. In relation to the subject matter under consideration, the focus of the article is on the company level with its sublevels and thus on the design of work (4.0) within (manufacturing) companies.

2.1 Macro-Ergonomic and Micro-Ergonomic Dimensions of Work Design

In recent times, it has been possible to see an expansion of technology-centered design approaches to the realization of the concept of Industry 4.0 to include human and organizational aspects of work design. This is certainly not least due to the commitment of the trade unions. The white paper on Work 4.0 by the BMAS (2017), recent contributions by the Jacobs Foundation and the German Academy of Technical Sciences (e.g., Jacobs et al. 2017), publications from Industry 4.0- or digitalization-related funding programs (e.g., Botthof and Hartmann 2015; Wischmann and Hartmann 2018c; Bauer et al. 2021; Mütze-Niewöhner et al. 2021a) as well as the more recent publications by the Gesellschaft für Arbeitswissenschaft (translation: Work Science Society) (cf. http://www.gesellschaft-fuer-arbeitswissenschaft.de and the Zeitschrift für Arbeitswissenschaft) outline a differentiated discussion in which the central role of humans and the need for a human-friendly digitalization of work are rarely questioned.

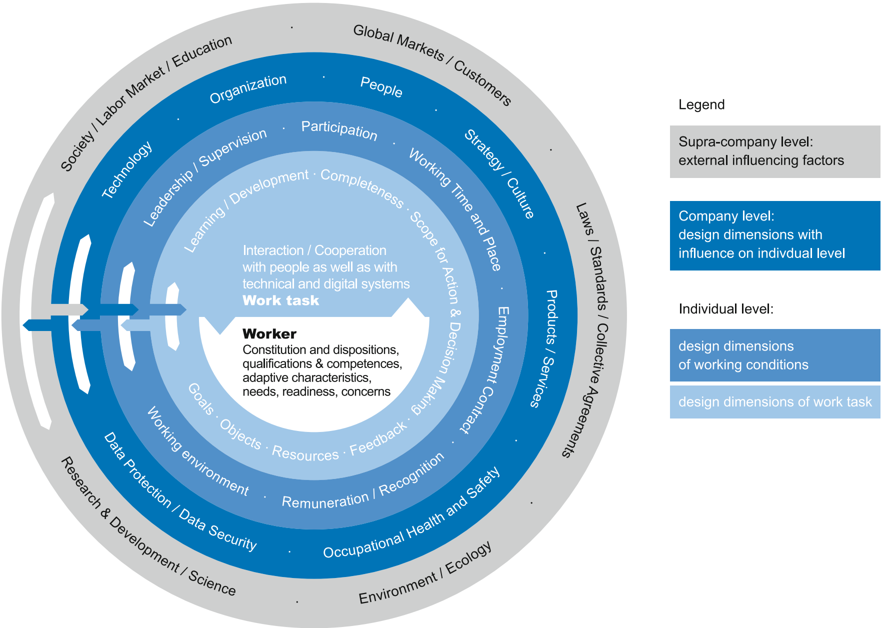


Fig. 1 Influencing factors and dimensions of human-centered work design in companies

The focus of human-centered work analysis and design is often on work system design, primarily at the level of the workplace. The work system models used in this context (see Schlick et al. 2018, p. 21 et seq.) cite the essential elements and relationships as well as input and output parameters. In particular, the extended work system model (see Schlick et al. 2018, p. 22) allows for a differentiated view of work forms, transformation processes (e.g., in the sense of processing objects using work equipment, energy, etc.) and interactions within socio-technical work systems and as such, it can still be recommended as a classification system for the design of workplaces or areas of work in the course of digitalization and automation projects.

If we look at the design of work in companies from a more macro-ergonomic perspective, other dimensions and influencing variables come into focus. These are relevant in more far-reaching redesigning processes in the context of company **transformation** processes and have an impact on the design of **work tasks and the conditions of people at work** (Fig. 1). The shell model presented here attempts to do justice to the complexity of the subject under consideration and show the main influencing factors from the corporate environment as well as the main design dimensions at the company and individual levels.

2.2 Design Aspects at the Level of the Person at Work

In terms of a human-centered design approach, the person at work with his or her individual characteristics, needs, dispositions and expectations is at the center of the model. The design level of the person at work comprises the two inner shells and as such, the design dimensions that directly affect the person at work and are perceived by him or her as work tasks and conditions with the resulting demands and stresses. In addition to the working environment, further relevant aspects are listed or highlighted that are (justifiably) given special attention in the discussion about work in the digital age: working hours and place of work, the employment contract, remuneration and other forms of recognition, direct management and supervision (e.g., through assistance systems), as well as the possibilities of **participation** in operational processes, which are usually not attributed to the main task, e.g., the redesigning of work processes or the adaptation of digital support systems (participative work design).

The **work task** is described in a more differentiated way, wherein design dimensions are cited that can be regarded as established findings of work science research, some of which have already found their way into various standards, and therefore have to be included in the consideration in the event of a planned change to the human-machine division of functions. On the one hand, the prominent position of **interaction** and **cooperation** considers the increased use of technical and digital systems and on the other hand, the associated need for ergonomically designed interfaces, etc. Furthermore, it notices developments and scenarios that suggest that communication, interaction, and cooperation between people—both as part of the work task in the context of self-directed group work and in the context of participative innovation work—will increase overall.

People are represented twice in the shell model in Fig. 1, on the one hand as individuals (in the center) and on the other hand as part of the entire workforce and/or certain groups of people who, for example, are or should be the target group of specific qualification offers or working time offers. Whereas the individual is viewed rather abstractly in the design phase (rough planning) of design or transformation processes as part of a group of people for whom certain characteristics or effects of stresses can be assumed as valid (preventive and prospective work design), in the detailed planning phase, the individual prerequisites and needs of those involved and those who are affected must be considered. The specific effects on the people at work must be evaluated during and after the implementation and/or testing phase to be able to adjust (corrective work design) or, as the case may be, even to justify the withdrawal of design interventions that have been tested on a pilot basis (see criteria in Sect. 3).

2.3 Design Aspects at the Company Level

The designing of Work (4.0) does not just take place at the societal or supra-company level, but also within manufacturing companies. In projects for the **reorganization** and further development of small and medium-sized enterprises, a classification scheme comprising the seven dimensions listed in Fig. 1 has proven itself to be useful, particularly in the orientation phase and in the rough planning phase. The classic dimensions of the MTO approach (MTO = Man, Technology, Organization, see Strohm and Ulich 1997) are taken up in the categories "**people**" (in the sense of the entire workforce, including managers), **technology** (including technical systems) and **organization** (with the sub-aspects of operational and structural organization). The following design dimensions must also be included from the beginning, especially in the planning of digitalization and automation projects: **strategy and culture** (e.g., complementary vs. technology-centered automation strategy, see Hirsch-Kreinsen 2016, as well as questions of management culture during and after the transformation), **products and services** (new business ideas and models that either drive the design process or whose development is initiated in parallel to compensate for expected substitution effects), **occupational health and safety**, as well as all sub-aspects around the topic of "**data**." Data protection and data security issues, which continue to play a central role in Industry 4.0 projects, are dealt with in detail elsewhere in this handbook.

The dimensions represent central subject areas of the design of work in the company context, which have a direct impact on the tasks and working conditions of the people at work. Each dimension can be assigned numerous sub-aspects that are relevant in the context of change processes and must be reflected on, discussed, processed and, if necessary, negotiated with the involvement of the works council and other company experts (e.g., data protection officers, occupational health and safety specialists, human resources managers). The application of this grid is not about the unambiguousness and correctness of the assignment, but about collecting and ordering the essential design aspects to then prioritize and transfer them to a participatory analysis and design process.

Selected design aspects are briefly highlighted below. For reasons of space, a detailed discussion of all of the dimensions must be dispensed with, but some of them are dealt with in other contributions.

2.4 Work Design Strategies in the Context of Industry 4.0

Around the turn of the millennium, Grote et al. (1999, 2000) and Wäfler et al. (1999, 2003) published their method for the complementary analysis and design of production tasks in socio-technical systems (KOMPASS). Reading the two volumes can be recommended here, as the authors dedicate themselves to the central questions of the division of tasks between humans and technology in the context of automation projects in production from a perspective that is at the same time, well-founded in occupational psychology and practice-oriented. “The aim of complementary system design is to take into account the complementary differences between humans and technology in the design of work systems in order to contribute to their safety and efficiency” (Grote et al. 1999, p. 255, translated from German). In this context, people, technology and organization must be considered together and coordinated with each other. In terms of **human-technology interaction**, the design approach demands the controllability of technical systems by the human operator as a prerequisite for taking responsibility (Grote 2018; Grote et al. 2000). The complementary automation strategy forms the contrast to a purely technology-centered strategy (see also Hirsch-Kreinsen 2015, 2016) and is thus in principle, suitable for achieving the goals of human-centered work design. The pursuit of a complementary strategy for the design of modern production work presupposes a (management) culture in which the image of a human prevails that classifies people as reliable and capable of mastering and controlling the process (Wäfler et al. 1999). It is obvious that depending on the desired end state (see scenarios and maturity models for Industry 4.0 cited above), there is more or less potential for realizing a complementary automation strategy.

The topic of “**participation**” is also a question of strategy and culture. Although this design dimension is explicitly placed at the level of the person at work in Fig. 1 (in the sense of existing participation opportunities outside of the main task), the extent and quality of employee participation in the redesign or reorganization of work systems and processes is usually decided at the management level. The application of the principle of participatory work design (Duell 1983) is especially aimed at the advantages of including the expert knowledge of the work persons and the associated increase in acceptance (see also Zink 2015). Various studies make a positive reference to the ability to innovate in companies (e.g., Biedermann et al. 2013; Myskovszky von Myrow et al. 2014, 2015) or prove—in accordance with motivational-psychological goal-setting theory—the productivity-increasing effect of participative management approaches (see e.g., Schmidt 2004; Pritchard et al. 2008). An example of the participatory, simulation-based design of a flexible, demographically robust assembly organization is described by Latos et al. (2017b).

2.5 Flexible and Cooperative Work Organization

It is undisputed that the vision of a so-called smart factory is accompanied by far-reaching changes in the organization of work and operations, which, for example, hold out the prospect of a move away from cyclic assembly line production towards island production (e.g., see Audi AG 2017). One driver of organizational change is the claim, often postulated with Industry 4.0, of being able to manufacture customized products in batch size 1. Rising variant diversity and the increasing networking of people and machines increase the complexity of systems and processes and require a high degree of flexibility in terms of the overall socio-technical system (e.g., see Latos et al. 2017a, 2018). At the same time, demographic developments demand organizational forms that do justice to different performance and competence profiles. Furthermore, there is a demand for "**organizational agility**," which allows companies to respond quickly, effectively and efficiently to changes in a complex, dynamic environment characterized by uncertainty (Jacobs et al. 2017, p. 15; Steireif et al. 2020; for example, concerning agile methods, see also Morris et al. 2014; more differentiated, critical reflections can be found in Sauer 2017; Neumer et al. 2018; Neumer and Nicklich 2021; Pfeiffer et al. 2019).

Various publications aim for a better understanding and definition of this "organizational agility" by providing individual conceptual and measurement models (see e.g., Sherehiy et al. 2007; Lindsjørn et al. 2016; Bergmann and Karwowski 2019). As most studies and approaches refer to software development or to the application of specific agile project management methods (Niederman et al. 2018; Saleh et al. 2019), further research should strive for a better contextualization of organizational agility and the investigation of its effects, in particular w.r.t. complex and technology-intensive development environments.

Team-based forms of work organization offer a great potential to meet the requirements in terms of responsiveness, flexibility and diversity, and in particular, also in terms of the ability to solve complex problems that require diverse, sometimes very specific technological or process-related knowledge (Mütze-Niewöhner et al. 2021b). **Group work** is already established in many manufacturing companies. The question here is to what extent the existing structures can be used for the implementation of Industry 4.0.

According to a study by Kinkel et al. (2007), in 2006, almost 75% of German industrial companies had introduced group work into their production, but only 12% used forms of group work that are characterized by a high degree of self-direction and are associated with complete, more demanding activities for the group and/or for the individual employee. The development and introduction of so-called holistic production systems was certainly linked to the hope of an accompanying increase in the share of higher-value forms of group work (e.g., Kinkel et al. 2008; Schlick et al. 2018). Although group work can be regarded as an integral part of such production systems (Dombrowski et al. 2006; Bikfalvi 2011), the forms of group work that are realized still do not meet the criteria of qualified group work equipped with a great scope for action and decision-making (see Abel and Ittermann 2014; Stranzenbach 2018). With regard to the realization of organizational scenarios that predict largely autonomous, flexible groups with a high level of competence, the existing group work structures nevertheless (or precisely because of this) offer a potential that could be exploited in the course of Industry 4.0 projects. Above all, positive effects on the target variables of group work can be expected from strategies in the direction of the "tool scenario" (Stranzenbach 2018; in this regard, it should be noted that Stranzenbach did not consider the "hybrid scenario").

One of the biggest challenges in modern production systems is managing **complexity**. Latos (2020) studied the effects of complexity on the performance of production groups and came up with interesting results: while task complexity had a positive effect on the quality of group work and its outcomes, group design complexity (e.g., heterogeneity, dynamics) led to negative effects. Based on his findings, the author derives specific recommendations for the complexity-related design of group work in production.

The interactions with other design dimensions become particularly clear in the area of organization. While decentralization efforts largely enable self-directed forms of teamwork, the increasing degree of automation reduces the scope for designing complete group tasks that are conducive to learning and which, in particular, provide for cooperative, possibly digitally assisted control by the people in the **human-machine team**. From a work science perspective, the demand for responsive human-machine teams equipped with high degrees of autonomy certainly holds potential. However, given the spectrum of divergent forms of work organization (Hirsch-Kreinsen 2015, 2016), it is still largely open as to whether these will primarily be groups whose members monitor complex systems and are merely "coupled" to each other by data technology, or whether "real" work groups will emerge in which humans work collaboratively with robots and digital assistants. An overview of the different forms of group and team work can be found in Schlick et al. (2018). For example, regarding the specific characteristics and challenges of distributed teams, see O'Leary and Cummings (2007) or Boos et al. (2017).

The continuous improvement of advanced information technologies allows for a higher temporal and spatial flexibility of work organization and sets new demands on leadership, communication and collaboration in distributed teams (Antoni and Syrek 2017). For handling the rising complexity in virtual settings, emerging leadership concepts such as shared leadership need to be considered, as they are positively associated with team performance in **virtual teams** (Hoch and Kozlowski 2014; Robert and You (2018). Such leadership concepts can capitalize on a variety of communication tools which support team collaboration and compensate the lack of face-to-face meetings to some degree (Mayer et al. 2020). High task interdependence often increases the complexity of collaboration in virtual team settings, so the design of agile work should be discussed.

Analyzing qualitative case studies, Pfeiffer et al. (2019) conclude that, under certain assumptions, agile methods can protect team members against stress and voluntary self-exploitation. However, they emphasize the fragility of this "safe space" and the need for accompanying organizational backing (ibid). Recommendations for the design of agile work can also be found in Wille and Müller (2018).

In the context of project management, there are numerous approaches for describing such complex working environments, which, however, predominantly focus on the complexity of a single project (see e.g., Geraldi et al. 2011; Botchkarev and Finnigan 2015). There is still a need for research regarding the measurement of complexity in multi-project management. In particular, it is necessary to investigate the effects of complexity on the success and stress of project managers and whether there are dependencies on the degree of agility of the organizational working environment (Harlacher et al. 2021).

For team-related research, two central research fields emerge, among others: on the one hand, there is the investigation of the effects of agility and complexity on the quality and results of teamwork, and on the other hand, there is the clarification of open questions in the field of human-machine cooperation.

The demand for agility is also associated with requirements for the **temporal and spatial flexibility** of employees (Jacobs et al. 2017). Under the aspects of time sovereignty and the work-life or work-domain balance, both opportunities and risks as well as the demands for more or less regulation are controversially discussed. According to an analysis by Piele and Piele (2017), flexible working time models are widespread in companies in the manufacturing sector, however, they only bring a partial gain in time sovereignty for employees, as the dissolution of rigid working time limits, not only allows for more self-determination, but also for external determination. The effects of externally determined flexibility are particularly evident in the case of trust-based working time, e.g., the forfeiture of working time, weekend work and the violation of legal working time regulations (Piele and Piele 2017).

Professionals who do so-called knowledge-based work are already largely independent of time and place thanks to the Internet, WLAN, VPN, laptop and smartphone. The high availability of information and communication technologies, including corresponding end devices, as well as the increasing digitalization of work objects, means and processes, lead to an overall increase in the share of work tasks that require the handling of knowledge, data and information (Beermann et al. 2017). It is therefore to be expected that forms of mobile working will also increase in manufacturing companies. Boosted by the pandemic, economy and society perceive working from home or alternatives to the office as increasingly necessary in times of globalization and also as increasingly feasible with modern technologies (Bonin et al. 2020). Nevertheless, it needs to be discussed to which extent mobile work can be successful for companies and their employees in the long-term. A high variety of tasks can be done from home, but face-to-face meetings stay important especially for the development of a trustful relationship with team members (Hill et al. 2008). Further, life circumstances and personality traits of employees should be considered for a **mobile working** setting to elicit job satisfaction (Lott 2020). Two publications by the Federal Institute for Occupational Health Safety deal with the health opportunities and risks of location-flexible and time-flexible working (Beermann et al. 2017) and the effects of work-related extended accessibility to life-domain balance and health (Pangert et al. 2016). You will find design approaches as well as indications of the need for further research there.

2.6 Use of Assistance Systems and Cooperative Robots

The studies cited above concerning the status of digitalization in companies suggest that small and medium-sized manufacturing companies in particular—in view of the investment costs, the existing uncertainties with regard to data protection, data security, technical reliability and occupational safety, and not least because of the smaller distance between company management and employees—will tend to develop in smaller steps (and not disruptively) in the direction of Industry 4.0. The approach to new technologies typically takes place within the framework of pilot projects that focus on specific work processes or workplaces with regard to digitalization or mechanization possibilities and seek to support them, e.g., through **assistance systems**. From a work science perspective, the potentials lie, for example, in the reduction of high levels of physical and mental stress, in the reduction of hazards and barriers (e.g., for people with disabilities or those who have altered performance), in the mastering of complexity (e.g., through the processing of information in line with requirements), in greater effectiveness and efficiency (e.g., through the reduction of search tasks) and in the use of new technologies (e.g., by reducing search complexity or simpler, more intuitive operation of work and operating equipment) as well as in the design of complete activities that are conducive to learning and motivation with room for maneuver and opportunities for cooperation (e.g., see Czerniak-Wilmes et al. 2017; Petruck et al. 2019; Paetzold and Nitsch 2014; Wischmann and Hartmann 2018b; as well as chapter "Criteria for analysis, assessment and design" in this article).

Klocke et al. (2017) explain the importance of assistance systems in the "Internet of Production" and in particular, emphasize the role of humans. While the focus in the production context has so far been on the technical support and automation of mechanical functions for physical relief, the focus today and in the future will be on the support of unstructured, variable tasks with a high cognitive load and with it, on the development of technical production assistance systems "that support human decision-making in the implementation of design, optimization and planning tasks on the basis of sensor and modeling data, processed by means of artificial intelligence algorithms, along the value chain" (Klocke et al. 2017, p. 270 et seq., translated from German; see also VDI/VDE-GMA 2016; examples of assistance systems can also be found in Apt et al. 2018).

In the report on health and safety at work (BMAS 2018), assistance systems are also attributed potential benefits in view of the reported stresses and causes of work-related illnesses (e.g., use of exoskeletons to reduce physical stresses with the aim of avoiding musculoskeletal disorders). Research results concerning the risks associated with the use of head-mounted displays (HMD) and recommendations for the safe and optimal use of HMD are provided by Wille et al. (2015), Theis et al. (2016), and Wille (2016).

Especially in the field of production and logistics, highly automated robotic platforms are increasingly being used to assist people directly in the workplace (referred to as "cooperative robots" in the following). The requirements for the design of cooperative robots differ significantly from those of traditional industrial robots. For example, since there are no separating protective devices here to protect humans from collisions with the robot, it is necessary to consider a usage concept for these workplaces that does not have a negative impact on safety, productivity or user acceptance (Leichtmann et al. 2018; Petruck et al. 2018).

A lack of consideration of humans in the development of cooperative robot systems can lead to the user not being able to master the functions, even in functionally well-designed systems, due to the complexity or lack of usability.

This can result in the developed systems not being successful in the market or simply not being used by employees (cf. Bröhl et al. 2019). High costs can also result from not considering the user, if complex systems require long learning phases and possibly re-learning times as a result of unlearning (Colceriu et al. 2020). Finally, serious operating errors can occur due to a lack of consideration of human perception, decision-making and action processes (Behnke et al. 2017).

Decades of work science research into the humane and economically compatible design of **human-machine systems** have already provided important findings and design recommendations. However, there is still a great need for research into the design of workplaces with highly automated cooperative robot systems. Currently, the following topics, among others, are focused in this field:

1. the role of the cooperative robot in the work system (e.g., cooperative robots can temporarily assume the role of a work tool as well as that of a work object or a work person, which in turn can have different effects on work processes)
2. the influence of robot movements on the user (e.g., the effects of speed, acceleration, movement patterns and distances on user acceptance)
3. the user-appropriate division of tasks between humans and robots (e.g., to avoid over-challenging and under-challenging)
4. intuitive user interfaces (e.g., for simplified programming), as well as
5. influencing factors in human-multirobot systems (e.g., changed awareness of situations)

With reference to the model of work design introduced above, it should be made clear at this point that even in projects relating to individual workplaces, all design dimensions (see Fig. 1) should be considered during the design and planning phase: Questions related to people and (working) conditions play an important role if, for example, the use of cooperative robots (dimension "technology") is planned to support workers in carrying out tasks. In the dimension "occupational health and safety", for example, changes in the load situation and possible collision hazards must be determined. Data protection aspects are also relevant depending on which data are to be provided to or, as the case may be, recorded on the robot (e.g., usage times in connection with quality data recording). In the dimension "people", questions arise about the acceptance and the subjective experience of stress as well as the qualification requirements that can result from handling the robot and from the changed work process. The modification and description of the work process concerns the dimension "organization". Here as well, questions related to task division and responsibilities (e.g., for control, maintenance, and servicing) must be clarified. The expected effects on task content, the scope of action and the decision-making ability of the persons at work must be considered, as well as potential effects on other company departments, e.g., if the cooperative robot takes over work tasks that were previously carried out by persons at work from different departments. Availability and reliability must be checked and interfaces to existing systems must be realized (dimensions "technology" and "organization"). References to "products and services" can also arise, e.g., in the form of requirements for product design. Limiting the application to certain components or variants or expanding it by using the system to extend the range of services is also conceivable. In the chosen example, aspects of corporate culture and leadership concern, among other things, overarching issues relating to the target Industry 4.0 maturity level, the handling of acceptance problems and job insecurity.

3 Criteria for Analysis, Assessment, and Design

The work science literature provides various systems of criteria and instruments for the analysis, assessment and design of work, (see overviews in Luczak 1997; Dunckel 1999; Greif and Hamborg 2018; Schlick et al. 2018; Schmidt et al. 2008) which can be assigned to five hierarchically linked **levels of assessment**: (1) harmlessness and tolerability (lowest level), (2) feasibility, (3) freedom from impairment and reasonableness, (4) satisfaction and personality development and (5) social compatibility (Luczak and Volpert 1987).

On compliance with the three lowest levels of assessment, reference is made to the relevant human factors and ergonomics literature which set out the relevant sub-criteria, findings and recommendations and refer to applicable standards and laws (e.g., Bridger 2018; Rogers et al. 2015; Schlick et al. 2018; Schmauder and Spanner-Ulmer 2014; Schmidt et al. 2008). It is worth recalling the instruments for risk assessment and the standards and recommendations of the accident insurance institutions that must be applied or observed when changing or redesigning workplaces (see the websites of the Federal Institute for Occupational Health and Safety (Bundesanstalt für Arbeitsschutz und Arbeitsmedizin) and the German Public Accident Insurance Umbrella Organization (Deutsche Gesetzliche Unfallversicherung)) particularly in connection with the assessment of potential impairment. However, it must be noted that studies concerning the effects of digital technologies and assistance systems—both on key business figures and in particular, on the health, well-being, performance, motivation, and satisfaction of people at work—are still ongoing, not least because they almost inevitably lag behind implementation. The need for research has already been pointed out in various places.

The "ErgoCAM" system can be mentioned here as an example of a digital assistance system that supports operational practitioners in the ergonomic analysis and evaluation of human movements and activities in real time (Johnen et al. 2018; for the theoretical and empirical foundations, see Brandl et al. 2016; Brandl 2017). Digitalization strategies are also being used within occupational science to better integrate occupational science instruments and data into the world of work.

When assessing the relevant human factors and work science literature, especially the literature on organizational psychology, the following criteria can be cited for work task design that promotes motivation and health, serves the need for personal development and is socially acceptable (see, among others, Algera 1990; Fried and Ferris 1987; Grote et al. 1999, 2000; Hacker and Sachse 2014; Hackman and Oldham 1976, Luczak et al. 2006a, b; Morgeson and Humphrey 2008; Nerdinger et al. 2014; Schuler and Sonntag 2007; Schlick et al. 2018, p. 689 et seqq; Spector 1986; Susman 1976; Ulich 2011):

- the completeness of the task with the partial aspects of sequential and hierarchical completeness (see also DIN EN 29241-2:1993; DIN EN 614-2:2008; DIN EN ISO 6385:2016)
- the scope for action in the performance of tasks (including degrees of freedom in terms of time and content or decision-making scope and authority in the sense of autonomy and self-determination as well as possibilities of self-control in the sense of work-based feedback)
- the transparency and designability of the work processes
- the significance of the task
- the variety of demands the task makes on the sensory, cognitive, and motor systems of the worker
- communication and cooperation requirements
- learning and development opportunities
- transparent and accepted objectives, and
- the availability of necessary resources

In connection with the complementary automation strategy, reference has already been made above to the KOMPASS procedure, which defines further criteria for the analysis and design of production tasks, e.g., information and execution authority, dynamic coupling and flexibility (Grote et al. 1999). Increasingly, in addition to classical ergonomic criteria, such as usability (see e.g., DIN EN ISO 9241-11:2018; Herczeg 2018; Rogers et al. 2015) and acceptance, further far-reaching ethical, legal, and social aspects or implications (ELSA or ELSI) are considered when designing work equipment and objects and human-machine interfaces. This is done, for example, to assess and avoid associated risks during the design phase (Mertens et al. 2018; Nelles et al. 2017; Wille et al. 2016). In the context of European research funding, the requirement for social responsibility in the research and development of innovations also includes aspects pertaining to environmental sustainability (see Responsible Research and Innovation in European Commission 2013).

4 Guiding Principles for Human-Centered Work Design in the Digital Age

With their core definition of work science, Luczak and Volpert (1987) have already provided a model for work design in the field of tension between rationalization and humanization. Here, this will be used for the design of work in the context of Industry 4.0 and digitalization. From the work science perspective, guiding principles for the human-centered design of work in the digital age ("Work 4.0") could be as follows (based on Luczak and Volpert 1987, own additions, omissions and modifications):

People work

- under feasible, harm-free, impairment-free, and health-promoting working conditions
- in effective, efficient, safe, and sustainable work processes and systems
- in which they are supported by ergonomically designed technical and digital systems as needed to accomplish full and meaningful tasks
- in which they see standards of ethical, legal, and social appropriateness fulfilled in terms of work content, scope, time, place, environment, pay, freedom of action and design, autonomy and informational self-determination, as well as feedback, appreciation, participation and cooperation, and
- in which they can use, maintain, and expand their qualifications and competences and develop their personality in interaction and cooperation with others

5 Summary

Despite all the heterogeneity, the studies concerning the status and development prospects of work, as a whole, leave no doubt that increasing digitalization, automation and networking will have, and already has an impact on gainful employment in manufacturing companies. The decisive factor will be whether we succeed in using the existing scope for shaping work at the various levels to exploit the diverse potentials of digital transformation for an innovative and social (both economic and human-centered) shaping of work in the digital age.

From a social perspective, the potentials of digitalization relate, for example, to the emergence of new types of tasks and professions (e.g., through technology, product and service innovations or the development of innovative business models) as well as the adaptation of existing professions, activities, requirement profiles and participation conditions (e.g., using adaptive assistance systems, robotic systems, or flexible, life-phase-oriented organization and work models).

At the company level, there are also opportunities for a prospective and balanced work design that considers technological, economic, and human criteria. For example, the (re-)design of work systems can be pursued to reduce stress and hazards as well as monotonous, under-demanding or over-demanding work tasks and instead, increasingly integrate complete activities with learning and development opportunities into the work process. Semi-autonomous, team-based forms of work organization and ergonomically designed, adaptable assistance systems that process data transparently and in a user-friendly way can, for example, help to control or reduce the increasing complexity as a result of networking, product diversity and individualization.

To tap the potential of a human-centered work design at the company level, the paper recommends a shell model that can be used to identify essential dimensions and effects on employees (including managers), to develop an economically and socially justifiable digitalization and automation strategy, to derive holistic target systems and to test and evaluate these within the framework of participatory design processes. One focus of workplace design should be on the macro-ergonomic and micro-ergonomic design of the interaction and cooperation between people and digital and technical systems, for which findings from work science research are already available. Reference should also be made here to the results of the research initiative "Arbeit in der digitalisierten Welt" (Work in the digitalized world), which involved 29 collaborative projects that were funded jointly by the Federal Ministry of Education and Research (BMBF) and the European Social Fund (see Bauer et al. 2021 for an overview over the results and further publications).

Although comprehensive, the findings are not yet sufficient in all research areas surrounding the future of work to derive generally applicable, empirically validated and practical recommendations for the design of work in the digital age. There is still a need for research, especially when it comes to determining medium to long-term effects on working people. At this point, reference should be made to funding initiatives from which valuable application-oriented findings and practical examples for work design can be expected in the short and medium term, e.g., the research and funding priorities initiated by the Federal Ministry of Education and Research (BMBF) "Zukunft der Arbeit: Mittelstand – innovativ und sozial" (Future of work: SME – innovative and social).

The conceptual model for the human-centered design of work in the digital age picks up on essential criteria that are already applied in the analysis and evaluation of plans for design, design outcomes and design alternatives in the field of work science and should be applied in the course of designing of our future world of work. With the proposed guiding principles, the authors would like to stimulate critical discussion beyond the field of work science and look forward to additional contributions, further developments and specific definitions, especially with regard to a more differentiated consideration of forms of work in which humans and machines manage work tasks cooperatively.

References

Abel J, Ittermann P (2014) Teilautonom oder re-taylorisiert? – Gruppenarbeit bei Ganzheitlichen Produktionssystemen. Arbeit 23:277–291

Ahrens D, Spöttl G (2018) Industrie 4.0 und Herausforderungen an die Qualifizierung von Fachkräften. In: Hirsch-Kreinsen H, Ittermann P, Falkenberg J (eds) Digitalisierung industrieller Arbeit: Die Vision Industrie 4.0 und ihre sozialen Herausforderungen, 2., aktualisierte und erweiterte Edition. Edition sigma, Nomos Verlag, Baden-Baden, pp 175–194

Algera JA (1990) The job characteristics model of work motivation revisited. In: Kleinbeck U, Quast HH, Thierry H, Hacker H (eds) Work motivation. Erlbaum, Hillsdale, NJ, pp 85–103

Antoni CH, Syrek C (2017) Digitalisierung der Arbeit: Konsequenzen für Führung und Zusammenarbeit. Gr Interakt Org 48:247–258. https://doi.org/10.1007/s11612-017-0391-5

Apt W, Bovenschulte M, Priesack K, Weiß C, Hartmann EA (2018) Einsatz von digitalen Assistenzsystemen im Betrieb. Bundesministerium für Arbeit und Soziales (eds) Forschungsbericht 502, Berlin

Arnold D, Arntz M, Gregory T, Steffes S, Zierahn U (2018) No need for automation angst, but automation policies. In: Neufeind M, O'Reilly J, Ranft F (eds) Work in the digital age: challenges of the fourth industrial revolution. London, pp 75–87

Arntz M, Gregory T, Lehmer F, Matthes B, Zierahn U (2016a) Dienstleister haben die Nase vorn: Arbeitswelt 4.0 - Stand der Digitalisierung in Deutschland. IAB-Kurzbericht 22, Institute for Employment Research (IAB), Nuremberg

Arntz M, Gregory T, Zierahn U (2016b) The risk of automation for jobs in OECD countries: a comparative analysis. OECD Social, Employment and Migration Working Papers, No. 189, Paris

Arntz M, Gregory T, Lehmer F, Matthes B, Zierahn U (2017) Technology and jobs in the fourth industrial revolution. http://conference.iza.org/conference_files/MacroEcon_2017/gregory_t8580.pdf. Accessed 12 Feb 2021

Arntz M, Gregory T, Zierahn U (2020) Digitization and the future of work: macroeconomic consequences. In: Zimmermann KF (ed) Handbook of labor, human resources and population economics. Springer International Publishing, Cham, pp 1–29. https://doi.org/10.1007/978-3-319-57365-6_11-1

Audi AG (2017) Smart factory. Dialoge-Magazin, Issue 2017. https://www.audi-mediacenter.com/de/publikationen/magazine/dialoge-smart-factory-2017-364. Accessed 11 Feb 2021

Autor DH, Levy F, Murnane RJ (2003) The skill content of recent technological change: an empirical exploration. Q J Econ 118:1279–1333

Bauer W, Mütze-Niewöhner S, Stowasser S, Zanker C, Müller N (eds) (2021) Arbeit in der digitalisierten Welt: Praxisbeispiele und Gestaltungslösungen aus dem BMBF-Förderschwerpunkt. Springer Vieweg, Berlin, Heidelberg

Beermann B, Amlinger-Chatterjee M, Brenscheidt F, Gerstenberg S, Niehaus M, Wöhrmann AM (2017) Orts- und zeitflexibles Arbeiten: Gesundheitliche Chancen und Risiken. 1. Edition. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Dortmund, Berlin, Dresden

Behnke G, Leichtmann B, Bercher P, Holler D, Nitsch V, Baumann M, Biundo S (2017) Help me make a dinner! Challenges when assisting humans in action planning. In: 2017 IEEE International Conference on Companion Technology (ICCT), pp 1–6

Bergmann T, Karwowski W (2019) Agile project management and project success: a literature review. In: Kantola JI, Nazir S, Barath T (eds) Advances in human factors, business management and society, vol 783. Springer International Publishing, Cham, pp 405–414

Bersch J, Diekhof J, Krieger B, Licht G, Wagner S (2018) Abnehmendes Produktivitätswachstum – zunehmende Produktivitätsunterschiede. ZEW policybrief No 4, Zentrum für Europäische Wirtschaftsforschung, Mannheim

Bersch J, Diekhof J, Krieger B, Licht G, Murmann S (2019) Productivity slowdown, innovation and industry dynamics. In: Lehmann EE, Keilbach M (eds) From industrial organization to entrepreneurship. Springer International Publishing, Cham, pp 229–241

Biedermann A, Dreher C, Gerstlberger W, Klippert J, Kopplin S, Meszléry K, da Mota Pedrosa A (2013) Partizipatives und ganzheitliches Innovationsmanagement. Sichtweisen von Betriebsleitungen und Betriebsräten. Reihe: Forschung aus der Hans-Böckler-Stiftung, Band 157. Edition sigma, Nomos Verlag, Baden-Baden

Bikfalvi A (2011) Teamwork in production: implementation, its determinants, and estimates for German manufacturing. Hum Factors Ergon Manuf Serv Ind 21(3):244–259. https://doi.org/10.1002/hfm.20230

Bonin H, Gregory T, Zierahn U (2015) Übertragung der Studie von Frey/Osborne (2013) auf Deutschland. Kurzexpertise Nr. 57, Zentrum für Europäische Wirtschaftsforschung, Mannheim

Bonin H, Eichhorst W, Kczynska J, Kümmerling A, Rinne U, Scholten A, Steffes S (2020) Verbreitung und Auswirkungen von mobiler Arbeit: Kurzexpertise. https://www.ssoar.info/ssoar/handle/document/70079. Accessed 11 Feb 2021

Boos M, Hardwig T, Riethmüller M (2017) Führung und Zusammenarbeit in verteilten Teams. Hogrefe, Göttingen

Bostrom N (2014) Superintelligence – paths, dangers, strategies. Oxford University Press, Oxford

Botchkarev A, Finnigan P (2015) Complexity in the context of information systems project management. Organ Project Manag 2:15. https://doi.org/10.5130/opm.v2i1.4272
Botthof A, Hartmann E (eds) (2015) Zukunft der Arbeit in Industrie 4.0. Springer Vieweg, Berlin, Heidelberg
Bouncken RB, Reuschl AJ (2016) Coworking-spaces: how a phenomenon of the sharing economy builds a novel trend for the workplace and for entrepreneurship. Rev Manag Sci 12:317–334. https://doi.org/10.1007/s11846-016-0215-y
Brandl C (2017) Ergonomische Analyse von Körperhaltungen in Produktionssystemen für eine computergestützte Arbeitsgestaltung und -organisation. Dissertation RWTH Aachen University, Shaker, Aachen
Brandl C, Bonin D, Mertens A, Wischniewski S, Schlick CM (2016) Digitalisierungsansätze ergonomischer Analysen und Interventionen am Beispiel der markerlosen Erfassung von Körperhaltungen bei Arbeitstätigkeiten in der Produktion. Z. Arb. Wiss. 70:89–98. https://doi.org/10.1007/s41449-016-0016-9
Bridger RS (2018) Introduction to human factors and ergonomics, 4th edn. Taylor & Francis, New York
Bröhl C, Nelles J, Brandl C, Mertens A, Nitsch V (2019) Human–robot collaboration acceptance model: development and comparison for Germany, Japan, China and the USA. Int J Soc Robot 11(5):709–726
Burstedde A, Risius P (2017) Studie 2/2017: Fachkräfteengpässe in Unternehmen: Regionale Fachkräftesituation und Mobilität. Institut der deutschen Wirtschaft Köln e.V. (eds) Cologne. https://www.iwkoeln.de/fileadmin/publikationen/2017/328843/IW-Gutachten_Regionale_Fachkraeftesituation_und_Mobilitaet.pdf. Accessed 11 Feb 2021
Colceriu C, Riedl M, Henrich D, Brell-Cokcan S, Nitsch V (2020) User-centered design of an intuitive robot playback programming system. In: Schüppstuhl T, Tracht K, Henrich D (eds) Annals of scientific society for assembly, handling and industrial robotics. Springer Vieweg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-61755-7_18
Cordes A, Gehrke B (2015) Industrielle Entwicklung und funktionale Verschiebungen in Europa – Eine empirische Analyse ausgeübter Tätigkeiten. Vierteljahrshefte zur Wirtschaftsforschung 84: 79–101. https://doi.org/10.3790/vjh.84.1.79
Czerniak-Wilmes JN, Brandl C, Mertens AW, Schlick C (2017) Innovative Mensch-Maschine-Interaktionskonzepte für den Facharbeiter der Zukunft in der Produktion 4.0. In: Spöttl G, Windelband L (eds) Industrie 4.0: Risiken und Chancen für die Berufsbildung. Bertelsmann Verlag, Bielefeld, pp 171–188
Daheim C, Wintermann O (2016) 2050: Die Zukunft der Arbeit. Ergebnisse einer internationalen Delphi-Studie des Millennium Project. Bertelsmann Stiftung, Gütersloh
De Stefano VM (2016) The rise of the 'just-in-time workforce': on-demand work, crowd work and labour protection in the 'gig-economy'. Comp Labor Law Policy J 37(3):471–504
Dengler K, Matthes B (2015) Folgen der Digitalisierung für die Arbeitswelt: Substituierbarkeitspotenziale von Berufen in Deutschland. IAB-Forschungsbericht 11, Institute for Employment Research (IAB), Nuremberg
Dengler K, Matthes B (2018a) Substituierbarkeitspotenziale von Berufen: Wenige Berufsbilder halten mit der Digitalisierung Schritt. IAB-Kurzbericht 4, Institute for Employment Research (IAB), Nuremberg
Dengler K, Matthes B (2018b) The impacts of digital transformation on the labour market: substitution potentials of occupations in Germany. Technol Forecast Soc Change 137:304–316. https://doi.org/10.1016/j.techfore.2018.09.024
DIN EN 29241-2 (1993) Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten, Teil 2: Anforderungen an die Arbeitsaufgaben. Beuth Verlag, Berlin
DIN EN 614-2 (2008) Sicherheit von Maschinen – Ergonomische Gestaltungsgrundsätze, Teil 2: Wechselwirkungen zwischen der Gestaltung von Maschinen und den Arbeitsaufgaben. Beuth Verlag, Berlin

DIN EN ISO 6385 (2016) Grundsätze der Ergonomie für die Gestaltung von Arbeitssystemen. Beuth Verlag, Berlin
DIN EN ISO 9241-11 (2018) Ergonomie der Mensch-System-Interaktion – Teil 11: Gebrauchstauglichkeit: Begriffe und Konzepte. Beuth Verlag, Berlin
Dispan J (2017) Entwicklungstrends im Werkzeugmaschinenbau 2017. Working Paper Forschungsförderung, Nummer 29, Hans-Böckler-Stiftung, Düsseldorf
Dombrowski U, Palluk M, Schmidt S (2006) Strukturelle Analyse Ganzheitlicher Produktionssysteme. Zeitschrift für wirtschaftlichen Fabrikbetrieb 101(3):114–118
Duell W (1983) Partizipative Arbeitsgestaltung: Bedingungen erfolgreicher Intervention. Psychosozial 20:71–90
Düll N, Bertschek I, Dworschak B, Meil P, Niebel T, Ohnemus J, Vetter T, Zaiser H (2016) Arbeitsmarkt 2030 Digitalisierung der Arbeitswelt: Fachexpertise zur Prognose 2016, Munich
Dunckel H (ed) (1999) Handbuch psychologischer Arbeitsanalyseverfahren. vdf, Hochschulverl. an der ETH, Zurich
Eichhorst W, Spermann A (2015) Sharing Economy – Chancen, Risiken und Gestaltungsoptionen für den Arbeitsmarkt. IZA Research Report No. 69. http://ftp.iza.org/report_pdfs/iza_report_69.pdf. Accessed 11 Feb 2021
European Commission (2013) Options for strengthening: responsible research and innovation. European Union, Luxembourg. https://doi.org/10.2777/46253
Federal Ministry of Labour and Social Affairs (BMAS) (eds) (2016) Monitor Digitalisierung am Arbeitsplatz. Berlin. https://inqa.de/SharedDocs/downloads/webshop/monitor-digitalisierung-am-arbeitsplatz?__blob=publicationFile. Accessed 11 Feb 2021
Federal Ministry of Labour and Social Affairs (BMAS) (eds) (2017) Weißbuch Arbeiten 4.0. Berlin
Federal Ministry of Labour and Social Affairs (BMAS) (eds) (2018) Sicherheit und Gesundheit bei der Arbeit: Berichtsjahr 2017. Bericht der Bundesregierung über den Stand von Sicherheit und Gesundheit bei der Arbeit und über das Unfall- und Berufskrankheitengeschehen in der Bundesrepublik Deutschland im Jahre 2017, Berlin
Federal Ministry of Labour and Social Affairs (BMAS) (eds) (2020a) Digitalisierung im Arbeitsalltag von Beschäftigten: Konsequenzen für Tätigkeiten, Verhalten und Arbeitsbedingungen. Berlin. https://www.bmas.de/SharedDocs/Downloads/DE/Publikationen/Forschungsberichte/fb-555-digitalisierung-im-arbeitsalltag-von-beschaeftigten.pdf?__blob=publicationFile&v=1. Accessed 12 Feb 2021
Federal Ministry of Labour and Social Affairs (BMAS) (eds) (2020b) The federal ministry of labour and social affairs' thematic reader. Berlin. https://www.bmas.de/SharedDocs/Downloads/DE/Publikationen/a811-eu2020-thematic-reader.pdf;jsessionid=B56B3F6A3AA806768941B13B0695DFC0.delivery2-replication?__blob=publicationFile&v=1. Accessed 12 Feb 2021
Frenz M, Heinen S, Schlick CM (2015) Industrie 4.0: Anforderungen an Fachkräfte in der Produktionstechnik. Berufsbildung in Wissenschaft und Praxis (BWP) 44(6):2–16
Frey CB, Osborne MA (2013) The future of employment: how susceptible are jobs to computerization? Oxford Martin School, University of Oxford, Oxford
Fried Y, Ferris GR (1987) The validity of the job characteristics model – a review and meta-analysis. Pers Psychol 40:287–322
Fuchs J, Kubis A, Schneider L (2019) Zuwanderung und Digitalisierung: Wie viel Migration aus Drittstaaten benötigt der deutsche Arbeitsmarkt künftig? Bertelsmann Stiftung, Gütersloh
Geraldi J, Maylor H, Williams T (2011) Now, let's make it really complex (complicated). Int J Oper Prod Manag 31:966–990. https://doi.org/10.1108/01443571111165848
Gerlmaier A, Gül K, Hellert U, Kämpf T, Latniak E (eds) (2016) Praxishandbuch lebensphasenorientiertes Personalmanagement. Fachkräftepotenziale in technischen Entwicklungsbereichen erschließen und fördern. Springer Gabler, Wiesbaden
Greif S, Hamborg K-C (eds) (2018) Methoden der Arbeits-, Organisations- und Wirtschaftspsychologie, 1. Edition. Enzyklopädie der Psychologie Methodologie und Methoden Psychologische Interventionsmethoden, Vol. 3. Hogrefe, Göttingen

Grote G (2018) Gestaltungsansätze für das komplementäre Zusammenwirken von Mensch und Technik in Industrie 4.0. In: Hirsch-Kreinsen H, Ittermann P, Niehaus J (2018) Digitalisierung industrieller Arbeit. 2. Edition. Nomos Verlag, Baden-Baden, pp 215–231

Grote G, Wäfler T, Ryser C, Weik S, Zölch M, Windischer A (1999) Wie sich Mensch und Technik sinnvoll ergänzen: Die Analyse automatisierter Produktionssysteme mit KOMPASS. vdf, Hochschulverl. an der ETH, Zurich

Grote G, Ryser C, Wäfler T, Windischer A, Weik S (2000) KOMPASS: a method for complementary function allocation in automated work systems. Int J Hum-Comput Stud 52:267–287

Hacker W, Sachse P (2014) Allgemeine Arbeitspsychologie. 3. Edition, Hogrefe, Göttingen

Hackman JR, Oldham GR (1976) Motivation through the design of work: test of a theory. Organ Behav Hum Perform 16:250–279

Hammermann A, Stettes O (2014) Lebensphasenorientierte Personalpolitik: Theoretisches Konzept und empirisches Evidenz, Studie. Institut der deutschen Wirtschaft Köln e.V., Cologne

Hammermann A, Stettes O (2018) Digitaler Wandel: Chance für ein lebensphasenorientiertes Personalmanagement. IW-Kurzbericht 5. Institut der deutschen Wirtschaft Köln e.V., Cologne

Harlacher M, Nitsch V, Mütze-Niewöhner S (2021) Komplexität im Projektmanagement. In: Mütze-Niewöhner S, Hacker W, Hardwig T, Kauffeld S, Latniak E, Nicklich M, Pietrzyk U (eds) Projekt- und Teamarbeit in der digitalisierten Arbeitswelt. Springer, Berlin, Heidelberg, pp 55–73

Hartmann E (2015) Arbeitsgestaltung für Industrie 4.0: Alte Wahrheiten, neue Herausforderungen. In: Botthof A, Hartmann E (eds) Zukunft der Arbeit in Industrie 4.0. Springer Vieweg, Berlin, Heidelberg, pp 9–20

Herczeg M (2018) Software-Ergonomie: Theorien, Modelle und Kriterien für gebrauchstaugliche interaktive Computersysteme. 4. Edition, De Gruyter Studium

Hill NS, Bartol KM, Tesluk PE, Langa GA (2008) Organizational context and face-to-face interaction: influences on the development of trust and collaborative behaviors in computer-mediated groups. Organ Behav Hum Decis Process 108:187–201. https://doi.org/10.1016/j.obhdp.2008.10.002

Hirsch-Kreinsen H (2014) Wandel von Produktionsarbeit – “Industrie 4.0”. WSI-Mitteilungen 67: 421–429

Hirsch-Kreinsen H (2015) Entwicklungsperspektiven von Produktionsarbeit. In: Botthof A, Hartmann E (eds) Zukunft der Arbeit in Industrie 4.0. Springer Vieweg, Berlin, Heidelberg, pp 89–98

Hirsch-Kreinsen H (2016) Digitization of industrial work: development paths and prospects. J Labour Market Res 49:1–14. https://doi.org/10.1007/s12651-016-0200-6

Hirsch-Kreinsen H (2018) Einleitung: Digitalisierung industrieller Arbeit. In: Hirsch-Kreinsen H, Ittermann P, Niehaus J (2018) Digitalisierung industrieller Arbeit. 2. Edition. Nomos Verlag, Baden-Baden, pp 13–32

Hirsch-Kreinsen H, ten Hompel M (2015) Digitalisierung industrieller Arbeit: Entwicklungsperspektiven und Gestaltungsansätze. In: Vogel-Heuser B, Bauernhansl T, ten Hompel M (eds) Handbuch Industrie 4.0 Vol. 3: Logistik. Springer-Verlag, Berlin, Heidelberg, pp 357–376

Hirsch-Kreinsen H, Ittermann P, Niehaus J (2018) Digitalisierung industrieller Arbeit. 2. Edition. Nomos Verlag, Baden-Baden

Hoch JE, Kozlowski SWJ (2014) Leading virtual teams: hierarchical leadership, structural supports, and shared team leadership. J Appl Psychol 99:390–403. https://doi.org/10.1037/a0030264

Holler M (2017) Verbreitung, Folgen und Gestaltungsaspekte der Digitalisierung in der Arbeitswelt. Auswertungsbericht auf Basis des DGB-Index Gute Arbeit 2016. Berlin

Howe J (2006) Crowdsourcing: a definition. https://crowdsourcing.typepad.com/. Accessed 11 Feb 2021

Huws U, Spencer NH, Syrdal DS, Holts K (2018) Working in the gig economy. In: Neufeind M, O’Reilly J, Ranft F (eds) Work in the digital age: challenges of the fourth industrial revolution. London, pp 153–162

Institut DGB-Index Gute Arbeit (2016) Wie die Beschäftigten die Arbeitsbedingungen in Deutschland beurteilen: Mit dem Themenschwerpunkt: Die Digitalisierung der Arbeitswelt - Eine Zwischenbilanz aus der Sicht der Beschäftigten. DGB-Index Gute Arbeit Der Report 2016, Berlin

Jacobs JC, Kagermann H, Spath D (2017) Arbeit in der digitalen Transformation - Agilität, lebenslanges Lernen und Betriebspartner im Wandel. acatech DISKUSSION. Herbert Utz Verlag, Munich

Jodlbauer H, Schagerl M, Brunner M (2016) Industrie 4.0 versus Automatisierung. ZWF 111:222–224. https://doi.org/10.3139/104.111496

Johnen L, Brunner O, Hellig T, Nitsch V, Mertens AW, Brandl C (2018) Interaktives Erleben und Erlernen von Ergonomie mit ErgoCAM. In: Dachselt R, Weber RG (eds) Mensch und Computer 2018. Tagungsband, pp 989–993

Johns T, Gratton L (2013) The third wave of virtual work. Harv Bus Rev 91:66–73

Kagermann H, Wahlster W, Helbig J (eds) (2013) Securing the future of German manufacturing industry. Recommendations for implementing the strategic initiative INDUSTRIE 4.0. Final report of the Industrie 4.0 Working Group. Berlin. https://www.din.de/blob/76902/e8cac883f42bf28536e7e8165993f1fd/recommendations-for-implementing-industry-4-0-data.pdf. Accessed 12 Feb 2021

Kenny GP, Yardley JE, Martineau L, Jay O (2008) Physical work capacity in older adults. Implications for the aging worker. Am J Ind Med 51(8):610–625

Kinkel S, Lay G, Jäger A (2007) Mehr Flexibilität durch Organisation. Mitteilungen aus der ISI-Erhebung - Modernisierung der Produktion 42. Fraunhofer ISI, Karlsruhe

Kinkel S, Lay G, Jäger A (2008) Wandlungsfähigkeit durch Organisation. Industrie Management 24:78–82

Klocke F, Bassett E, Bönsch C, Gärtner R, Holstein S, Jamal R, Jurke B, Kamps S, Kerzel U, Mattfeld P, Shirobokov A, Stauder J, Stautner M, Trauth D (2017) Assistenzsysteme in der Produktionstechnik. In: Brecher C, Klocke F, Schmitt R, Schuh G (eds) Internet of Production für agile Unternehmen: AWK Aachener Werkzeugmaschinen-Kolloquium 2017, 18. bis 19. Mai, 1. Edition Apprimus Verlag, Aachen

Kropp P, Dengler K (2019) The impact of digital transformation on regional labour markets in Germany: substitution potentials of occupational tasks. In: Proceedings of the Weizenbaum Conference 2019 "Challenges of digital inequality - digital education, digital work, digital life", Berlin, pp 1–8. https://doi.org/10.34669/wi.cp/2.8

Latos BA (2020) Auswirkungen von Komplexität auf die Performance von Produktionsgruppen. Dissertation, RWTH Aachen University, Shaker, Aachen

Latos BA, Harlacher M, Przybysz PM, Mütze-Niewöhner S (2017a) Transformation of working environments through digitalization: exploration and systematization of complexity drivers. In: IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). IEEE, pp 1084–1088

Latos BA, Holtkötter C, Brinkjans J, Kalantar P, Przybysz PM, Mütze-Niewöhner S (2017b) Partizipatives und simulationsgestütztes Vorgehen zur Konzeption einer flexiblen und demografierobusten Montagelinie. Zeitschrift für Arbeitswissenschaft 72:90–98

Latos BA, Harlacher M, Burgert F, Nitsch V, Przybysz PM, Mütze-Niewöhner S (2018) Complexity drivers in digitalized work systems: implications for cooperative forms of work. Adv Sci Technol Eng Syst J 3:171–185

Leichtmann B, Schnös F, Rinck P, Zäh M, Nitsch V (2018) Work system analysis for the user-centered development of cooperative mobile robots. In: Gesellschaft für Arbeitswissenschaft (GfA), Dortmund (ed) Frühjahrskongress 2018, Frankfurt a. M. ARBEIT(s).WISSEN.SCHAF(f)T – Grundlage für Management & Kompetenzentwicklung

Leimeister JM, Zogaj S, Durward D (2015a) New forms of employment and IT -- Crowdsourcing (2015). 4th Conference of the Regulating for Decent Work Network. https://ssrn.com/abstract=2736953. Accessed 13 Feb 2021

Leimeister JM, Zogaj S, Durward D, Blohm I (2015b) Arbeit und IT: Crowdsourcing und Crowdwork als neue Arbeits- und Beschäftigungsformen. In: ver.di – Vereinte Dienstleistungsgewerkschaft (eds) Gute Arbeit und Digitalisierung: Prozessanalysen und Gestaltungsperspektiven für eine humane digitale Arbeitswelt, Berlin, pp 66–79

Lerch C, Jäger A, Maloca S (2017) Wie digital ist Deutschlands Industrie wirklich? Arbeit und Produktivität in der digitalen Produktion. Mitteilungen aus der ISI-Erhebung - Modernisierung der Produktion 71. Fraunhofer ISI, Karlsruhe

Lichtblau K, Stich V, Bertenrath R, Blum M, Bleider M, Millack A, Schmitt K, Schmitz E, Schröter M (2015) Industrie 4.0 Readiness. Aachen, Cologne. http://industrie40.vdma.org/documents/4214230/5356229/Industrie%204.0%20Readiness%20Study%20English.pdf/f6de92c1-74ed-4790-b6a4-74b30b1e83f0. Accessed 12 Feb 2021

Lindsjørn Y, Sjøberg DI, Dingsøyr T, Bergersen G, Dyba T (2016) Teamwork quality and project success in software development: a survey of agile development teams. J Syst Softw 122:274–286

Lott Y (2020) Work-Life Balance im Homeoffice: Was kann der Betrieb tun? Welche betrieblichen Bedingungen sind für eine gute Work-Life Balance im Homeoffice notwendig? WSI Report Nr. 54. WSI, Düsseldorf

Luczak H (1997) Task analysis. In: Salvendy G (ed) Handbook of human factors and ergonomics, 2nd edn. Wiley-Interscience, New York, pp 340–416

Luczak H, Volpert W (1987) Arbeitswissenschaft. Kerndefinition – Gegenstandskatalog – Forschungsgebiete. RKW-Verlag, Eschborn

Luczak H, Hinrichsen S, Mütze-Niewöhner S (2006a) Humans in work system environment. In: Marras WS, Karwowski W (eds) Fundamentals and assessment tools for occupational ergonomics. The occupational ergonomics handbook, 2nd edn. CRC Taylor & Francis, London, Great Britain, Chapter 5.1–5.52

Luczak H, Kabel T, Licht T (2006b) Task design and motivation. In: Salvendy G (ed) Handbook of human factors and ergonomics, 3rd edn. John Wiley & Sons, New Jersey, pp 384–427

Mayer C, Mütze-Niewohner S, Nitsch V (2020) Empirical classification of advanced information technology towards their support of leadership behaviors in virtual project management settings. In: 2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). IEEE, pp 275–279

Mertens A, Schäfer K, Rasche P, Bröhl C, Theis S, Seinsch T, Brandl C, Wille M (2018) Shaping ethical, legal and social implications of the digital revolution through participation: the new interdisciplinary research paradigm of Aachener DenkfabrEthik. In: Bagnara S, Tartaglia R, Albolino S, Alexander T, Fujita Y (eds) Proceedings of the 20th Congress of the International Ergonomics Association (AISC 824). Springer, Cham, pp 1493–1497

Morgeson FP, Humphrey SE (2008) Job and team design – toward a more integrative conceptualization of work design. Res Pers Hum Resour Manag 27:39–91

Morris L, Ma M, Wu PC (2014) Agile innovation: a revolutionary approach to accelerate success, inspire engagement, and ignite creativity. Wiley, Hoboken, New Jersey

Müller-Thur K, Angerer P, Körner U, Dragano N (2018) Work with digital technologies, psychosocial demands/stress and health implications: are there any correlations? ASU Arbeitsmed Sozialmed Umweltmed 53:388–391

Mütze-Niewöhner S, Hacker W, Hardwig T, Kauffeld S, Latniak E, Nicklich M, Pietrzyk U (eds) (2021a) Projekt- und Teamarbeit in der digitalisierten Arbeitswelt: Herausforderungen, Strategien und Empfehlungen. Springer Vieweg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-62231-5

Mütze-Niewöhner S, Latniak E, Hardwig T, Nicklich M, Hacker W, Harlacher M, Pietrzyk U, Kauffeld S (2021b) Projekt- und Teamarbeit in der digitalisierten Arbeitswelt. In: Mütze-Niewöhner et al (eds) Projekt- und Teamarbeit in der digitalisierten Arbeitswelt: Herausforderungen, Strategien und Empfehlungen. Springer Vieweg, Berlin, Heidelberg, pp 1–31. https://doi.org/10.1007/978-3-662-62231-5_1

Myskovszky von Myrow T, Mütze-Niewöhner S, Duckwitz S (2014) Towards a conceptual model of team-based innovation: enablers and obstacles of innovation productivity. In: Proceedings of the 2014 International Conference on Industrial Engineering and Operations Management, pp 1030–1036

Myskovszky von Myrow T, Przybsz PM, Mütze-Niewöhner S, Schlick CM (2015) Partizipatives Innovationsmanagement in altersgemischten Teams. In: Jeschke S, Richert A, Hees F, Jooß C (eds) Exploring Demographics. Transdisziplinäre Perspektiven zur Innovationsfähigkeit im demografischen Wandel. Springer Fachmedien, Wiesbaden, pp 541–553

Nedelkoska L, Quintini G (2018) Automation, skills use and training. OECD Social, Employment and Migration Working Papers. No. 202, OECD Publishing, Paris. https://doi.org/10.1787/2e2f4eea-en.

Neef A, Burmeister K (2005) Die Schwarm-Organisation – Ein neues Paradigma für das e-Unternehmen der Zukunft. In: Kuhlin B, Thielmann H (eds) Real-Time Enterprise in der Praxis. Springer-Verlag, Berlin, Heidelberg, pp 563–572

Nelles J, Kohns S, Spies J, Bröhl C, Brandl C, Mertens A, Schlick CM (2017) Best-practice approach for a solution-oriented technology assessment: ethical, legal, and social issues in the context of human-robot collaboration. In: Schlick CM, Duckwitz S, Flemisch F, Frenz M, Mertens A, Mütze-Niewöhner S (eds) Advances in ergonomic design of systems, products and processes: Proceedings of the Annual Meeting of GfA 2016. Springer, Berlin, Heidelberg, pp 1–14

Nerdinger FW, Blickle G, Schaper N (2014) Arbeits- und Organisationspsychologie. 3. Edition, Springer-Verlag Berlin, Heidelberg

Neumer J, Nicklich M (2021) Fluide Teams in agilen Kontexten – Grenzziehung und innere Strukturierung als Herausforderung für Selbstorganisation. In: Mütze-Niewöhner S et al (eds) Projekt- und Teamarbeit in der digitalisierten Arbeitswelt. Springer Vieweg, Berlin, Heidelberg, pp 31–53. https://doi.org/10.1007/978-3-662-62231-5_2

Neumer J, Porschen-Hueck S, Sauer S (2018) Reflexive scaling as a way towards agile organization. JIMS 18:27–38. https://doi.org/10.18374/JIMS-18-2.3

Niederman F, Lechler T, Petit Y (2018) A research agenda for extending agile practices in software development and additional task domains. Project Manag J 49:3–17. https://doi.org/10.1177/8756972818802713

O'Leary, Cummings (2007) The spatial, temporal, and configurational characteristics of geographic dispersion in teams. MIS Q 31:433. https://doi.org/10.2307/25148802

Oldham GR, Fried Y (2016) Job design research and theory: past, present and future. Organ Behav Hum Decis Process 136:20–35

Paetzold K, Nitsch V (2014) Beschreibung eines kompetenzorientierten Ansatzes für die Gestaltung technischer Unterstützungssysteme. In: Weidner R, Redlich T (eds) Erste Transdisziplinäre Konferenz zum Thema "Technische Unterstützungssysteme, die die Menschen wirklich wollen". HSU, Hamburg, pp 19–29

Pangert B, Pauls N, Schüpbach H (2016) Die Auswirkungen arbeitsbezogener erweiterter Erreichbarkeit auf Life-Domain-Balance und Gesundheit. 2. Edition. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Dortmund, Berlin, Dresden

Petruck H, Faber MD, Mertens AW, Brandl C, Nitsch V (2018) Mensch-Roboter-Kooperation in der Praxis. ASU Arbeitsmedizin, Sozialmedizin, Umweltmedizin 10:646–648

Petruck H, Mertens AW, Nitsch V (2019) Adaptive Mensch-Roboter-Kollaboration in der Montage: Manuelle Montage im Internet of Production: Unsicherheit in einer digital vorsimulierten Welt? wt Werkstattstechnik online 109(9):1436–4980

Pfeiffer, Sauer, Ritter (2019) Agile methods as stress management tools? An empirical study. Work Organ Labour Glob 13:20. https://doi.org/10.13169/workorgalaboglob.13.2.0020

Piele C, Piele A (2017) Flexible Arbeitszeiten - Arbeitszeitmodelle und Flexibilitätsanforderungen: Eine Analyse des verarbeitenden Gewerbes auf Basis der IG Metall-Beschäftigtenbefragung 2017. Fraunhofer IAO, Stuttgart

Pongratz HJ, Bormann S (2017) Online-Arbeit auf Internet-Plattformen: Empirische Befunde zum 'Crowdworking' in Deutschland. Arbeits- und Industriesoziologische Studien 10:158–181
Pritchard RD, Harrell MM, DiazGranados D, Guzman MJ (2008) The productivity measurement and enhancement system – a meta-analysis. J Appl Psychol 93:540–567
Reuschl AJ, Bouncken RB (2017) Coworking-Spaces als neue Organisationsform in der Sharing Economy. In: Bruhn M, Hadwich K (eds) Dienstleistungen 4.0. Springer Gabler, Wiesbaden, pp 185–208. https://doi.org/10.1007/978-3-658-17552-8_8
Reuter C, Gartzen T, Prote J-P, Fränken B (2016) Industrie-4.0-Audit. VDI-Z Integrierte Produktion, Ausgabe 26. VDI-Fachmedien. https://www.vdi-z.de/2016/Ausgabe-06/Forschung-und-Praxis/Industrie-4.0-Audit?page=2. Accessed 27 Feb 2019
Robert LP, You S (2018) Are you satisfied yet? Shared leadership, individual trust, autonomy, and satisfaction in virtual teams. J Assoc Inf Sci Technol 69:503–513. https://doi.org/10.1002/asi.23983
Rogers Y, Sharp H, Preece J (2015) Interaction design: beyond human-computer interaction, 4th edn. John Wiley, Hoboken, NJ, USA
Saleh SM, Huq SM, Rahman MA (2019) Comparative study within Scrum, Kanban, XP focused on their practices. In: 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE). IEEE, pp 1–6
Sauer S (2017) Agile project management as a generator of stress relief and efficiency? Int J Strateg Manag 17:67–74. https://doi.org/10.18374/IJSM-17-2.5
Schlick C, Mütze-Niewöhner S (2010) Editorial: Altersdifferenzierte Arbeitssysteme. Zeitschrift für Arbeitswissenschaft 64(3):133–135
Schlick CM, Mütze-Niewöhner S, Köttendorf N (2009) Unterstützung von zukunftsorientierten Unternehmensstrategien durch professionelle Demographie-Beratung. In: Speck P (ed) Employability - Herausforderungen für die strategische Personalentwicklung. 4., aktualisierte und erweiterte Edition Gabler, Wiesbaden, pp 43–60
Schlick C, Frieling E, Wegge J (eds) (2013) Age-differentiated work systems. Springer, Berlin
Schlick CM, Duckwitz S, Flemisch F, Frenz M, Mertens A, Mütze-Niewöhner S (eds) (2017) Advances in ergonomic design of systems, products and processes: Proceedings of the Annual Meeting of GfA 2016. Springer, Berlin, Heidelberg
Schlick CM, Bruder R, Luczak H (2018) Arbeitswissenschaft. 4. Edition, Springer, Berlin
Schmauder M, Spanner-Ulmer B (2014) Ergonomie: Grundlagen zur Interaktion von Mensch, Technik und Organisation, 1. Edition, REFA-Fachbuchreihe Arbeitsgestaltung. Carl Hanser, Munich
Schmidt K-H (2004) Förderung von Gruppenleistungen mit dem Partizipativen Produktivitätsmanagement (PPM). In: Schuler H (ed) Beurteilung und Förderung beruflicher Leistung. Hogrefe, Göttingen, pp 238–253
Schmidt FA (2017) Digital labour markets in the platform economy: mapping the political challenges of crowd work and gig work. In: Friedrisch-Ebert-Stiftung (eds). http://www.bollettinoadapt.it/wp-content/uploads/2020/10/13164.pdf. Accessed 12 Feb 2021
Schmidt L, Schlick CM, Grosche J (eds) (2008) Ergonomie und Mensch-Maschine-Systeme. Springer-Verlag, Berlin, Heidelberg
Schröder L, Schwemmle M (2014) Gute Arbeit in der Crowd? In: Schröder L, Urban H-J (eds) Jahrbuch Gute Arbeit. Bund-Verlag, Frankfurt, pp 112–124
Schuler H, Sonntag K (eds) (2007) Handbuch der Arbeits-und Organisationspsychologie. Hogrefe Verlag, Göttingen
Shephard R (2000) Aging and productivity. Some physiological issues. Int J Ind Ergon 25(5):535–545
Sherehiy B, Karwowski W, Layer JK (2007) A review of enterprise agility: concepts, frameworks, and attributes. Int J Ind Ergon 37:445–460. https://doi.org/10.1016/j.ergon.2007.01.007
Spath D, Ganschar O, Gerlach S, Hämmerle M, Krause T, Schlund S (2013a) Produktionsarbeit der Zukunft - Industrie 4.0: Studie. Fraunhofer Verlag, Stuttgart

Spath D, Ganschar O, Gerlach S, Hämmerle M, Krause T, Schlund S (2013b) Manufacturing work of the future – Industry 4.0: management summary. Fraunhofer, Stuttgart. https://www.produktionsarbeit.de/content/dam/produktionsarbeit/de/documents/Management_Summary.pdf. Accessed 12 Feb 2021

Spector PE (1986) Perceived control by employees: a meta-analysis of studies concerning autonomy and participation at work. Hum Relat 39:1005–1016

Steireif N, Schirmer M, Schnitzler M, Mütze-Niewohner S (2020) Towards an extended team model for agile development of complex products. In: 2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). IEEE, pp 280–284

Stettes O (2018) Keine Angst vor Robotern: Beschäftigungseffekte der Digitalisierung - eine Aktualisierung früherer IW-Befunde. IW-Report 11, Cologne

Stranzenbach R (2018) Analyse und Gestaltung von Gruppenarbeit in ganzheitlichen Produktionssystemen. Dissertation, RWTH Aachen University, Shaker, Aachen

Strohm O, Ulich E (eds) (1997) Unternehmen arbeitspsychologisch bewerten. Ein Mehrebenenansatz unter besonderer Berücksichtigung von Mensch, Technik und Organisation. vdf, Hochschulverl. an der ETH, Zurich

Susman GI (1976) Autonomy at work. A sociotechnical analysis of participative management. Praeger, New York

Tempel J, Ilmarinen J (2013) Arbeitsleben 2025: Das Haus der Arbeitsfähigkeit im Unternehmen bauen. VSA-Verlag, Hamburg

Theis S, Pfendler C, Alexander T, Mertens A, Brandl C, Schlick CM (2016) Head-Mounted Displays - Bedingungen des sicheren und beanspruchungsoptimalen Einsatzes: Physische Beanspruchung beim Einsatz von HMDs. 1. Edition. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Dortmund

Uhlmann E, Hohwieler E, Kraft M (2013) Selbstorganisierende Produktion mit verteilter Intelligenz: Intelligente Werkstücke steuern ihren Weg durch die Fertigung. wt Werkstattstechnik online 103:114–117

Ulich E (2011) Arbeitspsychologie. 7. Edition Schäffer-Poeschel Verlag, Stuttgart

VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik (GMA) (ed) (2016) Statusreport: Arbeitswelt Industrie 4.0. https://www.vdi.de/ueber-uns/presse/publikationen/details/arbeitswelt-industrie-40. Accessed 12 Feb 2021

ver.di - Bereich Innovation und Gute Arbeit (ed) (2015a) Gute Arbeit und Digitalisierung: Prozessanalysen und Gestaltungsperspektiven für eine humane digitale Arbeitswelt, ver.di – Vereinte Dienstleistungsgewerkschaft Bereich Innovation und Gute Arbeit, Berlin

ver.di (2015b) 'Good work' and good services in the digital world. Resolution of the 4th Federal Congress of the United Services Union ver.di, September 2015. verdi-congress-resolution2015-good_work_and_good_services_in_digital_world.pdf. Accessed 13 Feb 2021

Vogel-Heuser B, Bauernhansl T, ten Hompel M (eds) (2018) Handbuch Industrie 4.0: Bd. 1: Produktion, 2., erweiterte und bearbeitete Edition Springer Reference Technik. Springer Vieweg, Berlin

Vogler-Ludwig K, Düll N, Kriechel B (2016) Arbeitsmarkt 2030: Wirtschaft und Arbeitsmarkt im digitalen Zeitalter, Prognose 2016, Kurzfassung: Analyse der zukünftigen Arbeitskräftenachfrage und des -angebots in Deutschland auf Basis eines Rechenmodells. Economix Resarch & Consulting, Munich

Wäfler T, Windischer A, Ryser C, Weik S, Grote G (1999) Wie sich Mensch und Technik sinnvoll ergänzen. Die Gestaltung automatisierter Produktionssysteme mit KOMPASS. vdf, Hochschulverl. an der ETH, Zurich

Wäfler T, Grote G, Windischer A, Ryser C (2003) KOMPASS: a method for complementary system design. In: Hollnagel E (ed) Handbook of cognitive task design. Mahwah, pp 477–502

Weber E (2016) Industry 4.0: job-producer or employment-destroyer? Aktuelle Berichte No 2/2016. Institute for Employment Research (IAB), Nuremberg. https://www.econstor.eu/bitstream/10419/161710/1/aktueller_bericht_1602.pdf. Accessed 13 Feb 2021

Weber E (2017) Digitalisierung als Herausforderung für eine Weiterbildungspolitik. Wirtschaftsdienst 97:372–374. https://doi.org/10.1007/s10273-017-2146-x
Weber E, Zika G, Wolter MI, Maier T (2017a) Wirtschaft 4.0 und die Folgen für die künftige Berufsfeldstruktur. In: Möller J, Walwei U (eds) Arbeitsmarkt kompakt: Analysen, Daten, Fakten. IAB-Bibliothek 363, Institute for Employment Research (IAB), Nuremberg, W. Bertelsmann Verlag, Bielefeld, pp 118–119
Weber E, Zika G, Wolter MI, Maier T (2017b) Wirtschaft 4.0 und die Folgen für die künftige Anforderungsstruktur. In: Möller J, Walwei U (eds) Arbeitsmarkt kompakt: Analysen, Daten, Fakten. IAB-Bibliothek 363, Institute for Employment Research (IAB), Nuremberg, W. Bertelsmann Verlag, Bielefeld, pp 120–121
WHO – World Health Organisation (1994) Aging and Work Capacity – Altern und Arbeit. WHO Technical Report Series 835. In: Bundesanstalt für Arbeitsmedizin (ed) Schriftenreihe der Bundesanstalt für Arbeitsmedizin, Übersetzung 3. Wirtschaftsverlag, NW Bremerhaven
Wille M (2016) Head-Mounted Displays - Bedingungen des sicheren und beanspruchungsoptimalen Einsatzes: Psychische Beanspruchung beim Einsatz von HMDs. 1. Edition. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Dortmund
Wille C, Müller N (2018) Gute agile Arbeit: Gestaltungsempfehlungen aus dem Projekt diGAP. Fact-Sheet 5, Vereinigte Dienstleistungsgewerkschaft, Bereich Innovation und Gute Arbeit, Berlin
Wille M, Mertens A, Adolph L, Theis S, Wischniewski S, Schlick C (2015) Regression analysis of age effects during prolonged work with head-mounted displays. Procedia Manuf 3:4322–4329
Wille M, Brandl C, Nelles J, Mertens AW, Schlick C (2016) AMICAI: Ein methodisches Vorgehen zur quantitativen Analyse von ethischen, rechtlichen und sozialen Auswirkungen anwendungsnaher Forschungsprojekte. In: Arbeit in komplexen Systemen - digital, vernetzt, human?! 62. Kongress der Gesellschaft für Arbeitswissenschaft, RWTH Aachen University, Institut für Arbeitswissenschaft (IAW), 2.-4. März 2016 / Gesellschaft für Arbeitswissenschaft e.V., Dortmund
Wilson AJ, Daugherty PR (2018) Collaborative intelligence: humans and AI are joining forces. Harv Bus Rev. July–August 2018. https://www.accenture.com/t20180828T222720Z__w__/ae-en/_acnmedia/PDF-84/Accenture-Collaborative-Intelligence-2018.pdf. Accessed 13 Feb 2021
Wischmann S, Hartmann EA (2018a) Zukunft der Arbeit in Industrie 4.0 – Szenarien aus Forschungs- und Entwicklungsprojekten. In: Wischmann S, Hartmann EA (eds) Zukunft der Arbeit – Eine praxisnahe Betrachtung. Springer Vieweg, Berlin, Heidelberg, pp 1–7. https://doi.org/10.1007/978-3-662-49266-6_1
Wischmann S, Hartmann EA (2018b) Prognostizierte Veränderungen der gestaltbaren Arbeitssystemdimensionen. In: Wischmann S, Hartmann EA (eds) Zukunft der Arbeit – Eine praxisnahe Betrachtung. Springer Vieweg, Berlin, Heidelberg, pp 9–31. https://doi.org/10.1007/978-3-662-49266-6_2
Wischmann S, Hartmann EA (eds) (2018c) Zukunft der Arbeit – Eine praxisnahe Betrachtung. Springer Vieweg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-49266-6
Wischmann S, Wangler L, Botthof A (2015) Industrie 4.0: Volks- und betriebswirtschaftliche Faktoren für den Standort Deutschland, Studie im Rahmen der Begleitforschung zum Technologieprogramm AUTONOMIK für Industrie 4.0. Bundesministerium für Wirtschaft und Energie (ed), Berlin. https://vdivde-it.de/system/files/pdfs/industrie-4.0-volks-und-betriebswirtschaftliche-faktoren-fuer-den-standort-deutschland.pdf. Accessed 11 Feb 2021

Wolter MI, Mönning A, Hummel M, Weber E, Zika G, Helmrich R, Maier T, Neuber-Pohl C (2016) Economy 4.0 and its labour market and economic impacts: scenario calculations in line with the BIBB-IAB qualification and occupational field projections. IAB-Forschungsbericht No. 13/2016. Institute for Employment Research (IAB), Nuremberg. http://doku.iab.de/forschungsbericht/2016/fb1316_en.pdf. Accessed 13 Feb 2021

Zika G, Helmrich R, Maier T, Weber E, Wolter MI (2018) Regionale Branchenstruktur spielt eine wichtige Rolle: Arbeitsmarkteffekte der Digitalisierung bis 2035. IAB-Kurzbericht 9, Institute for Employment Research (IAB), Nuremberg

Zimmermann V (2017) 2017 Digitalisierung der Wirtschaft: breite Basis, vielfältige Hemmnisse. Unternehmensbefragung. KfW Bankengruppe (ed), Frankfurt am Main. https://www.kfw.de/PDF/Download-Center/Konzernthemen/Research/PDF-Dokumente-Unternehmensbefragung/Unternehmensbefragung-2017-—-Digitalisierung.pdf. Accessed 12 Feb 2021

Zink KJ (2015) Beteiligung und Integration als Voraussetzung erfolgreicher Veränderung. In: Zink K, Kötter W, Longmuß J, Thul M (eds) Veränderungsprozesse erfolgreich gestalten. VDI-Buch. Springer Vieweg, Berlin, Heidelberg, pp 1–10

The Relevance of Labour Law in the Process of Industry 4.0

Rüdiger Krause

Contents

1 Industry 4.0, Work 4.0, Labour Law 4.0

The keywords Industry 4.0, Work 4.0 and Labour Law 4.0 represent fundamental changes in the economy, work and labour law. While **Industry 4.0** refers to the concept of end-to-end digital networking of the industrial value chain ('smart factory'), which was launched a couple of years ago (Kagermann et al. 2011; Forschungsunion/acatech 2013; Spath 2013; BITKOM/VDMA/ZVEI 2015; Vogel-Heuser et al. 2017), **Work 4.0** concerns the comprehensive **transformation of the world of work**, which goes far beyond the industrial sector and, while it is largely driven by **digitalisation** (Brandt et al. 2016; Eurofound 2018; Funken and Schulz-Schaeffer 2008; IBA Global Employment Institute 2017; Stettes 2016), it also encompasses other upheavals such as demographic change or cultural change

R. Krause (✉)
Chair of Civil Law and Labor Law, Georg-August-University Göttingen, Göttingen, Germany
e-mail: lehrstuhl.krause@jura.uni-goettingen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_53

(BMAS 2015, pp. 14 et seq.; BMAS 2016, pp. 18 et seq.). Finally, the buzzword **Labour Law 4.0** is intended to bundle the effects of these processes on labour law in a handy way (Arnold and Günther 2018; Baker McKenzie 2017; Benecke 2018b; Däubler 2016a; Giesen and Kersten 2017; Günther and Böglmüller 2015; Hanau 2016; Haußmann 2019; Kramer 2019; Krause 2016, 2017; Oetker 2016; Thüsing 2016; Uffmann 2016a).

This article, which deals specifically with the relationship between Industry 4.0 and (German and European) labour law, thus singles out a specific section of the current debate, which is all about the future of the economy and work. The consequences of Industry 4.0 and the use of increasingly powerful industrial robots associated with this for the labour *market* are therefore excluded from the outset (Boston Consulting Group 2019; IAB 2015; World Economic Forum 2018, pp. 7 et seq.), as this is a topic that precedes labour law as a regulation of existing or prospective employment relationships. Furthermore, the central question of the concrete **organisation of work** in company practice and thus in the implementation of Industry 4.0 is not addressed as such, but only to the extent that labour law sets substantial or—in the form of the works council's participation rights—procedural regulatory barriers. Finally, essentially only the statutory labour law framework will be addressed, although labour law in the broader sense also includes collective bargaining standards, i.e. **collective agreements** and **works agreements**, and the successful implementation of the Industry 4.0 concept not only depends on solving technical and organisational problems but can hardly succeed without a collective bargaining framework at the sectoral and, above all, at the company level to ensure acceptance and support of the changes on the part of the workforce (cf. Hirsch-Kreinsen 2014, pp. 427 et seq.). At least it should be pointed out in this context that some German trade unions, in particular the Industrial Union of Metalworkers (IG Metall), increasingly try to promote the conclusion of company-focussed so-called 'Zukunftsvereinbarungen' ('future agreements') in order to strengthen the influence of employees on the process of the introduction of Industry 4.0 (Bosch et al. 2020; Haipeter et al. 2018; Röder et al. 2021).

However, the focus on the labour law framework of Industry 4.0 would still be boundless because there is no special labour law related to this model (see, however, the demand by Schimmelpfennig 2016, p. 70), but rather the entire labour law is applicable in the affected companies. Accordingly, many of the labour law issues discussed in connection with Industry 4.0 in particular (cf. Arnold and Günther 2018; Giesen et al. 2016; Pfrogner 2018) are not limited to the industrial sector, but also affect other sectors of the economy. Nevertheless, several labour law issues can be identified that are particularly affected by the technical and organisational changes envisaged by Industry 4.0.

For a better understanding of the labour law implications of Industry 4.0, the key features of this concept should be recalled in advance (see Henseler-Unger 2017; Hirsch-Kreinsen 2014; Hirsch-Kreinsen et al. 2018; Paul 2016; Simon 2016, 2018; Vogel-Heuser et al. 2017). As already outlined at the beginning, this is about linking the entire chain of industrial value creation through information and communication technology, from product development to manufacturing and the assembly of components to sales, service and recycling, with the aim of using all available capacities as efficiently as possible. At the centre of the guiding principle, which is thus related to the entire life cycle of products, is the connection of the virtual computer world with real production resources to form **cyber-physical systems**. Specifically, all machines, robots, equipment, transport and storage systems, work-pieces, assistance systems, etc. associated with industrial production should collect and process data largely autonomously and exchange it with each other in real time (**Internet of Things**) to make production more flexible, faster and more resilient (on the limits of the digitalisation of work from an occupational science perspective, Huchler 2017).

However, from the perspective of labour law, it is not technological innovations as such that are significant. Rather, the topic concerning labour law is the various effects of this transformation on employees. There is no question that there will be such effects, since Industry 4.0 is not about the 'deserted factories' often conjured up in the earlier debate on computer-integrated manufacturing, which would remove the ground from labour law, but about an intensive interaction between workers and digitalised production systems (Forschungsunion/acatech 2013, pp. 61, 92, 99; Botthoff and Hartmann 2015; Seidel 2017, pp. 15 et seq.; Wischmann and Hartmann 2018), which places considerable demands on a humane design of work. For example, the very keyword of **assistance systems** shows that Industry 4.0 will in many cases lead to interaction between workers and a digitally designed working environment, in which personal data is processed in addition to purely company-related data to increase efficiency. In all of this, the digital networking of value chains is to take place vertically and horizontally, both transnationally and across companies or groups, i.e. also integrating suppliers and customers or machine manufacturers and technology users, which raises additional questions and challenges in view of the traditional approach of labour law to normatively order the legal relationships between the employer and the individual employee established by the employment contract or the legal relationships between the employer and the works council as representative of the workforce of a company.

The implementation of Industry 4.0 at the company level is still in its infancy in many cases (Evers and Oberbeck 2017, pp. 62 et seq.; Pfeiffer and Huchler 2018, pp. 170 et seq.). Nevertheless, it makes sense to assess the relevance of the changes envisaged by this concept in terms of labour law. This is true even more as labour law does not merely act as a 'repair shop' to correct changes in the world of work that have already taken place, but rather to steer the actions of the actors in relation to the employees along certain paths from the outset. In this context, the various problems relevant to labour law can be systematically assigned to three major areas, which admittedly partially overlap: First, there are issues related to the concrete use of labour (Sect. 2), second, the protection of employees from specific risks in the context of Industry 4.0 (Sect. 3) and third, the institutionalised influence of the employees' side on the processes of change in the company associated with this model, which can also include the issue of ongoing qualification (Sect. 4). However, for reasons of space, the question of whether the digitalisation of value chains will also lead to changes in the area of remuneration, in particular through an increase in performance-related remuneration components, and what framework conditions labour law will provide in this respect, will not be dealt with (see Uffmann 2016b, pp. 31 et seq.). The same applies to the phenomenon of (external) crowdworking, which is only loosely connected to Industry 4.0 as a new form of coordination of gainful employment (on the practical irrelevance to date, see Bosch et al. 2017, p. 9), even if a connecting line is sometimes drawn in labour law scholarship (Neighbour 2017, paras. 9 et seq.). Finally, the numerous questions that can arise from the cross-border nature of cases relevant to labour law are not to be addressed here.

2 Use of the Workforce in Connection With Industry 4.0

The linking of all links in the entire value chain through information and communication technology envisaged by Industry 4.0 means that, on the associated integration of labour, two labour law issues particularly come into focus, namely working time law (Sect. 2.1) and the right of the employer to issue instructions (Sect. 2.2).

2.1 Limits on Working Time

The relevant working time limits are found in the **Working Time Act** (Arbeitszeitgesetz = ArbZG), which in turn is largely determined under European law by the **Working Time Directive** 2003/88/EC. From the multitude of individual regulations contained therein, two points of contention in particular have come to the fore in recent discussions, namely, on the one hand, the maximum **daily working time** and, on the other, the **daily rest period**. As far as the maximum working time on a working day is concerned, German law in principle provides for a maximum of 8 h (Section 3 sentence 1 ArbZG). An extension to 10 h is possible with appropriate compensation (Section 3 sentence 2 ArbZG), a further extension is generally only possible under the twofold condition that this is permitted in a collective agreement or based on a collective agreement in a works agreement and that the working time includes a considerable amount of standby duty or on-call duty (Section 7(1) no. 1 lit. a, (2a) ArbZG). In contrast, European working time law only recognises a maximum weekly working time of 48 h (Article 6 Directive 2003/88/EC), thus in principle allowing longer working days than German working time law. On the other hand, both legal levels agree on the question of the minimum daily rest period, in that they both require 11 h in principle (Section 5(1) ArbZG and Article 3 RL 2003/88/EC).

In particular, the cross-time-zone linking of value chains in real time and the associated need for communicative coordination processes in relation to manufacturing plants or suppliers or customers in Asia on the one hand and America on the other has led to calls from the employers' side to adapt the more rigid German regulation of maximum working hours on workdays to the more liberal European regulation in this respect (BDA 2015, pp. 6 et seq.; BDA 2018, p. 23; vbw 2016, pp. 7 et seq.; also Sachverständigenrat 2017, paras. 78, 777). In addition, a debate has arisen as to whether every work-related activity outside regular working hours, such as a brief call from a superior, colleague or customer as well as checking e-mails in the evening, leads to an interruption of the rest period with the consequence that a minimum rest period of 11 h begins to run again, i.e. a corresponding activity at 11 p.m. means not being allowed to start work again until 10 a.m. the next day. In this respect, some authors argue for softening of the ArbZG in two ways, namely, first for the irrelevance of **short interruptions of the rest period**, whereby partial disturbances of the employee of up to 15 min are apparently declared to be harmless in terms of rest period without any further restrictions on their location and frequency (Arnold and Winzer 2018, paras. 28 et seq.; Kramer 2019, B paras. 970 et seq., in each case with further references). Second, work-related activities of employees that have neither been expressly nor impliedly ordered by the employer are to be excluded from the scope of application of working time law as **'voluntary work'** from the outset (Arnold and Winzer 2018, paras. 38 et seq.; Kramer 2019, B paras. 972 et seq., in each case with further references). Both theses are of considerable importance due to the increasing prevalence of **mobile work**, which, although not exclusively taking place within the framework of Industry 4.0, certainly plays a role in this concept (cf. Forschungsunion/acatech 2013, p. 27) and, moreover, has received a considerable boost from the COVID-19 pandemic which has triggered the repeated temporary legal introduction of mandatory home office.

However, these relativisations of the current working time law are not convincing (see Krause 2016, pp. B 35 et seq., B 41 et seq. with further references). After all, despite the practical relevance of the topic, it is a theoretical discussion insofar as cases in which these questions are raised are not usually brought before the courts, i.e. the actors regularly refrain from settling internal working time conflicts externally. On the legislative level, too, there are no signs of a general liberalisation of the ArbZG, especially since a general reduction of the statutory rest period is blocked by Union law anyway. Nevertheless, the **coalition agreement** between the current governing parties SPD, BÜNDNIS 90/DIE GRÜNEN and FDP of 2021 provides for a limited possibility to deviate from the currently existing regulations of the Working Time Act if collective agreements or works agreements based on collective agreements allow this within the framework of an **experimentation clause** (Coalition Agreement 2021, p. 68; already in this direction BMAS 2016, pp. 124 et seq.). However, the context of the political announcement makes it clear that the aim is not merely to adapt employees' personal working time more closely to business requirements, for example in the course of Industry 4.0, but also and especially to improve the compatibility of employees' professional and private needs (**work-life balance**).

Remarkably, the 2018 **collective agreement of the metal and electrical industry on mobile work** contains a loosening of the rest period regulations, which provides for a reduction of the rest period between the end and the resumption of the daily working time to as low as 9 h if the employee can determine the end on the day in question or the start of the daily working time on the following day himself. The parties to the collective agreement in Germany's most important industrial sector have thus indicated that they would accommodate flexibility needs to a certain extent if the flexibility granted adequately balances the interests of both parties to the employment contract. At the company level, a pragmatic approach to working time law seems to prevail. Complaints by employees directly related to compliance with the limits set by working time law are practically non-existent. Industrial sociological surveys suggest that works councils which can complain to the employer about individual violations of the Working Time Act, invoking their general right of supervision under Section 80(1) no. 1 of the Works Constitution Act (Betriebsverfassungsgesetz = BetrVG), and press for redress (BAG, judgment of 21 March 2017 – 7 ABR 17/15, NZA 2017, 1014, para. 17) tolerate singular but counteract frequent violations of maximum working hours and minimum rest periods, which are regularly an expression of a disproportion between the allocation of tasks and staffing. However, it is not only between the employer and the employee that conflicts arise in working time law. Rather, the works council's claim to control working time sometimes collides with the, at least temporary, working time wishes of some employees themselves, who do not always see the sense of a strict legal limit on work-related activities because deadlines are pressing, projects have to be completed and target agreements have to be met (cf. Möhring-Hesse 2007). Working time law is therefore facing an increasing erosion process in a digitalised world of work (Däubler 2016b, p. 331).

2.2 Inter-Employer and Depersonalised Instructions

Another problem specific to labour law that may arise because of Industry 4.0 solutions concerns the employer's right to issue **instructions**, which is inherent in every employment contract. On the one hand, it is conceivable that suppliers or customers will directly access the work capacity of individual employees as a result of a networked value chain. On the other hand, the issuing of work-related instructions by assistance systems that are not limited to the information technology support of human decisions, but in which complex and adaptive algorithms make autonomous decisions based on **'artificial intelligence'** (on the conceptual foundations BITKOM 2017, p. 9; Enquete Commission 2020, pp. 48 et seq.; Ernst 2017, p. 1027; Kirn and Müller-Hengstenberg 2014; Niederée und Nejdl 2020; Reichwald and Pfisterer 2016). The latter is apparently no longer a dream of the future. For example, it is reported that a Japanese electronics company (Hitachi) has begun to use artificial intelligence systems to assign work tasks to employees (BITKOM 2017, p. 25; Wildhaber 2016, pp. 330 et seq.).

The normative basis for the employer's right of direction is Section 106 of the Trade, Commerce and Industry Code (Gewerbeordnung = GewO). According to this, the employer is entitled to determine in more detail the employee's duty to perform, which is only outlined in the employment contract. Of course, the employer does not have to exercise the right to issue instructions in person but can be represented by company employees (superiors). The only decisive factor is that the person specifically exercising the right to issue instructions can rely on a legitimation that goes back to the employer as a party to the employment contract. Since the individual instruction can be qualified as a unilateral declaration of intent requiring receipt (BAG, judgment of 16 April 2015 – 6 AZR 242/14, NZA-RR 2015, 532, para. 24), the attribution in these cases results from an application of the right of representation under Sections 164 et seq. of the German Civil Code (Bürgerliches Gesetzbuch = BGB). If instructions are issued by a third party not belonging to the company (supplier, customer), the legal situation is somewhat more complicated. In principle, representation is legally permissible in such a situation. However, the third party will regularly act in its own name and not in the name of the contracting employer. In addition, there will usually be no corresponding authorisation on the part of the contracting employer. The prerequisite of a transfer of legal power to the supplier or customer will generally also prevent an authorisation of the third party to exercise the right of direction, which is generally recognised by case law (cf. BAG, judgment of 8 August 1958 – 4 AZR 173/55, BAGE 6, 232, 243). Furthermore, due to the fundamental non-transferability of the claim to work performance under Section 613 sentence 2 BGB, this would require a corresponding provision in the employment contract. Therefore, the only legal construction that remains is for the third party and the contractual employer to agree that the third party is entitled to make declarations to the contractual employer that concretise performance within the framework of the relationship with the contractual employer, which is usually based on a contract of sale or a contract for work and services, while the contractual employer in turn expresses to its employees that it has adopted the third party's work orders as a blanket fulfilment of an instruction that it has issued itself. Since the number of such transactions will be limited despite the production-related linkage of the different companies, since there will be no integration of the employee concerned into the operations of the supplier or customer and, above all, since there is unlikely to be constant access to the same employee, there is also no risk that this process will be qualified as covert provision of workers to the third party within the meaning of Section 1(1) of the Temporary Employment Act (Arbeitnehmerüberlassungsgesetz = AÜG). Accordingly, there has not yet been sufficient reason to approach the construct of a multipolar employment relationship (Bücker 2016; Krause 2017, p. 35).

On **instructions by robots** or other technical systems with artificial intelligence (**management by algorithm**), the first relevant opinions are to the effect that such a form of exercising the right of direction is permissible from a purely labour law perspective (Arnold and Winzer 2018, para. 236; Däubler 2020, § 10 para. 10; Dzida 2016, p. 146; Groß and Gressel 2016, p. 994; Günther and Böglmüller 2017, pp. 55 et seq.; Neighbour 2017, para. 60; Wildhaber 2016, p. 331). Here, it may be left to one's own devices whether the corresponding instructions of autonomous systems are directly attributed to the employer in accordance with the traditional view or whether, in line with newer statements in the literature, the right of representation is also used in this respect (Schirmer 2016, pp. 663 et seq.; Teubner

2018, pp. 177 et seq.). In any case, the chain of legitimation reaches back to the employer. In this context, it is sometimes seen as problematic that an instruction under Section 106 GewO must comply with equitable discretion, but that a robot cannot currently perform such a weighing (Arnold and Winzer 2018, para. 237; Lembke 2020, Section 106 GewO para. 6; Schiefer and Worzalla 2019, p. 1905). However, since an instruction is not already inadmissible because no corresponding mental operations have been carried out, but only if the result is inequitable (BAG, judgment of 30 November 2016 – 10 AZR 11/16, NZA 2017, 1394, para. 28), this aspect will not play a role in the majority of cases, especially since the employee is entitled to ignore the instruction in such a constellation (BAG, judgment of 18 October 2017 – 10 AZR 330/16, BAGE 160, 296).

More important, however, is an objection resulting from data protection law. If one puts the general data protection requirements for Industry 4.0 applications back for a moment (see Sect. 3.2), the question is whether a fundamental prohibition on issuing instructions in an automated form can be derived from Article 22 of the **General Data Protection Regulation** (GDPR). However, such a prohibition can only be considered if personal data is processed, i.e. intelligent software does not merely use machine or operational data and generates a work order from it in the abstract and without regard to the respective employee, so that the models of **digital work management** practised to date (see Kuhlmann et al. 2018) should not be problematic from a data protection perspective. However, if, for example, the current location of the employee in the company is included in the instruction for a maintenance order, this requirement is fulfilled. If, in addition, an assessment of individual personality aspects such as the previous work performance of the employee or his or her reliability in performing the task is carried out to calculate ‘optimal’ instructions, this is in fact **profiling** (cf. Article 4 no. 4 GDPR), which also falls within the scope of application of the provision under the narrowest interpretation of Article 22(1) GDPR (cf. Abel 2018, p. 305; Buchner 2020, Article 22 GDPR paras. 20 et seq. in each case with further references). In these cases, it therefore depends on whether the issuance of an instruction can be qualified as a ‘decision’ that has legal effect towards the employee or significantly affects him in a similar way. This is sometimes denied on the grounds that an instruction does not reshape the employment relationship, but merely concretises the duty to work. A legally disadvantageous decision is only to be seen in a warning or in a termination as a reaction of the employer to the (repeated) non-compliance with an algorithm-based instruction (Arnold and Winzer 2018, para. 238; Günther and Böglmüller 2017, p. 56). However, since, as mentioned, an instruction represents a declaration of intent by which the duty to work, which was initially only outlined in abstract terms, is determined, so that it is now in principle legally binding which activity the employee is specifically obliged to perform, it is difficult to doubt a legal effect of the instruction generated by artificial intelligence (Däubler 2020, § 10 para. 12; Neighbour 2017, paras. 61 et seq.; also, apparently Giesen 2018, p. 442; Groß and Gressel 2016, p. 994). This corresponds to the fundamental meaning and purpose of Article 22(1) GDPR, which is to prevent individuals from becoming the object of opaque automated decision-making processes to protect their privacy (Buchner

2020, Article 22 of the GDPR, para. 1; Ernst 2017, pp. 1030 et seq.; Hoeren and Niehoff 2018, p. 53; Martini 2017, p. 1019), an idea that is also relevant in the employment relationship and may not be set aside merely to make work processes more efficient. Still unclear is an exceptional admissibility under the aspect of necessity for the fulfillment of the employment contract under Article 22(2)(a) GDPR (for this, on AI-generated instructions concerning the daily workflow Schwarze 2020, § 8 para. 34; similar Höpfner and Daum 2021, pp. 485 et seq.; in contrast Wedde 2020, p. 38, there also on the meaning of Article 22(2)(b) GDPR in conjunction with Section 31 of the Federal Data Protection Act (Bundesdatenschutzgesetz = BDSG)). Admissibility based on the express **consent** of the employee concerned under Article 22(2)(c) GDPR is likely to fail in most cases due to the lack of voluntariness (cf. Article 7(4) GDPR). In contrast, there would be no problems with regard to Article 22(1) GDPR if the automatically generated 'instructions' were legally qualified merely as non-binding recommendations for optimal task performance, with which the non-compliance by the employee would not entail any negative consequences (Klebe 2019, p. 134).

3 Protection Interests of Employees in the Context of Industry 4.0

Another immediately adjacent area concerns specific protection interests of employees in connection with Industry 4.0 applications. In this respect, technical occupational safety and health at the man-machine interface (Sect. 3.1) and the broad field of employee data protection (Sect. 3.2) should be mentioned.

3.1 Occupational Safety and Health at the Man-Machine Interface

The increasing use of robots that are expected to 'leave their cages' and collaborate directly with workers (**advanced robotics**) raises questions of technical occupational safety and health (see May 2014 for more details). In this respect, the manufacturer is already called upon to consider the health hazards for the operating personnel in addition to the general product safety requirements when designing and manufacturing robots that are often individually adapted (Section 2 no. 12 Machinery Ordinance in conjunction with Annex I Directive 2006/42/EC No. 1.1.6.). In addition to the general requirements of the **Occupational Safety and Health Act** (Arbeitsschutzgesetz = ArbSchG), the employer in turn must above all observe the requirements of the **Ordinance on Industrial Safety and Health** (Betriebssicherheitsverordnung, BetrSichV), because assisting robots are to be qualified as work equipment within the meaning of Section 5 BetrSichV. To eliminate or reduce sources of danger, cooperation between the employer and the manufacturer is also required for reasons of occupational safety and health (Kohte 2015, p. 1419). Another key factor is the **risk assessment** required by occupational safety and health law (Section 5 ArbSchG, Section 3 BetrSichV), whereby the Institute for Occupational Safety and Health of the German Statutory Accident Insurance has developed recommendations for the use of collaborative robots (cobots) (BG/BGIA 2011). As far as the subsequent use of assistive robots is concerned, concrete instructions for action exist in EN ISO 10218. Comparable principles apply to other assistive systems such as exoskeletons (Martini and Botta 2018, pp. 635 et seq.).

The increasing infiltration of industrial manufacturing with digital technologies also makes the issue of the associated **mental stress** increasingly important (Baikcioglu 2015; Sasse and Schönfeld 2016; especially with regard to Industry 4.0 Diebig et al. 2017). With the extension of the risk assessment to mental stress introduced in 2013 by Section 5(3) no. 6 ArbSchG, labour law provides a legal framework that addresses this problem in a comparatively general way. A legal ordinance concretising these abstract requirements ('anti-stress ordinance') has been discussed for several years but has not yet been able to gain acceptance in the political arena (Krause 2016, pp. B 70 et seq.).

3.2 Employee Data Protection in Value Networks

Industry 4.0 not only leads to an increasingly extensive processing of machine and operational data, but also builds on the processing of personal data to a considerable extent. Examples include the automated instructions just mentioned, but also assistance systems that adjust to the personal circumstances of employees, such as anthropometrically optimally shaped workstations (Hofmann 2016, p. 13). It is also conceivable to process **employee data to monitor performance and behaviour** to be able to trace errors in the production of workpieces on this basis, for example by linking machine data with shift schedules. To increase productivity and reduce malfunctions, it is also possible to record which employees carry out which maintenance tasks and malfunctions particularly quick and reliable (**smart maintenance**).

Since 2018, the legal framework for the processing of employee data has primarily been provided by the GDPR and Section 26 **BDSG**, which was enacted to fill the opening clause of Article 88(1) GDPR. These regulations are overarched at the European level by Article 7 (right to respect for private life) and Article 8 (right to protection of personal data) of the Charter of Fundamental Rights (CFR) and by Article 8 (right to respect for private life) of the European Convention on Human Rights (ECHR) and at the constitutional level by the fundamental right to informational self-determination derived from the general right of personality under Article 2(1) in conjunction with Article 1(1) Basic Law (Grundgesetz = GG) (Franzen 2019, pp. 22 et seq.). In this context, Section 26 BDSG and the relevant provisions of the GDPR are of primary importance for the practical application of the law.

Employee data protection issues are always raised when it comes to processing personal data, i.e. any form of handling information that relates to an identified or identifiable employee (Article 4 no. 1 and 2 GDPR, Section 26(6) BDSG). If only machine or operational data is processed, on the other hand, there are no data protection problems from the outset. The storage, transmission and evaluation of such data in machines or workpieces to optimise the flow of materials or to avoid production interruptions is irrelevant under data protection law (Seifert 2018, pp. 179 et seq.). The same applies if such data is transferred across company boundaries throughout the entire value chain or centrally outsourced in a cloud worldwide, as is apparently happening as part of a cooperation between Volkswagen and Amazon (Industrial Cloud). On the other hand, as soon as employee data is processed, for example by linking it to machine or operational data, employee data protection law comes into play. According to this, data processing is only lawful if certain legal requirements are met (Article 6 GDPR). The central legal basis for the area of Industry 4.0 applications thus is Section 26(1) sentence 1 BDSG, which declares the processing of personal data of employees to be lawful in principle if this is necessary for the performance of the employment relationship. As far as the concretisation of these abstract requirements is concerned, the fundamental **prohibition of total surveillance** has been recognised for a long time, which was recently confirmed by the Federal Labour Court (Bundesarbeitsgericht = BAG) for the previous version of the BDSG based on the comprehensive surveillance of employees with the help of a keylogger to control the entire usage behaviour of an employee on a computer (BAG, judgment of 27 July 2017 – 2 AZR 681/16, BAGE 159, 380, para. 33; also Kort 2018, p. 25). Therefore, employee data may not be collected, stored and analysed in a way that enables all-round monitoring. Furthermore, it must be examined in each individual case which employee data must be processed at all to achieve certain legitimate purposes. In this context, increasing efficiency is a legitimate purpose in itself that can justify processing employee data (Hofmann 2016, p. 15; Hornung and Hofmann 2018, pp. 240 et seq.). It will also be permissible to store data to determine a need for further qualification (cf. BAG, judgment of 14 November 2006 – 1 ABR 4/06, BAGE 120, 146, para. 39; Hofmann 2016, p. 17). In all of this, however, the principles of transparency and data minimisation in particular must be observed (Article 6(1)(a) and (c) GDPR). In addition, stricter requirements apply if special categories of personal employee data are processed, including biometric data and health data (cf. Article 9(1) GDPR, Section 26(3) BDSG), which may play a role in assistance systems under certain circumstances (on data protection in assistance systems, see Steidle 2005, pp. 189 et seq. for details). The **consent** of the data subject, which in principle can constitute a justification for processing personal data (Article 6(1)(a), 7 GDPR, Section 26(2) BDSG) and is also not excluded from the outset with regard to special categories of personal data (Article 9(2)(a) GDPR, Section 26(3) BDSG), is unlikely to play a significant role in the area of Industry 4.0, because it is difficult to make the optimisation of production processes that rely on the processing of employee data dependent on whether all affected employees have declared their consent in a legally valid manner and have not revoked it. Moreover, consent would only be effective if it is a voluntary decision by the employee concerned (Article 7(4) GDPR, Section 26 (2) BDSG), which again will be doubtful in many cases.

Finally, as far as the inter-company transfer of employee data is concerned, the transfer of data within groups of companies (Article 4 no. 19 GDPR) is not permitted at will, but it is privileged (cf. Recital 48 sentence 1 GDPR). However, this privilege does not extend to companies that are part of a digitally linked value network (suppliers, customers) but do not belong to the group of companies (Seifert 2018, p. 182).

The mandatory **co-determination of the works council** under Section 87(1) no. 6 BetrVG in the case of the introduction of technical equipment intended to monitor the performance or behaviour of employees constitutes a procedural protection mechanism. This provision has always been interpreted broadly by case law and comes into play even if the technical device in question is merely suitable for monitoring employees (BAG, judgment of 13 Dezember 2016 – 1 ABR 7/15, BAGE 157, 220, para. 22; details in Krause 2020). In practice, this right of co-determination in particular often proves to be a sharp sword to protect employees from too intensive monitoring of their performance or other conduct. Thus, the works agreements concluded on this basis, or the decisions of the conciliation committee often regulate in great detail which employee data accruing in connection with production may be used for which purposes or for which purposes the use of such data should not be permitted.

4 Institutionalised Influence on the Transformation Process

The above-mentioned right of co-determination in relation to technical equipment suitable for monitoring is already a central provision that provides the employee side with an institutionalised influence on certain components of the transformation process towards Industry 4.0. In addition, various other participation rights come into consideration in this context (Sect. 4.1), which also relate in particular to the area of workforce training, which is central to the success of this concept (Sect. 4.2), but which do not amount to comprehensive co-regulation of the transformation process.

4.1 Co-Design of New Work Processes by the Works Council

The works council has the strongest participation rights in the sense of equal participation in the catalogue of topics under Section 87(1) BetrVG, which, under the—not very meaningful—heading of social matters, primarily summarises matters that involve the implementation of specifications related to the production process at the level of work technology with a view to the employees. From the specific perspective of Industry 4.0, in addition to the aforementioned Section 87(1) no.

6 BetrVG, the rights of co-determination in **occupational health protection** under Section 87(1) no. 7 BetrVG and in **group work** under Section 87(1) no. 13 BetrVG are likely to be of the greatest importance. In addition, for all forms of **flexible working time models**, such as mobile work, there is the right of co-determination under Section 87(1) no. 2 BetrVG and moreover under the new Section 87(1) no. 14 BetrVG, introduced by the Works Councils Modernisation Act of 2021.

In the area of **work design**, however, the works council's participation rights are much weaker. With regard to the planning of technical equipment, work procedures, work processes and workplaces, Section 90 BetrVG only provides for the right to information and consultation, whereby the parties are at least supposed to consider the **established findings of ergonomics on the humane design of work**. There is no question that the developments planned with the digital networking of the entire value chain through Industry 4.0 will also lead to considerable changes with regard to the areas mentioned in Section 90 BetrVG. In this respect, the Works Councils Modernisation Act explicitly expanded the catalogue in Section 90 BetrVG to include the use of **artificial intelligence**. To be able to better assess the processes, some of which are difficult to understand, and to consult properly with the employer, the works council may have a **right to consult an expert** under Section 80 (3) BetrVG, which will be facilitated in future by the new Section 80(3) sentences 2 and 3 BetrVG, also introduced by the Works Councils Modernisation Act. Section 91 BetrVG provides a more extensive right of participation. Under this section, the works council may, under certain conditions, demand remedial measures and, if necessary, enforce them through the conciliation board. However, the hurdles for this right of co-determination are so high that the provision has hardly gained any practical significance so far. For example, it is necessary that the changes to workplaces, work processes or the working environment obviously contradict established scientific knowledge and that the employees are particularly burdened by them. In this way, therefore, only those Industry 4.0 solutions can be influenced that manifestly fail to achieve the self-imposed goal of this concept of also contributing to improved working conditions (cf. Forschungsunion/acatech 2013, pp. 5, 27 et seq., 56 et seq.). Furthermore, the works council can take the further development in the direction of Industry 4.0 as an opportunity to address future personnel planning, although it is only entitled to information and consultation rights under Section 92 BetrVG. In individual cases, the intended increase in efficiency can also trigger **employment risks** and thus activate the participation right under Section 92a BetrVG, which is also quite weak. In addition, in companies with regularly more than one hundred permanently employed workers, the existing **economic committee** must be involved at an early stage in the planning for a conversion to Industry 4.0, whereby in this respect Section 106(3) no. 5 (fabrication and working methods, in particular the introduction of new working methods) as well as no. 9 BetrVG (change of company organisation) are relevant as subjects of mandatory information and consultation. In individual cases, the prerequisites of a **change in operations** under Section 111 sentence 3 no. 4 (fundamental change in business organisation) or no. 5 BetrVG (introduction of fundamentally new working methods and production processes) may also be fulfilled, so that the works council must be comprehensively informed about the plans and, under Section 112 BetrVG, a reconciliation of interests and, under certain circumstances, a **'qualification social plan'** may result (on this, Göpfert and Wenzler 2020; Günther 2018, pp. 81 et seq.; Röder and Gebert 2017). In addition, the parties can conclude company agreements on all company digitalisation processes under Section 77 BetrVG, but in practice this encounters numerous obstacles, such as the speed and complexity of the changes (Matuschek and Kleemann 2018).

A completely different question is whether Industry 4.0 will undermine the legal preconditions for co-determination in companies, in that the company as the basis of the works council institution will lose its contours. However, the discussion to date does not provide sufficient evidence that the creation of continuous data flows within the entire value chain, despite the envisaged decentralisation of self-controlling **cyber-physical systems** by means of edge computing (Forschungsunion/acatech 2013, p. 96; Hirsch-Kreinsen 2014, p. 421; Pfeiffer and Huchler 2018, p. 170; Spath 2013, p. 22), the responsibilities for matters relevant to participation on the employer side, which are ultimately decisive for the structure of works constitution law, are also eroding and the traditional concept of establishment is therefore outdated (see Benecke 2018a, paras. 8 et seq.; Franzen 2016; Günther and Böglmüller 2015, p. 1027; Krause 2016, pp. B 89 et seq.). Depending on the case, the legal concept of joint operation of several companies, which has been known for a long time and is anchored in Section 1(2) BetrVG, can also help, or the possibility of transferring tasks to working groups under Section 28a BetrVG, which has been provided for by law since 2001, can be mobilised. In addition, under Section 3 BetrVG, there is the possibility of a largely customised regulation of the works constitution by collective agreement or works agreement (on this, Benecke 2018a, paras. 25 et seq.; Franzen 2016; Krause 2016, pp. B 92 et seq.).

4.2 Continuing Professional Qualification as an Ongoing Process Close to the Workplace

One of the key success factors for the implementation of Industry 4.0 is undoubtedly the issue of **qualification** and **further training** of employees to keep pace with technical developments and the resulting constantly changing work requirements (Forschungsunion/acatech 2013, pp. 6 et seq., 59, 61; Hammermann and Klös 2016, pp. 90 et seq.; Hirsch-Kreinsen 2014, p. 424; Kärcher 2015, pp. 52 et seq.; with regard to artificial intelligence, see Enquete Commission 2020, p. 321). Even if numerous measures of continuing vocational qualification and training already take place at the company level (Heidemann 2015), it is necessary to ask what support labour law provides in this regard, not least against the empirically confirmed background that learning close to the workplace is more effective and efficient than other forms of continuing professional qualification and training (Stettes 2016, p. 69).

However, at the level of individual employment law, there is essentially a lack of action. In addition to the apparently not very effective educational leave laws of the Länder, Section 81 (4) BetrVG provides for a right of the employee to a discussion with the employer if it is clear that the employee's activities will change due to a redesign of the work environment and that his or her professional knowledge and skills will no longer be sufficient to fulfil his or her tasks. In contrast, the provision does not provide employees with a right to further training (Günther 2018, p. 65; Kleinebrinck 2018, p. 256).

At the level of collective labour law (see Kleinebrinck 2018 and Krause 2021 for more details), the works council can, under Section 96(1) sentence 2 BetrVG, first demand that the employer determine the **vocational training** needs of the company. Furthermore, under Section 96(1) sentences 2 and 3 BetrVG, the works council has a general right of consultation and proposal in matters of vocational training. In this respect, the Works Councils Modernisation Act introduced the new Section 96 (1a) BetrVG which provides that the works parties should also agree on vocational training measures during consultations. In case of disagreement, both sides are to be given the right to appeal to the conciliation board for mediation, although the board can only work towards an agreement between the parties without being able to make a binding decision itself. In addition, the law contains a graduated system of works council involvement. First, the general right to consult is condensed into a duty to consult independent of the works council's request, insofar as, inter alia, the introduction of in-company vocational training measures and participation in external vocational training measures are concerned (Section 97(1) BetrVG). If the employer plans or implements general company measures which have the consequence that the activities of the employees concerned change and their vocational knowledge and skills are no longer sufficient to fulfil their tasks, the works council also has a genuine right of co-determination, including a right of initiative, on the introduction of company vocational training measures (Section 97(2) BetrVG). However, this provision, which aims to adjust qualifications, has rarely been the subject of court decisions, so that numerous details have not yet been conclusively clarified (see Fitting 2020, Section 97 BetrVG paras. 20 et seq.; Franzen 2001). Nevertheless, the provision can provide works councils that are regularly interested in this issue (cf. Bosch et al. 2017, p. 18) with a means of enforcing in-company vocational training measures, especially during the transition to Industry 4.0, to ensure that employees are able to cope with future work requirements and, in particular, do not have to fear for their jobs because their qualifications are no longer adequate (Göpfert and Seier 2019). Finally, as far as the implementation of in-company vocational training measures is concerned, the works council also has a comprehensive right of co-determination (Section 98 BetrVG), which relates, for example, to the timing and the selection of the participants in the training measure, while the employer can decide on the amount of funds used without co-determination (BAG, judgment of 24 August 2004 – 1 ABR 28/03, BAGE 111, 350, 358). In addition, the issue of qualification is increasingly becoming the focus of the parties to collective agreements, with the metal and electrical industry in particular setting new standards at the sectoral level with its 2015 **collective agreement on qualification**. Although the focus of all these efforts is on securing employment, the function of vocational qualification as a means of securing personal autonomy should not be completely ignored.

5 Conclusion and Perspectives

The introduction of Industry 4.0 may be a new industrial revolution, yet it does not trigger a revolution in labour law. Nevertheless, the question arises for some sub-areas of labour law as to whether and to what extent they still do justice to the fundamental goal of labour law regulations of creating adequate institutional protection for dependent employees against the risks of the work process and the forces of the labour market. Admittedly, it is too early at present to give a conclusive answer to these questions. Nor should the adaptability of labour law and especially of the parties to collective agreements at the sectoral and company level be underestimated. Works constitution law particularly provides a legal framework that enables at least energetic works councils to shape company labour relations together with the employer in such a way that employee interests are not neglected with regard to the day-to-day demands of work, while it is primarily the task of the sectoral social partners to split the digitisation dividend appropriately between the employer side and the employee side. However, if it turns out that the existing labour law framework no longer takes sufficient account of the economic and social requirements of digitalised value chains, a further development of the legal regulations should also be discussed in the sense of a 'protective purpose 4.0' (Kurt 2018, p. 142).

References

Abel RB (2018) Automatisierte Entscheidungen im Einzelfall gem. Art. 22 DS-GVO. ZD 8:304–307

Arnold C, Günther J (eds) (2018) Arbeitsrecht 4.0. München

Arnold C, Winzer T (2018) Flexibilisierung im individuellen Arbeitsrecht. In: Arnold C, Günther J (eds) Arbeitsrecht 4.0, München, Kapitel 3

Baikcioglu J (2015) Psychische Erkrankungen am Arbeitsplatz. NZA 32:1424–1433

Baker McKenzie (ed) (2017) Arbeitswelt 4.0. Berlin

Benecke M (2018a) Kollektives Arbeitsrecht 4.0. In: Arnold C, Günther J (eds) Arbeitsrecht 4.0, München, Kapitel 7

Benecke M (ed) (2018b) Unternehmen 4.0. Arbeitsrechtlicher Strukturwandel durch Digitalisierung. Baden-Baden

BITKOM (ed) (2017) Künstliche Intelligenz verstehen als Automation des Entscheidens. Berlin

BITKOM/VDMA/ZVEI (eds) (2015) Umsetzungsstrategie Industrie 4.0. Ergebnisbericht der Plattformindustrie 4.0. Berlin, Frankfurt am Main

Bosch G, Bromberg T, Haipeter T, Schmidt J (2017) Industrie und Arbeit 4.0. IAQ-Report. Duisburg, Essen

Bosch G, Schmitz J, Haipeter T, Spallek A-C (2020) Gestaltung von Industrie 4.0 durch gewerkschaftliche Betriebspolitik. Arbeit 29:3–23
Boston Consulting Group (2019) Advanced robotics in the factory of the future. Boston
Botthoff A, Hartmann EA (eds) (2015) Zukunft der Arbeit in Industrie 4.0. Berlin
Brandt A, Polom L, Danneberg M (2016) Gute digitale Arbeit. Auswirkungen der Digitalisierung im Dienstleistungsbereich. Bonn
Buchner B (2020) Artikel 22. In: Kühling J, Buchner B (eds) Datenschutz-Grundverordnung, 3rd edn. München
Bücker A (2016) Arbeitsrecht in der vernetzten Arbeitswelt. Industrielle Beziehungen 23:187–225
Bundesministerium für Arbeit und Soziales (BMAS) (2015) Grünbuch Arbeiten 4.0. Berlin
Bundesministerium für Arbeit und Soziales (BMAS) (2016) Weißbuch Arbeiten 4.0. Berlin
Bundesvereinigung der Deutschen Arbeitgeberverbände (BDA) (2015) Arbeitswelt 4.0 – Chancen nutzen, Herausforderungen meistern. Positionen der BDA zum Grünbuch "Arbeiten 4.0" des Bundesministeriums für Arbeit und Soziales. Berlin
Bundesvereinigung der Deutschen Arbeitgeberverbände (BDA) (2018) Germany Reloaded – Wie Wirtschaft und Beschäftigte von der Digitalisierung profitieren können. Berlin
Däubler W (2016a) Digitalisierung und Arbeitswelt. SR Sonderausgabe 6:1–44
Däubler W (2016b) Herausforderungen für das Arbeitsrecht – Deregulierung, Globalisierung, Digitalisierung. AuR 64:325–334
Däubler W (2020) Digitalisierung und Arbeitsrecht, 7th edn. Frankfurt am Main
Diebig M, Müller A, Angerer P (2017) Psychische Belastungen in der Industrie 4.0. ASU 52:832–839
Dzida B (2016) Wearables am Arbeitsplatz. Wie tragbare Computersysteme Einzug in die Betriebe halten. ArbRB 16:146–149
Enquete-Kommission "Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale", Bericht (2020), BT-Drucksache 19/23700
Ernst C (2017) Algorithmische Entscheidungsfindung und personenbezogene Daten. JZ 72:1026–1036
Eurofound (2018) Automation, digitisation and platforms: implications for work and employment. Luxembourg
Evers M, Oberbeck H (2017) Industrie 4.0 aus sozialwissenschaftlicher Perspektive: die Bedeutung betrieblicher Akteure und Interessenlagen. DRV 72:61–70
Fitting K (2020) Betriebsverfassungsgesetz, 30th edn. München
Forschungsunion/acatech (ed) (2013) Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Abschlussbericht des Arbeitskreises Industrie 4.0. Frankfurt am Main
Franzen M (2001) Das Mitbestimmungsrecht des Betriebsrats bei der Einführung von Maßnahmen der betrieblichen Berufsbildung nach § 97 II BetrVG. NZA 18:865–871
Franzen M (2016) Folgen von Industrie 4.0 für die Betriebsverfassung – Betriebsbegriff und Vereinbarungen nach § 3 BetrVG. In: Giesen R, Junker A, Rieble V (eds) Industrie 4.0 als Herausforderung des Arbeitsrechts. München, pp 107–126
Franzen M (2019) Persönlichkeitsrecht und Datenschutz im Arbeitsverhältnis. ZFA 50:18–39
Funken C, Schulz-Schaeffer I (Hrsg) (2008) Digitalisierung der Arbeitswelt: Zur Neuordnung formaler und informeller Prozesse in Unternehmen. Wiesbaden
Giesen R (2018) Robotik und Arbeitsrecht. RW 9:431–452
Giesen R, Kersten J (2017) Arbeiten 4.0. Arbeitsbeziehungen und Arbeitsrecht in der digitalen Welt. München
Giesen R, Junker A, Rieble V (eds) (2016) Industrie 4.0 als Herausforderung des Arbeitsrechts. München
Göpfert B, Seier J (2019) Die "Transformation-Einigungsstelle": Inhalt und Grenzen eines "Qualifizierungs-Sozialplans". NZA 36:588–594
Göpfert B, Wenzler F (2020) Erzwingbarer "Qualifizierungs"-Sozialplan? NZA 37:15–20
Groß N, Gressel J (2016) Entpersonalisierte Arbeitsverhältnisse als rechtliche Herausforderung – Wenn Roboter zu Kollegen und Vorgesetzten werden. NZA 33:990–996

Günther J (2018) Digitalisierung der Arbeitswelt – Veränderungen in der Belegschaft. In: Benecke M (ed) (2018b) Unternehmen 4.0. Arbeitsrechtlicher Strukturwandel durch Digitalisierung. Baden-Baden, pp 63–90

Günther J, Böglmüller M (2015) Arbeitsrecht 4.0 – Arbeitsrechtliche Herausforderungen der vierten industriellen Revolution. NZA 32:1025–1031

Günther J, Böglmüller M (2017) Künstliche Intelligenz und Roboter in der Arbeitswelt. BB 72:53–58

Haipeter T, Korflür I, Schilling G (2018) Neue Koordinaten für eine proaktive Betriebspolitik. WSI Mitteilungen 71:219–226

Hammermann A, Klös H-P (2016) Digitalisierung und "Industrie 4.0": Auswirkungen auf das Personal- und Qualifizierungsmanagement. In: Giesen R, Junker A, Rieble V (eds) Industrie 4.0 als Herausforderung des Arbeitsrechts. München, pp 77–101

Hanau H (2016) Schöne digitale Arbeitswelt? NJW 69:2613–2617

Haußmann K (2019) Arbeit 4.0 – Update Arbeitsrecht?! RdA 72:131–134

Heidemann W (2015) Trendbericht: Betriebliche Weiterbildung, 2nd edn. Düsseldorf

Henseler-Unger I (2017) Auf dem Weg zur Industrie 4.0. In: Sassenberg T, Faber F (eds) Rechtshandbuch Industrie 4.0 und Internet of Things. München, Teil 1

Hirsch-Kreinsen H (2014) Wandel von Produktionsarbeit – "Industrie 4.0". WSI Mitteilungen 67: 421–429

Hirsch-Kreinsen H, Ittermann P, Niehaus J (eds) (2018) Digitalisierung industrieller Arbeit. Die Vision Industrie 4.0 und ihre sozialen Herausforderungen, 2nd edn. Baden-Baden

Hoeren T, Niehoff M (2018) KI und Datenschutz – Begründungserfordernisse automatisierter Entscheidungen. RW 9:47–66

Hofmann K (2016) Smart Factory – Arbeitnehmerdatenschutz in der Industrie 4.0. ZD 6:12–17

Höpfner C, Daum JA (2021) Der "Robo-Boss" – Künstliche Intelligenz im Arbeitsverhältnis. ZFA 52:467–501

Hornung G, Hofmann K (2018) Datenschutz als Herausforderung der Arbeit in der Industrie 4.0. In: Hirsch-Kreinsen H, Ittermann P, Niehaus J (eds) Digitalisierung industrieller Arbeit. Die Vision Industrie 4.0 und ihre sozialen Herausforderungen, 2nd edn. Baden-Baden, pp 233–255

Huchler N (2017) Grenzen der Digitalisierung von Arbeit – Die Nicht-Digitalisierbarkeit und Notwendigkeit impliziten Erfahrungswissens und informellen Handelns. ZArbWiss 71:215–223

Institut für Arbeitsmarkt- und Berufsforschung (IAB) (2015) Industrie 4.0 und die Folgen für Arbeitsmarkt und Wirtschaft. Nürnberg

Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung (2011) BG/BGIA-Empfehlungen für die Gefährdungsbeurteilung nach Maschinenrichtlinie. Gestaltung von Arbeitsplätzen mit kollaborierenden Robotern. Sankt Augustin

International Bar Association (IBA) Global Employment Institute (ed) (2017) Artificial intelligence and robotics and their impact on the workplace. London

Kagermann H, Lukas W-D, Wahlster W (2011) Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur vierten industriellen Revolution, vdi-Nachrichten (13):2

Kärcher B (2015) Alternative Wege in die Industrie 4.0 – Möglichkeiten und Grenzen. In: Botthoff A, Hartmann EA (eds) Zukunft der Arbeit in Industrie 4.0. Berlin, pp 47–58

Kirn S, Müller-Hengstenberg CD (2014) Intelligente (Software-)Agenten: Von der Automatisierung zur Autonomie? Verselbständigung technischer Systeme. MMR 17:225–232

Klebe T (2019) Künstliche Intelligenz – eine Herausforderung für die Mitbestimmung. SR 9:128–137

Kleinebrinck W (2018) Initiativrechte des Betriebsrats auf digitale Weiterbildung der Arbeitnehmer. DB 71:254–259

Koalitionsvertrag zwischen SPD, BÜNDNIS 90/DIE GRÜNEN und FDP (2021). Berlin

Kohte W (2015) Arbeitsschutz in der digitalen Arbeitswelt. NZA 32:1417–1424

Kort M (2018) Neuer Beschäftigtendatenschutz und Industrie 4.0. RdA 71:24–33

Kramer S (ed) (2019) IT-Arbeitsrecht, 2nd edn. München

Krause R (2016) Digitalisierung der Arbeitswelt – Herausforderungen und Regelungsbedarf, Gutachten B zum 71. Deutschen Juristentag. München
Krause R (2017) Digitale Transformation der Arbeitswelt und arbeitsrechtliche Regulierung. JbArbR 54:23–50
Krause R (2020) Sozialverträgliche Arbeitnehmerüberwachung – Technikbasierte Beschäftigtenkontrolle als Gegenstand betrieblicher Mitbestimmung im digitalen Zeitalter. In: Gräfl E, Lunk S, Oetker H, Trebinger Y (eds) 100 Jahre Betriebsverfassungsrecht. München, pp 353–368
Krause R (2021) Der betriebsverfassungsrechtiche Rahmen für die berufliche Weiterbildung als "Infrastrukturrecht". In: Klapp M, Linck R, Preis U, Reinhard B, Wolf R (eds) Die Sicherung der kollektiven Ordnung. Festschrift für Ingrid Schmidt. München, pp 231–247
Kuhlmann M, Splett B, Wiegrefe S (2018) Montagearbeit 4.0. Eine Fallstudie zu Arbeitswirkungen und Gestaltungsperspektiven digitaler Werkerführung. WSI Mitteilungen 71:182–188
Kurt S (2018) Der Schutzzweck des Arbeitsrechts in der digitalen und globalen Welt – eine rechtethische Annäherung. In: Benecke M (ed) Unternehmen 4.0. Arbeitsrechtlicher Strukturwandel durch Digitalisierung. Baden-Baden, pp 115–147
Lembke M (2020) § 106 GewO. In: Henssler M, Willemsen HJ, Kalb H-J (Hrsg) Arbeitsrecht Kommentar, 9th edn. Köln
Martini M (2017) Algorithmen als Herausforderung für die Rechtsordnung. JZ 72:1017–1025
Martini M, Botta J (2018) Iron Man am Arbeitsplatz? – Exoskelette zwischen Effizienzstreben, Daten- und Gesundheitsschutz. NZA 35:625–637
Matuschek I, Kleemann F (2018) "Was man nicht kennt, kann man nicht regeln". Betriebsvereinbarungen als Instrument der arbeitspolitischen Regulierung von Industrie 4.0 und Digitalisierung. WSI Mitteilungen 71:227–234
May E (2014) Robotik und Arbeitsschutzrecht. In: Hilgendorf E (ed) Robotik im Kontext von Recht und Moral. Baden-Baden, pp 99–118
Möhring-Hesse M (2007) Wie Rechtsbrechern zu ihrem Recht verhelfen? Arbeitsrecht unter den Bedingungen subjektivierter Arbeit. KJ 40:347–357
Neighbour K (2017) Arbeitsrecht – Realität und Herausforderungen. In: Sassenberg T, Faber F (eds) Rechtshandbuch Industrie 4.0 und Internet of Things. München, Teil 2 G
Niederée C, Nejdl W (2020) § 2 Technische Grundlagen der KI. In: Ebers M, Heinze C, Krügel T, Steinrötter B (eds) Künstliche Intelligenz und Robotik, München
Oetker H (2016) Digitalisierung der Arbeitswelt – Herausforderungen und Regelungsbedarf. JZ 71: 817–824
Paul H (2016) Industrie 4.0 – Annäherung an ein Konzept, IAT Forschung Aktuell. Gelsenkirchen
Pfeiffer S, Huchler N (2018) Industrie 4.0 konkret – vom Leitbild zur Praxis? WSI Mitteilungen 71: 167–173
Pfrogner J (2018) Flexible Arbeitszeiten in der Industrie 4.0. BB 73:500–505
Reichwald J, Pfisterer D (2016) Autonomie und Intelligenz im Internet der Dinge. Möglichkeiten und Grenzen autonomer Handlungen. CR 32:208–212
Röder G, Gebert C (2017) Technologischer Wandel und Betriebsänderung – Bringen Industrie 4.0 und E-Mobilität den "Qualifizierungssozialplan"? NZA 34:1289–1296
Röder G, Günther J, Gerigk M (2021) Zukunftsvereinbarungen – ein neues Spielfeld für Gewerkschaften und Betriebräte? DB 74:1741–1744
Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2017) Für eine zukunftsorientierte Wirtschaftspolitik. Jahresgutachten 2017/2018. Wiesbaden
Sasse S, Schönfeld J (2016) Rechtliche Aspekte psychischer Belastungen im Arbeitsverhältnis. RdA 69:346–357
Schiefer B, Worzalla M (2019) Moderne Arbeitswelt (Teil 1): Spannungsfeld zwischen technischer Entwicklung und Individualarbeitsrecht. DB 72:1904–1908
Schimmelpfennig H-C (2016) Diskussionsbeitrag. In: Giesen R, Junker A, Rieble V (eds) Industrie 4.0 als Herausforderung des Arbeitsrechts. München, pp 70–71
Schirmer J-E (2016) Rechtsfähige Roboter? JZ 71:660–666

Schwarze R (2020) § 8 Arbeitsrechtliche Probleme von KI und Robotik. In: Ebers M, Heinze C, Krügel T, Steinrötter B (Hrsg.) Künstliche Intelligenz und Robotik, München

Seidel U (2017) Industry 4.0 and law – experiences from AUTONOMICS. In: Hilgendorf E, Seidel U (eds) Robotics, autonomics, and the law. Baden-Baden, pp 11–26

Seifert A (2018) Arbeit in der digitalen Fabrik am Beispiel von Wertschöpfungsnetzwerken. In: Hornung G (ed) Rechtsfragen der Industrie 4.0. Baden-Baden, pp 175–197

Simon C (2016) Digitale Transformation von Unternehmen durch Industrie 4.0. In: Giesen R, Junker A, Rieble V (eds) Industrie 4.0 als Herausforderung des Arbeitsrechts. München, pp 15–24

Simon C (2018) Digitalisierung und Industrie 4.0. In: Arnold C, Günther J (eds) Arbeitsrecht 4.0. München, Kapitel 1

Spath D (ed) (2013) Produktionsarbeit der Zukunft – Industrie 4.0, Stuttgart

Steidle R (2005) Multimedia-Assistenten im Betrieb. Datenschutzrechtliche Anforderungen, rechtliche Regelungs- und technische Gestaltungsvorschläge für mobile Agentensysteme. Wiesbaden

Stettes O (2016) Arbeitswelt der Zukunft. Wie die Digitalisierung den Arbeitsmarkt verändert, IW-Analysen Nr. 108. Köln

Teubner G (2018) Digitale Rechtssubjekte? Zum privatrechtlichen Status autonomer Softwareagenten. AcP 218:155–205

Thüsing G (2016) Digitalisierung der Arbeitswelt – Impulse zur rechtlichen Bewältigung der Herausforderung gewandelter Arbeitsformen. SR 6:87–108

Uffmann K (2016a) Digitalisierung der Arbeitswelt. NZA 33:977–984

Uffmann K (2016b) Möglichkeiten und Grenzen erfolgsabhängiger Vergütung. In: Giesen R, Junker A, Rieble V (eds) Industrie 4.0 als Herausforderung des Arbeitsrechts. München, pp 31–53

Vereinigung der Bayerischen Wirtschaft (vbw) (2016) Arbeitszeit – Handlungsfähigkeit der Unternehmen stärken. München

Vogel-Heuser B, Bauernhansl T, ten Hompel M (eds) (2017) Handbuch Industrie 4.0, Band 4, Allgemeine Grundlagen, 2nd edn. Berlin

Wedde P (2020) Automatisierung im Personalmanagement – arbeitsrechtliche Aspekte und Beschäftigtendatenschutz. Berlin

Wildhaber I (2016) Die Roboter kommen – Konsequenzen für Arbeit und Arbeitsrecht. ZSR 135: 315–351

Wischmann S, Hartmann EA (eds) (2018) Zukunft der Arbeit – Eine praxisnahe Betrachtung. Berlin

World Economic Forum (2018) The future of jobs report. Geneva

The Gender Blind Spot: Reflections on the German Discourse on Gender, Work, and Digitalization—Work 4.0

Carmen Leicht-Scholten

Contents

C. Leicht-Scholten (✉)
Chair of Gender and Diversity in Engineering, RWTH Aachen University, Aachen, Germany
e-mail: carmen.leicht@gdi.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_54

1 Introduction[1]

Particularly against the backdrop of the COVID-19 pandemic that began in 2020, technological change with the suffix "4.0" is attracting even more attention from business, politics, the media, and the public. As the numerous contributions to this book show, the effects and prognoses are being discussed in a wide variety of contexts, from production, transport, law, and medicine, to education, and are generating not only headlines but also heated debates about the opportunities and risks. The COVID-19 pandemic has fostered these discussions, as social distancing measures in the interest of infection suppression have led to an increasing use of and need for digital infrastructure. This is particularly relevant in the professional sphere, which is at the center of many of these discourses. Particularly the rapidly increased adoption of working from home as common practice leads to a higher level of digitalization for individuals. And while some discourses on the next generation of Work 4.0 have up until this point been dominated by the issue of the elimination of jobs through automation processes, others have emphasized a growing need for new knowledge workers. There is agreement and excitement that increasing digitalization will be associated not only with economic but also with major social transformation processes. Yet, technological innovations have had a decisive impact on social development since the earliest beginnings of industrialization. The innovations of Industry 2.0 and 3.0 also brought massive changes to our lives and work. If we follow the picture painted in the media, the new technologies will once again completely reorganize our lives. What insights can be drawn from empirical hindsight? And what does this mean for existing gender relations?

Following these questions, *Sect. 2* begins by approaching the relevance of a gendered view of work in general. It provides an overview of key differences in the occupational integration of men and women, as well as the main issues and areas of action associated with such differences. This provides a sound starting point for the subsequent consideration in the context of Work 4.0.

Using two much-discussed studies as examples, *Sect. 2.1* examines the extent to which the different positions of and consequences for men and women are considered in the context of the transformation of work through digitalization and automation and reflects on the results. Following this examination, *Sect. 2.2* expands the focus, providing an overview of where in Work 4.0 aspects and dimensions relevant to gender relations are evident, and an attempt is made to classify them. Looking at the actual situation under COVID-19 enables a litmus test and demonstrates the urgency of considering gendered perspectives.

---

[1] The article is an actualized version of Leicht-Scholten and Bouffier (2019): Mind the Gap Gender trifft Industrie 4.0 taking into account actual discussions. I would like to thank Marie Decker who contributed to the actualization of this article. She is a PhD candidate at the GDE with research interests in the fair application of data science and machine learning methods, particularly concerning issues of inequality reduction.

In *Sect. 3*, a comparative look is taken at the discourses on gender equity that were conducted in the context of previous transformation processes in the course of Industry 2.0 and 3.0. Have these changed or are we continuing discourses that have already begun? Can digitalization create conditions under which structures are broken up and a renegotiation of the gender order becomes possible? At the center of this thought process is the question of whether the potential for greater gender equity resides in digitalization, and if so, how. Examples are used to show how addressing gender issues broadens the view of central issues of social justice and reveals their close intersectional entanglement with further categories of social inequality.

In the final conclusions in *Sect. 4*, an attempt is made to provide an answer to the initial question posed about the potential contribution of digitalization to greater gender justice. It is argued that an active discussion and inclusion of the issue of gender equity can serve as a litmus test for a broader discourse on a Society 4.0. A society that addresses a multitude of dimensions of social justice and uses the initiated transformation processes of Industry 4.0 for a comprehensive societal change, creating new realms of opportunity for a good life for all through a changed humane concept of work, instead of remaining in a Society 3.0 with a social structure that is increasingly drifting apart.

2 Gender Aspects in Work 4.0: An Overview

When we speak of "**Industry 4.0,**" we usually mean more than the digital transformation of the *industrial* sector of our economy. Rather, the term with the numerical abbreviation is often used synonymously with other descriptions for the fourth form of industrialization, "Digitalization," in which the transformation process of our entire economy as well as our society as a whole is undergoing. To speak only of Industry 4.0 therefore seems an inadmissible oversimplification. For this reason, the term "**Digitalization**" is used to refer to the entire development within the fourth industrial revolution. This is because it affects nearly all areas of life and branches of industry. It is changing production methods, business models, value chains, work cultures, everyday practices, ways of thinking and thus, of course, also gender issues!

The starting point of our discussion about the digitalized "Society 4.0" is the fact that the character of our work and working society are undergoing fundamental changes. The concept of work, in the sense of **gainful employment**, forms the center of social institutions and structures, and thus shapes all areas of our lives. Our social network as well as the structure and logic behind the **course of our lives** are formed around the concept of gainful employment (Berger and Konietzka 2001; Mayer 1998). In our social structure, gainful employment is the norm for securing individual livelihood, is subjectively meaningful, and a fundamental part of social integration and social recognition.

In contrast to **paid work** or gainful employment (**Market Work**), **unpaid work** is crucial for social cohesion. This includes unpaid overtime, but above all civic engagement, political and social activities, and **reproductive work**. The latter refers to work that is essential for the preservation and continuation of our society, i.e., all activities that are performed around the family, such as raising children, caring for the needy (**Care Work**) or simply the daily recurring unpaid housework. The dual dichotomy of paid versus unpaid work is closely interwoven with the prevailing **gender stereotypes** in Germany, which are still strongly influenced by the industrial culture of the past 150 years. The resulting **gender relations** are therefore characterized by the closely intertwined dualisms of work and life, public and private, and production and reproduction (Hausen 1976).

Against the societal backdrop of educational expansion, an increasing **employment of women** (especially mothers) can be observed. In 2017, 78.9% of men and 71.5% of women were employed (Bundesagentur für Arbeit 2018, p. 7f.). With this percentage, Germany was not only above the European average of 62.4%, but also shared second place with Denmark in the European comparison of the employment rate of women aged 15 to under 65, surpassed only by Sweden (75.4%) (ibid., p. 8).[2]

On the other hand, it must be noted that around half of these women work **part-time** and temporarily interrupt their careers after the birth of a child (ibid., p. 9). While only 11% of men were employed part-time in 2016, 47.7% of women were (ibid.). These figures have increased for both men and women compared to 1991. It is furthermore noteworthy that the difference between the two sexes has increased significantly. While one-third of employed women worked part-time in 1991, nearly one in two women has done so consistently since 2006. Among men, however, the number of part-time workers rose almost continuously from 2 to 11% (WSI GenderDatenPortal 2018a). The reason for the differences between the sexes is, on the one hand, the strong increase in the participation of women in the labor force overall, which has made an employment-centered lifestyle a cross-gender norm and thus hardly suitable as a criterion for gender differentiation anymore. On the other hand, the large percentage of female part-time workers demonstrates that significant disparities still exist in the distribution of caregiving tasks between the genders (WSI GenderDatenPortal 2018b), which delineates the **compatibility** of gainful employment and family labor as a "women's issue."

Men's and women's employment also differs in the fields of work in which they are employed. A large proportion of women work in the social and health sectors, in education and administration, and in the service sector. Globally, 70% of health workers are women and they make up an even higher percentage of care-related occupations such as nursing, midwifery, and community health work (Azcona et al. 2020a, p. 5).

The COVID-19 crisis not only demonstrates the fragility of the care system but also reveals that women are the ones to suffer in multiple ways during the pandemic.

[2] For men ages 15 to 65, the employment rate in 2017 was 78.9%, also well above the EU average of 72.9% (Bundesagentur für Arbeit 2018, p. 7).

Their high exposure in care-related occupations increases the risk of infection significantly, especially in the personal care sector where women from marginalized groups are highly represented. In Germany, Italy, Spain and the United States, COVID-19 cases among female health workers have been two to three times higher than among male health workers (ibid., p. 4). In a study with 18 countries, it could be shown that COVID-19 is intensifying women's workloads at home and an Ipsos poll found out that "nearly half of all women with children at home say they spend more than 5 hours each day in childcare." (See: Azcona et al. 2020b)

On the other hand, male-dominated areas are still above all the construction industry, the mining, energy, water and waste disposal sector, the manufacturing industry, the transport and warehousing sector, as well as agriculture and forestry, and the proportion of men in the information and communication sector is also over 60% (see Bundesagentur für Arbeit 2018, p. 10).

Engineering studies, like other STEM subjects,[3] represent an additional field that is significantly dominated by men and in which women represent a minority. Germany, among others, has therefore the "National Pact for Women in STEM Professions,"[4] as well as numerous higher education policies and activities to increase the proportion of women in STEM fields. As Fig. 1 illustrates, the proportion of women is nevertheless increasing only extremely moderately. A review shows that most of the measures implemented in the form of accompanying measures start at the individual level, namely with female pupils or students, instead of adapting organizational and institutional framework conditions to implement **gender equity** (GWK 2011).

This is despite the massive increase in high school graduates due to the **expansion of education** since the end of the 1950s. While in 1960 only 6% of a graduating class took up university studies, this number increased to 23 in 2010 (Geissler 2014, p. 55) and rose to 50% by 2015 (Schneider et al. 2017, p. 42). The Vocational Education and Training Report 2017 continues to show that, starting in 2014, slightly more women than men are now taking up studies in Germany. This also applies to the years 2015 and 2016 (BMBF 2017, pp. 36 and 45).

Despite their high professional qualifications, women still experience wage inequality (**Gender Pay Gap**[5]) and are underrepresented in **leadership positions** in business, politics, and society, which puts Germany at the bottom of the list in this respect in a Europe-wide comparison (Destatis 2018; World Economic Forum 2016; Wrohlich 2017).

In view of the above-described societal changes in the context of work, as well as a changing gender system dynamic and an increasing diversity of gender roles,

---

[3] STEM is referring to Science, Technology, Engineering, and Mathematics. The German abbreviation is MINT.

[4] See the related website under: https://www.komm-mach-mint.de/.

[5] The unadjusted gender pay gap for Germany was around 21% in 2016. In a European comparison, Germany thus has the third largest difference. The difference between the gross hourly earnings of women and men was smallest in Romania and Italy, at 5% each (Destatis 2018, p. 44).

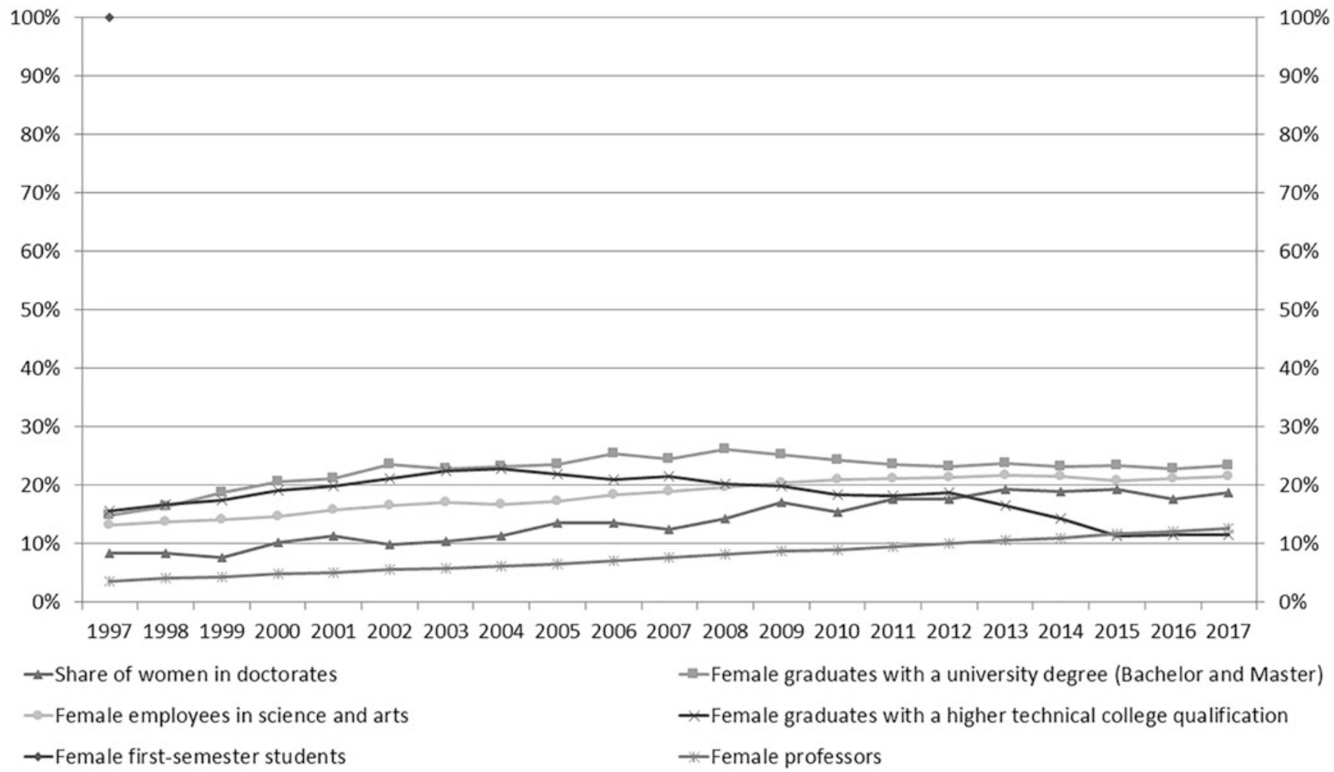


Fig. 1 Percentage of women in higher education in engineering, 1997 to 2017 (data basis: Destatis 2019, own calculations)

divisions of labor, and family and lifestyle models, the question arises as to the significance of digitalization in this transformation process.

If we begin by examining the current discussion about the impact of digitalization on the labor market, two important points are readily recognizable, irrespective of the flood of publications and discussions: Firstly, up until the widespread outbreak of COVID-19, there were no reliable prognoses that could provide information about the upcoming developments on the labor market. They were not available because, apart from indicators and circumstantial evidence, there is no solid baseline against which reliable statements can be made about impending changes. And no corresponding gender-sensitive models have been used either. There are still too many variables that cannot be sufficiently narrowed down, making a wide variety of scenarios possible. On the other hand, however, it is undisputed that increasing digitalization will *absolutely* massively and lastingly change our lives, whatever the scenario.

The debate regarding different scenarios and their significance for gender equality will be, with the aid of two studies on the probability of job automation, critically discussed in the following section.

2.1 Blind Spot Gender: Exemplary Consideration of Two Studies on the Probability of Job Automation

The publication of the American study by Frey and Osborne (2013) was cause for sensation, due to its conclusion that 47% of all employees in the United States would be very likely to lose their jobs due to **automation** and be replaced by machines. Bonin et al. (2015) transferred the format of the study to Germany and came to a similar conclusion of 42%. At the same time, however, they criticized the approach of the study, which was based among other things on expert assessments that overestimated the technical automation potential. They also criticized the fact that only certain activities could be automated by machines, but not entire professions, as these consist of a multitude of different areas of activity (ibid., p. i). Their correspondingly adjusted analysis of this set of facts concludes that in the USA 9% of the jobs have activity profiles with a *relatively high* probability of automation, while in Germany this characterizes 12% of the jobs. For both countries, the share of jobs with a *high* probability of automation is even lower. Bonin et al. (2015) emphasize, as do many authors of other studies, that technical automation will not just cost jobs, instead it will primarily change them and even create new ones. There will remain a plethora of activities that are difficult to automate and for which employees will have to be qualified, which further underlines the enormous importance of training and advanced qualification.

Both the study by Frey and Osborne and the one by Bonin et al. were selected because they are typical of the current preoccupation with the topic in several respects. Firstly, although gender is one of the variables collected, it is not differentiated in the analysis of the data. In this respect, the results do not allow any conclusions to be drawn about the extent to which the development affects women and men differently, although this can be safely assumed. Insufficient gender-specific differentiation in many studies on the impact of digitalization on employment makes it difficult to make solid estimates.

Secondly, most studies focus on traditionally male occupational areas, especially in the fields of industrial work, mechanical engineering, construction, and logistics, while there are hardly any studies on other sectors such as the service sector, in which a particularly large number of women are employed. For example, the caregiving growth sector shows that, on the one hand, it is a strongly gendered area of work with a disproportionately large percentage of women (Riegraf 2018; Aulenbacher et al. 2014). At the same time, other characteristics of social inequality such as origin, immigrational background, or quality of employment become relevant when predominantly Eastern European women permanently living in private households perform 24-hour support in the scope of live-in-caregiving without a sufficient degree of labor protection, or privacy. In addition to the issue of pay equity, considering these developments the further issues of precarious working conditions and the quality of employment in general must also be addressed, instead of truncating the discourse by reducing it to the issue of gender.

There is also a lack of precise research on commercial and administrative occupations, although these are among the occupations with the highest probability of digitalization and affect the employment reality of many women (Oliveira 2017, p. 25). The reason for this is probably the historical over-dominance of "**Industry 4.0**" in the terminology, which implies a confinement of the concept of "Work 4.0" to the industrial sector, although digitalization will have massive effects on the entire economic and employment sector. It can therefore be stated that the research situation shows a strong imbalance in favor of the industrial sector, while other economic sectors and fields of employment receive significantly less attention. This contrasts starkly with the clear importance of said sectors for the overall labor market and implies a bias to the detriment of so-called "women's occupations" (Scheele 2018). A blind spot we currently see when discussing the consequences of the pandemic on diverse social groups and gender. As many others the German sociologist Jutta Allmendinger formulated even for Germany that the pandemic will foster traditional role models and produce a backlash for gender equality for decades.[6]

A third weak point is the focus in above-mentioned studies on the purely quantitative consequences of digitalization for employment scenarios. Other equally important qualitative aspects of the digital transformation of work cannot be analyzed in this way and remain therefore excluded (Oliveira 2017, p. 68). This concerns the transition in practices and forms of work, the flexibilization of time and place, as well as the change in work cultures and hierarchies. However, it is these qualitative aspects that in sum describe the "logics" according to which work

[6]For more information see: https://newsroom.iza.org/de/archive/news/geschlechterrollen-in-corona-zeiten-kommt-es-zur-retraditionalisierung.

functions in different branches and positions in a company's organizational structure. It is precisely these aspects that reveal the subjective changes brought about by digitalization, which both men and women clearly feel in their employment relationships when, for example, activities change, people are allowed or expected to work from home, digital technologies impact communication, or further training expands the scope of tasks. Here in particular, are a multitude of gender-dependent aspects and dimensions to be found, which can lead to an inequitable development of this trend for men and women.

The examination of the two studies on the probability of job automation shows examples of shortcomings that require more detailed gender-differentiating analyses to describe the effects of digitalization on Work 4.0 for women and men with diverse social and professional backgrounds. Gender is one of several social **dimensions of inequality** such as education level, social background, and ethnicity, to name a few. Within the research literature on Industry 4.0 and work, however, gender and other categories of inequality are rarely considered. Notable exceptions are the work of Ahlers et al. (2018), Oliveira (2017), Bergmann et al. (2017), Bultemeier and Marrs (2016), Wischermann and Kirschenbauer (2015), and Matuschek (2016) and the many current studies on the global consequences of the pandemic on women all over the world.[7]

2.2 Gender-Sensitive Factors and Dimensions in the Context of Work 4.0

In addition to expert reports and studies on the topic of Work 4.0, there are a plethora of very different testimonies regarding the extensive debate on the "Megatrend" of digitalization and Work 4.0. Although gender-specific considerations rarely play a role in these, the illustrative study in *Sect. 2.1* showed that quantitative research approaches are unable to shed light on the "soft" factors that are required for an investigation of gender-specific effects in the context of Work 4.0. Therefore, it is necessary to sharpen the focus on qualitative characteristics and to consistently consider gender-specific effects in their societal context. This *blind spot* is where Fig. 2 comes into play.

Following the research for this article, central aspects and factors with relevance for a gender-sensitive approach to this topic were compiled. Figure 2 assigns them to the central fields of action "changing forms of work" and "changing work methods." Both areas have a direct impact on individuals—women and men in different

[7]For more actual studies see: the UN Data report; Azcona et al. (2020a, b). The United Nations (edts.) 2020: Shared responsibility, global solidarity: Responding to the socio-economic impacts of COVID-19 and the digital library: https://www.unwomen.org/en/digital-library/publications/2020/09/gender-equality-in-the-wake-of-covid-19 (accessed 20.01.21).

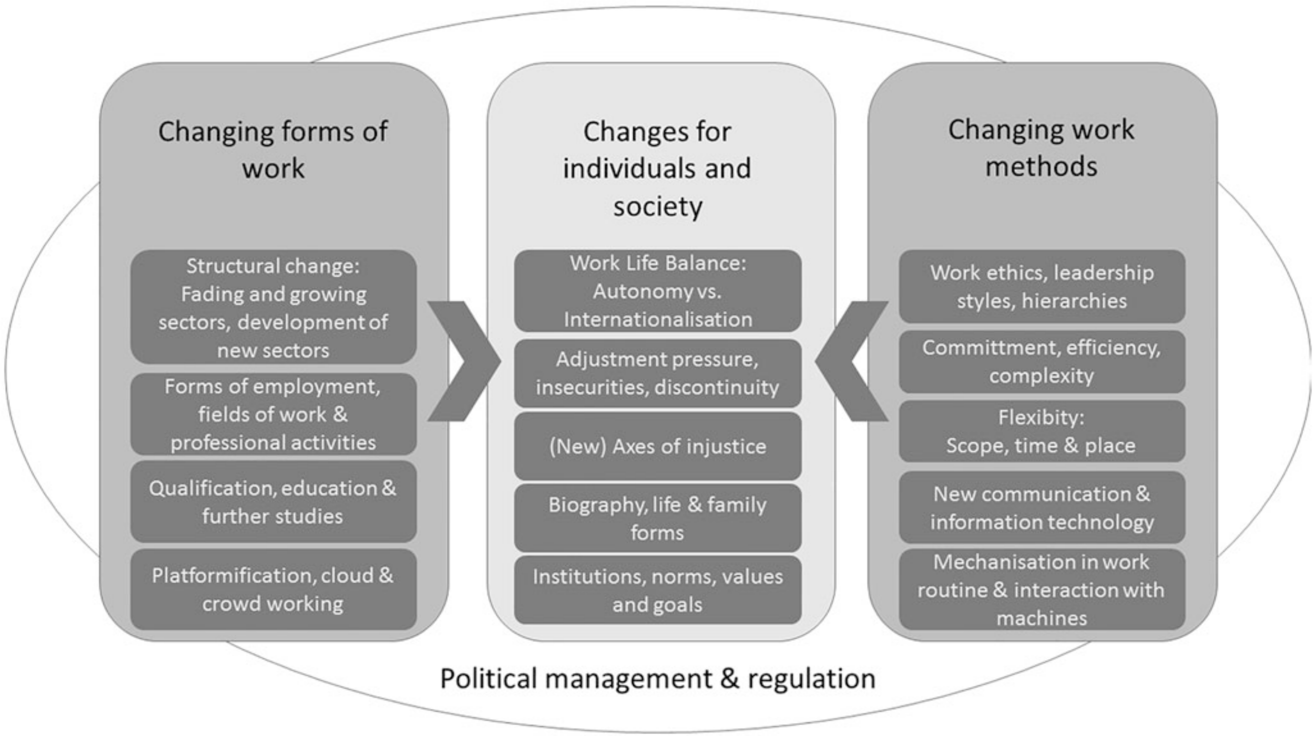


Fig. 2 Gender-sensitive factors and dimensions in the context of Work 4.0 (own illustration)

ways—as well as on society as a whole, which results in the third and central portion of the figure. Of great importance, especially for ensuring social justice, is the political framing of the processes through timely and appropriate elements of regulation. Unfortunately, it is precisely in this area, so integral for accompanying the processes of change, that politics have proven to be structurally incapable, due to lengthy and overly complex decision-making processes, of responding appropriately to the multi-layered challenges of digitalization.

The global consequences of the pandemic and the rapidly installed digitalization processes in many sectors clearly display the evidence of the above-mentioned dimensions. The pandemic has intensified domestic violence, whereas reduced income, limited social contact, and diminished access to support have significantly reduced the ability to execute any exit strategy (UN Women 2020).[8] Regarding a change in working methods, the sudden increase of people working from home also showed a reinforcement of known behavioral patterns in terms of domestic work. An increased unpaid workload due to homeschooling, increased meal preparation and caregiving activities, disproportionately carried out by women, has had an impact on the professional careers of many women, affecting their financial resources and inhibiting the transformation of income structures in the long-term (Keitel et al. 2020).

2.2.1 Changing Forms of Work

The *first* area of action comprises innovations and changes in *forms of work* that go hand in hand with structural change within the economy. While employees in the IT sector do not have to fear for their jobs, tax consultants will probably be replaced by optimized software. Studies on the probability of job automation and other contributions to quantitative changes on the labor market fall under this category.

This includes the occupation with industries that will change together with the **economic structural change**, but also the shift in forms of employment. For example, repeatedly emphasized is the fact that the profiles and requirements of many jobs and entire occupational fields are changing with the integration of new technologies, software, and automation. This has immense implications for the field of **qualification**. This is because employees must be trained to meet new requirements (Kagermann et al. 2013; Leicht-Scholten 2018). Access to education and especially to **continuing education**, which is experiencing a major increase in importance in the context of **lifelong learning**, is particularly vulnerable to disadvantages. The ability to provide suitable offers for diverse target groups will be integral in determining whether it will be possible in the long-term to meet the demand for labor on the one hand and, on the other hand, to assign as large a proportion of the working population as possible to a job that secures and satisfies

[8]For more information see: https://data.unwomen.org/features?field_theme_target_id=71.

their livelihood. This poses major challenges for our education system, especially in the technical sciences (Leicht-Scholten 2018).

The last point in the field of "changing forms of work" concerns entirely new forms of employment such as "**Cloud Work**" or "**Crowd Sourcing**" or **platformization**, in which work is performed via digital platforms (for example: Uber, airbnb, Deliveroo, Helpling, UpWork, MyHammer, etc.) which do not fit into our traditional categories of employment. Thus, workers are not employed, but merely mediated or flexibly contracted as needed. They themselves have the status of freelancers or solo-self-employed, not infrequently on the border of pseudo-self-employment and often in an extremely precarious economic situation. While these forms of work offer great flexibility and autonomy in time and place, they also mean a loss of commitment, security, and continuity. They affect, in different ways, low-skilled or unskilled workers, as well as well-qualified workers with specialized, in-demand skills (Scholz and Müller 2014). Here is the importance of a consideration of social justice beyond the category of gender particularly evident. In comparison to traditional forms of employment, such new forms of employment make it especially difficult for political management to fairly regulate and protect employees in terms of occupational health and safety, social security, and socio-political responsibility and legal liability for services, on the part of the platform providers (ibid.; Boes et al. 2014).

2.2.2 Changing Work Methods

The *second* key area of action in Fig. 2 refers to the aspects and dimensions of Work 4.0 that arise due to shifts in the way work is designed. These include changing **cultures of work**, leadership styles, and company hierarchies, but also aspects of **labor organization**. It is about *how*, *where*, *when* and *with what* work is accomplished. Such changes are triggered by new corporate strategies and goals or by the introduction of new **information and communications technology**, new software, machines, and robotics. These provide new opportunities for both companies, and their employees. "**Home office**," "sabbaticals," flat hierarchies, and working time accounts have the potential to be of great advantage to employees, or, in conjunction with massively increased performance requirements and availability expectations, to utterly disrupt their "**work-life balance**." The success or failure of an employee within this framework depends on the one hand, on the corresponding practices and contexts in which they are embedded, on the other hand, the individual preconditions they bring with them. As such, women and people in unpaid caregiving positions in addition to their gainful employment have particularly great chances, but also must contend with unequal opportunity structures to meet shifting requirements and expectations (Gesis 2020; Azcona et al. 2020a, b; Staab 2020).

This leads us to the *third* and central area of action in Fig. 2, which shows where and how the described changing forms of work and ways of working have an impact on the individual and societal level.

2.2.3 Changes for Individuals and Society

By way of explanation of the shift in forms of work and ways of working in the two preceding sections, several decisive points have already been mentioned, which can be used to systematize the effects on employees as individuals. The concept of work-life balance plays a central role, in which the ambivalent new innovations and opportunities find varied and abundant expression. Whether the latter proves to be an opportunity, or a risk is determined, on the one hand, by the surrounding contexts, such as the family situation, economic conditions, or the work environment in which they are embedded. On the other hand, individual conditions such as personality, social status, level of education, family of origin, ethnicity, physical limitations, religious affiliation, or even personal goals determine the way a person deals with these changes. Along with these factors, new **axes of inequality** beyond gender emerge, which must not be ignored in a well-founded discussion of the topic through inadmissible oversimplifications and the reduction of these complex issues to simply an issue of gender.

In *Sect. 3*, central discourses in this context will be illustratively illuminated and reflected upon.

3 Tomorrow's Technology, Yesterday's Discourses? Social Gender Discourses Between Redundancy and Abridgement

> Discourses about technology are always processes of legitimation and negotiation as well as part of a process of understanding and coping. They refer to models of work, society, and the human condition. Therefore, it is important for historical research as well as for current debates to understand that since the 1950s, similar patterns of argumentation can always be found. They not only illustrate the high status of gainful employment in society, but also reveal concepts of labor relations and content, ideas and expectations about the meaning of people in the labor process, and the respective concept of the labor society. (Heßler 2016, p. 19)

As initially mentioned, discussions about prospects and challenges of increasing automation of work processes are not only conducted in discourse around digitalization, but can be traced back to the 1950s, and from that point increasingly to the 1980s. Similar to the current discussions, two positions can be identified in the discourses of that time, which are correspondingly opposed to one another. On the one hand, the advantages of automation are emphasized and portrayed as a necessity for prosperity and progress by the business community, corporations, management, and engineers, while on the other hand, trade unions, social scientists, and the media emphasize the dangers of automation (ibid., p. 18).

There is agreement, however, that the new technologies will change the scope of occupational activity and professional profiles, and that new professions will be created. In terms of debates around the prospects for women, new chances are often identified due to, for example, greater **flexibility**, better compatibility, and flatter hierarchies (Ahlers et al. 2018). In terms of gender-specific attributes, there could indeed be promising potential here to break up existing inequality structures and create new design spaces to make work more gender-equitable. But what is meant by gender-equitable work? And under what conditions can these opportunities for women be turned into real improvements?

One of the central questions of gender justice concerns the question of changing gender-specific attributions. For though the technologies are new, the industries, contexts, and structures in which they are applied are not. At least in the interim, those who work with them will not be new either. At every step in this process, the new meets the old. Consequently, we are confronted with a situation in which technical innovations provide new impetus to change the familiar, thereby creating opportunities. Whether and how they are used, however, is largely determined by the existing structures and parameters of the evolved society with its norms and values.

Central themes of gender studies in the past decades in the context of gender equity and work were, and still are essentially the gender-specific division of productive and reproductive work, paid labor (market labor) and unpaid caregiving labor, the existing gender pay gap, as well as the underrepresentation of women in leadership positions.

Numerous studies on the separation of the spheres of professional and caregiving work and the associated gender segregation during industrialization have analyzed the impact of new technologies on the lives of women and men in the past (see, for example, Becker-Schmidt 2018; Weber 2018; Soiland 2018).

Since the beginning of the industrial revolutions, discourses can be traced throughout the technological progress, discussing the potential for changing gender relations in the labor process. Starting with discussions around the preoccupation with the mechanization of households and the idealization of housework and motherhood in the industrial revolution, through the obsession with efficiency through household technology and the stabilization of the middle-class nuclear family, to household technology as a symbol of modernity. However, a look at the history of technology in the home and its effect on gender relations reveals an ambiguity. On the one hand, the conception of housework as a purely female contribution to society enabled many women to enter the political sphere and thus made the, until this point, dichotomously opposed public and private sectors permeable. On the other hand, this reinforced the notion of exclusively feminine and masculine characteristics and thus solidified the middle-class gender-roles and their associated inequalities (Hausen 1976). In view of today's debates, the question of such ambiguity and risks arises again.

Discourses around gender equity and labor, as explained in great detail in the past (Scheele 2018; Riegraf 2018; Motakef and Wimbauer 2018; Aulenbacher and Décieux 2018), as well as in the current debates around digitalization, can be tied to three central themes: Equal opportunities, gainful employment versus reproductive work, and flexibilization of work.

The regulating obligation to implement or advance gender equality is now state of the art for most companies. Unfortunately, it seems this regulation overshadows the potential that women represent as a resource for companies. A study by the World Economic Forum states: "Female talent remains one of the most underutilized business resources [...]" (World Economic Forum 2016, p. 1). According to this 2016 study, the most common reason for companies to promote women is the ethical imperative of "fairness and equality" (42% of companies surveyed) (ibid., p. 1). In contrast, the expansion or increase of the talent pool, i.e., the use of the numerous existing highly qualified specialists, is cited by only a few companies as an argument for greater participation by women. This is particularly surprising given that in the information-technology sector, which has an especially high demand for skilled workers, women make up an average of 32% of junior staff, 21% of middle management, 11% at management level and only 5% of CEOs (ibid., p. 3). This means that even if companies have invested heavily in their employees, they are not able to develop and retain these talents. The lack of integration of women is therefore a conspicuous economic factor in the management of human capital.

When asked about the importance of strategies to better integrate women in companies, the three most frequently mentioned are: promoting work-life balance (38%), setting target numbers (33%), and further education and leadership training for women (ibid., p. 7). Yet more interesting and revealing, when asked about barriers that hinder women's careers, the same companies named, in addition to work-life balance, unconscious prejudices and stereotypes (**unconscious biases**) among managers, with 44% in each case. However, strategies for eliminating these stereotypical prejudices among managers are not to be found among the integration strategies mentioned (ibid., p. 8).

3.1 The Debate on Quotas

The recourse of monocausal explanatory and action approaches, which paint groups with a broad brush instead of considering the widely studied, mutually reinforcing, and complex structural mechanisms of discrimination, must be recognized as counterproductive. This can be seen, for example, in the highly emotional discussion about quota regulations. The constitutionally formulated starting point, according to which the quota only applies to applicants with equivalent qualifications, in fact excludes the preferential treatment of less qualified applicants, just as the concept of "equivalent qualifications" can always provide leeway for any selection decision (Leicht-Scholten 2000). However, the defamation of women as "token women" in these discussions reduces systematic issues to a matter of individual chance.

This is symptomatic of a truncation of the discussion around of equal opportunities, which neglect the existing macro- as well as micro-economic potential for the organizations themselves outside the normative discourse. As such, the existing human capital (see *Sect.* 3) and the potential of the positive correlation between gender diversity and corporate success, which has been stated in numerous studies, remains economically underutilized (McKinsey 2007; Catalyst 2004; Smith et al. 2006). In view of the opportunities presented by the digitalization process, it is an economic imperative to, through a broadening of perspectives, realize this potential to the benefit of all stakeholders (Leicht-Scholten et al. 2011).

3.2 The Substitute Discourse on Compatibility

The current discussion on compatibility (see also *Sect.* 2) obscures the dichotomous juxtaposition of paid work and unpaid care work and makes the latter a personalized "problem" of the individual. Women still take on the majority of caregiving tasks and experience a "double socialization" by assuming responsibility for both professional and family labor (Becker-Schmidt 2010). Instead of discussing the relevance of caregiving work for *all* employees, who take on or would like to take on such tasks, a stereotypical substitute discourse is conducted that ignores systematic contexts and obfuscates underlying origins. As a result of demographic change, the caregiving sector will continue to grow in importance. This, and the already existing shortage in caregiving fields, makes abundantly clear the urgent need for innovative perspectives that go beyond individual responsibility, and an understanding of these tasks as societal in nature, as is already the case in Sweden (Haberkern and Brandt 2010, p. 191). Current concepts in various industries in Sweden also show that a reduction in working hours with unchanged wages has, without causing a drop in productivity, precipitated an enormous increase in employee satisfaction. The reduction in working hours makes it possible to distribute tasks more flexibly and thus also break down stereotypical **gender roles**. In Germany, too, there are signs that traditional roles and responsibilities are being dissolved due to a **social value shift**. However, the strong structural tendency towards inertia counteracts this development and hinders positive changes in practice.

The utilization of parental leave by men has increased since the introduction of "Partnermonate," or partner months, in 2007 (Hobler et al. 2017). Young fathers, who increasingly want to participate equally in family and care work, thus face similar problems as young mothers. The issue of reconciliation between family and work-life thus no longer so much separates women and men, as parents and non-parents, caregivers of family members and non-caregivers.

And here the next level of this discourse emerges. As numerous recent studies on generational behavior show, work habits and the standards for quality of life have changed significantly over the decades. For many, the ideal of the successful professional has given way to the idea of the good life, in which work is one of several equal spheres, but no longer the central anchor point in life. The demands on the workplace itself have also changed, as Fig. 3 shows in detail. Employees increasingly see their work as a form of interpersonal exchange of understanding, recognition, and support, in which fairness and solidarity should be paramount (Kock and Kutzner 2014). At the same time, professional and personal biographies are becoming more diverse and disruptive, life trajectories more individualized.

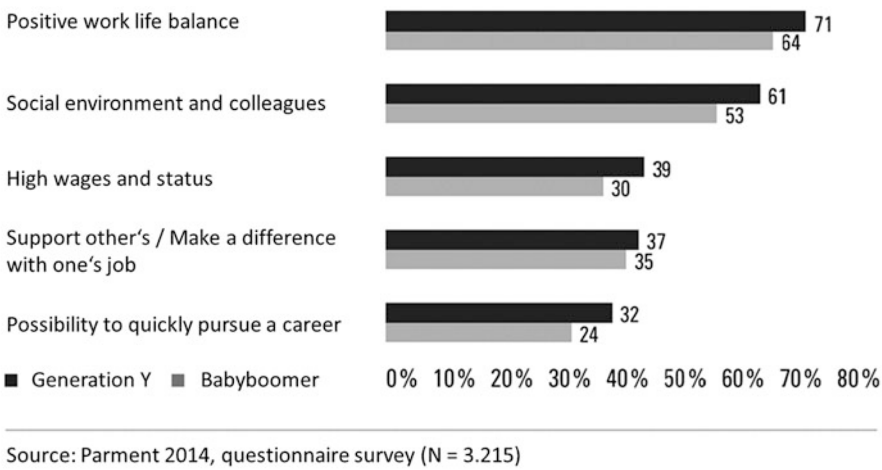


Fig. 3 Important aspects of the employer choice of baby boomers and generation Y (Lott 2017, p. 3)

3.3 *The Conversation About Flexibilization*

The promise of increasing flexibility in work processes is often discussed, especially in the context of gender, as a method to better reconcile work and family life. Work can now be done from anywhere and at any time. Irrespective of the aforementioned logical abridgement of the argument, it is evident that the traditional distribution of roles between the sexes has, on a fundamental level, hardly changed at all. Men are typically the main providers through full-time employment and especially in the case of families, while women usually take on the reproductive and caretaking tasks. And this despite the fact that normative gender roles are changing. While in 1982 only 32% of women (and men) in West Germany expressed egalitarian attitudes toward the division of roles between women and men, 89% of women and 84% of men in 2016 endorsed egalitarian gender roles (Blohm and Walter 2018). However, traditional gender images still dominate the social reality and women experience systematic discrimination in the labor market. Therefore, couples often distribute gainful employment and care work unequally to the detriment of women's employment and pension entitlements. The tax model of marital splitting, which is oriented toward the traditional distribution of roles, also has a structurally reinforcing effect here, offering little incentive for a dual-parent model and structurally counteracting individual efforts by couples to participate equally in work and family life. The political mechanisms, based on traditional role models, thus undermine social change.

4 Conclusions and Outlook

Whether specific technological changes will lead to social change is not determined by the technology alone, but rather by the normative and structural framework as well as the actors involved and their interaction with technological development. This implies as well that social transformations owing to technological developments absolutely require the inclusion of all members of society to be successful.

For example, in an exploration of caregiving work, it was made clear that an oversimplification of the discourse to merely a "compatibility issue for women" does not begin to address the urgent issue of our rapidly aging society. Furthermore, a shift towards understanding reproductive work not as a private, but as a societal task, combined with a fundamental societal discussion about the value a society attaches to reproductive work would not only "support" women. Indeed, political and structural models in Scandinavian countries show that such measures also make economic sense and benefit all members of society. In Sweden, for example, there is a social consensus that caregiving work, like reproductive work, should not be viewed as an individual task belonging only to those affected, but as a social task, which is also associated with the willingness to finance it. Furthermore, Sweden exemplifies the fact that a reduction in hours in the work week, without a reduction in pay, absolutely does not harm the economy of the country, but instead fosters compatibility between work and home life.

The example of gender and Work 4.0 has shown the challenges and opportunities that digitalization offers for social transformation processes. The major challenges lie in the fact that discourses on technology should be understood as legitimation and negotiation processes (Heßler 2016) and as part of a process of understanding and coping, which must also be understood and actively shaped as such by all contributing parties. This requires existing structural frameworks as well as socially established practices to be put to the test from the perspective of a "good life" for all.

It has also been shown that, despite all the normative discussions about equal opportunities, companies have not yet succeeded in fostering the advancement of excellently trained women in their companies. Steps towards achieving a gender-equitable organization must primarily begin on a systematic, structural scale within said organizations, not on the individual scale with women. For example, family-friendly structures in companies attract and retain not only women, but also men, who in turn want to take on family responsibilities. Furthermore, in this way individual categorization and stereotypes can be dismantled in the long-term, securing fairness and liberty for all employees.

Societal roles and divisions of labor are becoming more flexible. While the inclusion of women in gainful employment is growing, the standard working habits around full-time employment, job continuity, and social security are increasingly dissipating. Simultaneously, among the new forms of work, discontinuous and precarious employment relationships with new work models are on the rise. Particularly in younger cohorts, employees are experiencing an increasingly high degree of insecurity when entering the workforce due to this shift and often work under temporary, instable working conditions. Furthermore, in many households the increased labor market participation of women compensates for the loss of income of men, especially in areas of work requiring minimal qualifications. Changing roles and family models are leading to shifts in household structures. The classic family is progressively being replaced by single households, single parents and/or patchwork families. Winners and losers in these processes of change can therefore be identified along different dimensions of inequality, which must also be considered in their intersectional interdependence.

With the flexibilization of work concepts, we, as a society, are faced with an opportunity to break down existing gender-role stereotypes, create new flexible work and living spaces for women and men alike, and to reconceptualize the way we work. Indeed, the creation of new realms of opportunity through the increasing flexibilization of the labor market has the potential to catalyze the transformation to a society with equal opportunities, if both employees and employers recognize and take advantage of this chance. The swiftly changing attitude towards work, on the one hand, and the new standard ideals of a good life on the other, with which young employees today evaluate and select their employers, makes it clear that digitalization represents a decisive part of a very complex transformation process in our society, which addresses numerous other diversity dimensions in addition to gender.

Especially during this transitional phase to a more fluid gender structure in which we find ourselves, it becomes obvious that with the flexibilization of divisions of labor and gender-roles, discussions about inequality beyond the category of gender must also take place. For example, exchange between the highly educated, entrepreneurs, and precariously employed migrants. New axes of social inequality affect not only women, but also men and unfold due to their individual life situations, for example, with regard to cultural origin, education, forms of work and work habits, as well as family commitments.

The inclusion of gender perspectives in these negotiation processes can therefore serve as a blueprint for the inclusion of further dimensions of social inequality, such that this next step into the new digital age becomes a chance—the chance to recognize the discussion around technology as a legitimation and negotiation process, in which technological innovations are utilized by all members of society as drivers for innovation, to cope with the current global societal challenges.

The COVID-19 pandemic shows the urgency for action as well as the urgency for segregated data acquisition to globally support gender equality and the realization of the sustainable development goals (SDG).

References

Ahlers E et al (2018) Genderaspekte der Digitalisierung der Arbeitswelt. Bearbeitete Neuauflage zum Internationalen Frauentag. Hans-Böckler-Stiftung Arbeitspapier Nr 311. 2. Aufl, Düsseldorf

Aulenbacher B, Décieux F (2018) Prekaritäten: internationale Forschung zu globalen Ungleichheiten, Ungleichzeitigkeiten und Geschlecht. In: Kortendiek B, Riegraf B, Sabisch K (eds) Handbuch Interdisziplinäre Geschlechterforschung. Geschlecht und Gesellschaft, Bd 65. Springer VS, Wiesbaden, pp. 813–823

Aulenbacher B, Riegraf B, Theobald H (2014) Sorge: Arbeit, Verhältnisse, Regime. Care: Work, Relations, Regimes. Soziale Welt Sonderband, Bd 20. Nomos, Baden-Baden

Azcona G et al (2020a) From insight to action. Gender equality in the wake of COVID-19. United Nations Entity for Gender Equality and the Empowerment of Women

Azcona G et al (2020b) Ipsos survey confirms that COVID-19 is intensifying women's workload at home. https://data.unwomen.org/features/ipsos-survey-confirms-covid-19-intensifying-womens-workload-home. Accessed 21 Jan 2021

Becker-Schmidt R (2010) Doppelte Vergesellschaftung von Frauen. Divergenzen und Brückenschläge zwischen Privat- und Erwerbsleben. In: Becker R, Kortendiek B (eds) Handbuch Frauen- und Geschlechterforschung. VS Verlag für Sozialwissenschaften, pp. 65–74

Becker-Schmidt R (2018) Produktion – Reproduktion: kontroverse Zugänge in der Geschlechterforschung zu einem verwickelten Begriffspaar. In: Kortendiek B, Riegraf B, Sabisch K (eds) Handbuch Interdisziplinäre Geschlechterforschung. Geschlecht und Gesellschaft, Bd 65. Springer VS, Wiesbaden, pp. 65–75

Berger PA, Konietzka D (eds) (2001) Die Erwerbsgesellschaft: Neue Ungleichheiten und Unsicherheiten. Springer, Wiesbaden

Bergmann N, Lechner F, Gassler H, Pretterhofer N (2017) Digitalisierung – Industrie 4.0 – Arbeit 4.0 – Gender 4.0. Endbericht, Dezember 2017. Wien. https://www.researchgate.net/publication/324861390_Digitalisierung-Industrie_40-Arbeit_40-Gender_40_Endbericht. Accessed 27 Feb 2019

Blohm M, Walter J (2018) Traditionelle und egalitäre Vorstellungen zur Rolle der Frau im Zeitverlauf. Bundeszentrale für politische Bildung. http://www.bpb.de/nachschlagen/datenreport-2018/werte-und-einstellungen/278576/traditionelle-und-egalitaere-einstellungen-zur-rolle-der-frau-im-zeitverlauf. Accessed 3 Feb 2019

Bundesministerium für Bildung und Forschung (BMBF) (2017) Berufsbildungsbericht 2017. Paderborn. https://www.bmbf.de/upload_filestore/pub/Berufsbildungsbericht_2017.pdf. Accessed 27 Feb 2019

Boes A, Kämpf T, Langes B, Lühr T, Steglich S (2014) Cloudworking und die Zukunft der Arbeit. Kritische Analysen am Beispiel der Strategie "Generation Open" von IBM. Kassel

Bonin H, Gregory T, Zierahn U (2015) Übertragung der Studie von Frey/Osborne (2013) auf Deutschland. Kurzexpertise Nr 57. Endbericht. ZEW, Mannheim

Bultemeier A, Marrs K (eds) (2016) Frauen in der digitalen Arbeitswelt von morgen. München

Bundesagentur für Arbeit (2018) Die Arbeitsmarktsituation von Frauen und Männern 2017. Berichte: Blickpunkt Arbeitsmarkt. Nürnberg

Catalyst (2004) The bottom line: connecting corporate performance and gender diversity. New York/San Jose/Toronto

Federal Statistical Office (Destatis) (2018) Arbeitsmarkt auf einen Blick – Deutschland und Europa– Ausgabe 2018. Wiesbaden. https://www.destatis.de/DE/Publikationen/Thematisch/Arbeitsmarkt/Erwerbstaetige/BroeschuereArbeitsmarktBlick0010022189004.pdf?__blob=publicationFile. Accessed 27 Feb 2019

Federal Statistical Office (Destatis) (2019) Statistisches Bundesamt. https://www.destatis.de. Accessed 12 Jan 2022

Frey CB, Osborne MA (2013) The future of employment: how susceptible are jobs to computerisation? University of Oxford

Geissler R (2014) Bildungsexpansion und Bildungschancen. https://www.bpb.de/izpb/198031/bildungsexpansion-und-bildungschancen?p=all. Accessed 27 Feb 2019

GESIS (2020) Gender und Geschlecht. https://www.gesis.org/fileadmin/cews/www/download/cov19undGeschlecht_2020-10-19.pdf. Accessed 20 Jan 2021

Gemeinsame Wissenschaftskonferenz (GWK) (2011) Frauen in MINT-Fächern. Bilanzierung der Aktivitäten im hochschulischen Bereich. Materialien der GWK Bd 21. Bonn
Haberkern K, Brandt M (2010) Intergenerationale und professionelle Unterstützung älterer Personen in Europa. In: WSI Mitteilungen 4/2010
Hausen K (1976) Die Polarisierung der "Geschlechtscharaktere" - eine Spiegelung der Dissoziation von Erwerbs- und Familienleben. In: Conze W (ed) Sozialgeschichte der Familie in der Neuzeit Europas. Ernst Klett Verlag, Stuttgart, pp. 363–393
Heßler M (2016) Zur Persistenz der Argumente im Automatisierungsdiskurs. ApuZ 18–19/2016, pp. 17–24
Hobler D, Klenner C, Pfahl S, Sopp P, Wagner A (2017) Wer leistet unbezahlte Arbeit? Hausarbeit, Kindererziehung und Pflege im Geschlechtervergleich. Aktuelle Auswertungen aus dem WSI GenderDatenPortal. WSI Report. 35/2017. Düsseldorf
Kagermann H, Wahlster W, Helbig J (eds) (2013) Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Deutschlands Zukunft als Produktionsstandort sichern. Abschlussbericht des Arbeitskreises Industrie 4.0. Frankfurt am Main
Keitel C et al (2020) Die Corona-Gesellschaft. Analysen zur Lage und Perspektiven für die Zukunft. X-Texte zu Kultur und Gesellschaft, Transcript Verlag, Bielefeld
Kock K, Kutzner E (2014) »Das ist ein Geben und Nehmen«. Eine empirische Untersuchung über Betriebsklima, Reziprozität und gute Arbeit. Edition sigma, Berlin
Leicht-Scholten C (2000) Das Recht auf Gleichberechtigung im Grundgesetz. Die Entscheidungen des Bundesverfassungsgerichts von 1949 bis heute. Campus Verlag, Frankfurt/New York
Leicht-Scholten C (2018) Technikwissenschaften und Ingenieurwesen: der Beitrag der Geschlechterforschung für sozial verantwortliche Forschung und Entwicklung. In: Kortendiek B, Riegraf B, Sabisch K (eds) Handbuch Interdisziplinäre Geschlechterforschung. Geschlecht und Gesellschaft, Bd 65. Springer VS, Wiesbaden, pp. 699–709
Leicht-Scholten C, Bouffier A (2019) Mind the Gap – Industrie trifft Gender. In: Frenz W (ed) Handbuch Industrie 4.0: Recht, Technik, Gesellschaft. pp 1239–1259
Leicht-Scholten C et al (2011) Going diverse: innovative answers to future challenges. Gender and diversity perspectives in science, technology and business. Budrich UniPress, Leverkusen
Lott Y (2017) Flexible Arbeitszeiten: eine Gerechtigkeitsfrage? In: Hans Böckler Stiftung (Hrsg) Forschungsförderung Report Nr 1, Düsseldorf. https://www.boeckler.de/pdf/p_fofoe_report_001_2017.pdf. Accessed 14 Feb 2019
Matuschek I (2016) Industrie 4.0, Arbeit 4.0 - Gesellschaft 4.0? Eine Literaturstudie. Rosa-Luxemburg-Stiftung. Berlin
Mayer KU (1998) Lebensverlauf. In: Schäfers B, Zapf W (eds) Handwörterbuch zur Gesellschaft Deutschlands. Leske+Budrich, Opladen, p. 441
McKinsey & Company (Hrsg) (2007) Woman matter. Gender diversity, a corporate performance driver. O.O.
Motakef M, Wimbauer C (2018) Prekarisierung von Arbeit: erweiterte Perspektiven von Geschlechterforschung. In: Kortendiek B, Riegraf B, Sabisch K (eds) Handbuch Interdisziplinäre Geschlechterforschung. Geschlecht und Gesellschaft, Bd 65. Springer VS, Wiesbaden, pp. 783–791
Oliveira D (2017) Gender und Digitalisierung. Wie Technik allein die Geschlechterfrage nicht lösen wird. Working Paper Forschungsförderung Hans-Böckler-Stiftung, Düsseldorf
Riegraf B (2018) Care, Care-Arbeit und Geschlecht: gesellschaftliche Veränderungen und theoretische Auseinandersetzungen. In: Kortendiek B, Riegraf B, Sabisch K (eds) Handbuch Interdisziplinäre Geschlechterforschung. Geschlecht und Gesellschaft, Bd 65. Springer VS, Wiesbaden, pp. 763–773
Scheele A (2018) Arbeit und Geschlecht: Erwerbsarbeit, Hausarbeit und Care. In: Kortendiek B, Riegraf B, Sabisch K (eds) Handbuch Interdisziplinäre Geschlechterforschung. Geschlecht und Gesellschaft, Bd 65. Springer VS, Wiesbaden, pp. 753–763

Schneider H et al (2017) Erwerb der Hochschulreife und nachschulische Übergänge von Studienberechtigten. Studienberechtigte 2015 ein halbes Jahr vor und ein halbes Jahr nach Schulabschluss. In: DZHW (eds) Forum Hochschule 4/2017. Hannover

Scholz C, Müller S (2014) Zukunft der Arbeit: Cloudworker – ein Modell mit Risiken und Nebenwirkungen.Computerwoche. https://www.computerwoche.de/a/cloudworker-ein-modellmit-risiken-und-nebenwirkungen, 3063439. Accessed 23 Feb 2019

Smith N, Smith V, Verner M (2006) Do women in top management affect firm performance? A panel study of 2500 Danish firms. Int J Prod Perform Manag 55(7):569–593

Soiland T (2018) Ökonomisierung – Privatisierung: die verdeckte Unterseite neoliberaler Restrukturierungen und ihre Implikationen für die Geschlechterforschung. In: Kortendiek B, Riegraf B, Sabisch K (eds) Handbuch Interdisziplinäre Geschlechterforschung. Geschlecht und Gesellschaft, Bd 65. Springer VS, Wiesbaden, pp 95–103

Staab S (2020) Social protection for women informal workers. In: Chen M, Carré F (Hrsg) The informal economy revisited. Examining the past, envisioning the future. Routledge, London, pp 215–220

UN Women (2020) Will the Pandemic Derail hard-won Progress on Gender Equality? Spotlight on Gender, COVID-19 and the SDGs. https://data.unwomen.org/sites/default/files/documents/Publications/Spotlight-Gender-COVID-19-SDGs.pdf. Accessed 12 Jan 2022

Wischermann U, Kirschenbauer A (eds) (2015) Geschlechterarrangements in Bewegung. Veränderte Arbeits- und Lebensweisen durch Informatisierung? transcript, Bielefeld

World Economic Forum (2016) In: World Economic Forum (eds) The industry gender gap. Women and work in the fourth industrial revolution. Cologny/Genf

Wrohlich K (2017) There is a lot left to do to reach gender equality in Germany. DIW Wochenbericht43, pp. 427 f. http://www.diw.de/documents/publikationen/73/diw_01.c.567687.de/diw_econ_bull_2017-43-1.pdf. Accessed 22 Feb 2019

Weber L (2018) Arbeit – Leben: wechselseitiges Verflechtungsverhältnis aus Sicht der Geschlechterforschung. In: Kortendiek B, Riegraf B, Sabisch K (eds) Handbuch Interdisziplinäre Geschlechterforschung. Geschlecht und Gesellschaft, Bd 65. Springer VS, Wiesbaden, pp 87–94

WSI GenderDatenPortal (2018a) Teilzeitquoten der abhängig Beschäftigten 1991–2017. Hans-öckler-Stiftung. https://www.boeckler.de/51985.htm#. Accessed 05 Feb 2019

WSI GenderDatenPortal (2018b) Gründe für Teilzeittätigkeit nach Elternschaft 2017. Hans-Böckler-Stiftung. https://www.boeckler.de/51973.htm. Accessed 05 Feb 2019

The Digitally-Supported On-Site University

Birgitta Wolff, Viktoria Trofimow, and Stephanie Dinkelaker

Contents

Contribution to the English-language version of the Handbuch Industrie 4.0 - Digitalisierung in der Universität, Version BW/SD 31.12.20.

B. Wolff · S. Dinkelaker (✉)
Goethe University, Frankfurt a.M., Germany
e-mail: wolff@pvw.uni-frankfurt.de; dinkelaker@pvw.uni-frankfurt.de

V. Trofimow
DB Systel GmbH, Frankfurt a.M., Germany
e-mail: Viktoria.Trofimow@deutschebahn.com

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_55

Digitalisation is not an end in itself but improves the quality and allows more multi-faceted implementation of teaching, research, university administration and science management. This article uses the example of Goethe University Frankfurt to illustrate how digitalisation can be used conceptually and through concrete solutions to improve various fields of activity. The article begins with university-specific opportunities and challenges of digitalisation (Sect. 1). Goals and requirements are derived from the development of Goethe University's vision of the future (Sect. 2), and concrete projects in teaching, research, administration and management are outlined (Sect. 3). At the end, a summarising outlook is given (Sect. 4).[1]

1 Opportunities and Challenges

1.1 Opportunities: We Are the Ones Who Invent the Future

Digitalisation is a megatrend getting plenty of hype that affects our entire society. It is changing framework conditions in technological as well as legal, economic and socio-demographic terms. There were already studies on the consequences of new information and communication technology (ICT) for organisations in the last century (cf. Picot and Reichwald 1994; Picot et al. 1996). Modern information technology make it possible to carry out service chains across organisations independent of location. Value chains become distributed all around the globe. It sometimes does not matter anymore, for example, where programmers or designers are located. Many a company has become more 'virtual' in this way, creating new platforms and facilitating new types of organisations. Some companies will become larger, not least due to the new communication and control possibilities and network effects, while others will use increasing options for outsourcing to reduce in-house activities, essentially becoming smaller and parts of extensive networks. Companies react agilely to competitive pressure. They analyse and identify the potential for their fields of activity in view of these new opportunities. Non-profit organisations and the public sector—such as universities—do the same. The COVID-19 pandemic resulted in increased pressure for universities to use digital tools consistently:

[1] We would like to thank Dr. Olaf Kaltenborn and Sebastian Keil for their support in the final editing of the earlier German version as well as Jennifer Hohensteiner for most of the translation into the English version.

digitalisation became even more visible and essential. But this direction was actually clear even before coronavirus: Dräger and Müller-Eiselt described the radical change in learning in 2015 in their book 'Die digitale Bildungsrevolution' (The Digital Education Revolution): 'Just as the industrial revolution changed far more than production processes, the **digital revolution** will change not only learning processes but social structures as well. These great opportunities are also accompanied by great risks' (pp. 8f). In 2016, the Higher Education Forum on Digitalisation, carried out by the German Rectors' Conference (Hochschulrektorenkonferenz—HRK), the Centre for Higher Education Development (CHE) and the Stifterverband, with funding from the Federal Ministry of Education and Research (BMBF), published a report entitled 'The Digital Turn,' which is intended to support higher education institutions and **science policymakers** in this process of change. The report describes the digital turn less as a revolution and more as an evolution, and sees higher education institutions themselves as central drivers:

> The fact that higher education institutions in particular should embark on this path with a love of experimentation is clear from their institutional self-image and their social mandate. As research institutions and real-life laboratories, they are involved in the development of technological innovations, they initiate and analyse social change processes, and as places of innovation they are part of and drivers of overall social change (Higher Education Forum on Digitalisation 2016, p. 8).

Similarly, the Conference of Ministers of Education and Cultural Affairs (KMK) published a strategy 'Education in the Digital World' at the end of 2016. This strategy formulates how young people in schools, vocational training and universities should be prepared for the creative opportunities in the digital world. The strategy sees higher education institutions as important users of digital opportunities and at the same time as drivers of digital development (KMK 2016, p. 44). Thus, digitalisation should serve to fulfil the core tasks of higher education in teaching and research and not be an end in itself. Digitalisation should also improve administration and cooperation between higher education institutions. Knowledge is generated and disseminated at higher education institutions, so digitalisation is not only relevant to the development of strategy formation to support their core tasks, but is itself the subject of research (KMK 2016). At Goethe University, for example, there are research projects on green IT, financial trade systems, legal issues of digitalisation and the Internet, and—in a certain reflexive way—on the **digitalisation of learning**.

In Germany, higher education institutions are mainly funded from federal states' budgets. For years, the proportion of 'volatile' funding has been increasing in the overall system, as it has been at Goethe University: At Goethe University Frankfurt, the share of third-party funding acquired in 2018 amounted to more than 200 million of a total budget of around 640 million euros. And the basic funding share of the state is also volatile in the sense that it is tied to variable indicators, target agreements and projects. However, temporary funds can neither finance permanent professorships nor create or operate permanent infrastructures. There is a growing tension between the financing structure and current requirements, of which politicians are becoming increasingly aware. For example, the coalition agreement of the new Hessian state government emphasises digitalisation as a separate field of investment (2018 coalition agreement) and the state of Hessen, for example, operates the Competence Center for High Performance Computing in Hessen (HKHLR) together with the universities. Operating costs are also increasingly perceived as a problem area in their own right, especially energy costs for computers. These challenges are starting to be addressed by the federal and state governments.

Higher education institutions are places where direct exchange between scientists who share common values and serve non-commercial goals takes place. Higher education institutions are not only described as units with their own goals, but are understood as networks in which members use their connections to achieve their individual goals (Ehlers 2018). Under Article 5 of the German Constitution, professors are—for good reasons—guaranteed a high level of individual freedom in their work. However, all members of a higher education institution benefit from the network and the infrastructure of the institution.

The success of higher education institutions is measured above all by the quality of their research and the quality of their graduates. Unlike companies, higher education institutions do not aim to achieve financial gains by investing in digitalisation. **Digital strategy building** helps the members of higher education institutions at another level: through good (including digital) infrastructures as well as through **science-supporting organisational cultures** and agile leadership styles, they can more effectively achieve their individual and collective goals. Digitally supported profiling of this kind can lead to increased attractiveness for outstanding scientists and scholars, as well as for students. This in turn promotes research success and third-party funding, as well as raising the quality of study conditions through digitally optimised teaching by renowned and highly qualified personalities.

1.2 Challenges: The Shackled Driver

With close to 50,000 students and a good 150 courses of study at 16 faculties in the humanities, social sciences, natural and life sciences and medicine, Goethe University is one of the largest universities in Germany. In 2012 a good 43,000 students were enrolled at Goethe University, and by 2017 the number had risen to over 48,000. Unfortunately, this has also led to a deterioration of the professor-student-ratio, as budget and staffing levels have not grown to the same extent. While the ratio of professors to students was 1 to 83 in 2012, it had dropped to 1 to 95 by 2017. In contrast, the ratio of total academic staff to students has remained almost constant at around 1:20. The non-professorial teaching staff grew disproportionately; at the same time, there is a lack of professorships. As the number of students increases, the heterogeneity of students at Goethe University has also increased. At a good 30%, the proportion of students with a migration background at Goethe University is significantly higher than the national German average, according to the second university-wide student survey at Goethe University in 2017/18 (Opitz and Lommel

2017). Moreover, almost two thirds (64%) of students hold jobs parallel to their studies. Of the students working while studying, 83% said that they were working to finance their living costs (Opitz and Lommel 2017).

The size of Goethe University, its personnel structure with its overall decreasing number of staff, and the growing heterogeneity of the student body bring special challenges in addition to opportunities in the implementation and use of digital technology. These have to do with issues such as functionality, data protection, communication, consistency and the relatively slow pace of transformation compared to companies and smaller organisations. Challenges also arise from a university-specific system paradox: a university is characterised by a great openness to the new and at the same time by a critical examination of this newness. 'Change management' is gaining a new significance through digitalisation, as changes have become more complex than before. In a survey of Hessen's higher education institutions on the chances of digitalisation for improved teaching from 2018, the greatest challenges cited were acceptance by students and teachers (15%), financing (16%) and questions of copyright and data protection (15%) (Hessian Ministry of Higher Education, Research, Science and the Arts 2018).

The implementation of a stringent digitalisation strategy in an organisation such as Goethe University requires considerable investment. As non-profit organisations, state universities have no means for achieving economic digitalisation gains. This distinguishes higher education institutions from companies which, by choosing a suitable digital strategy, gain a competitive advantage over competitors on the market. Also, contrary to some speculation, it is not to be expected that digitalisation will save funds at higher education institutions: the services provided by the organisations will become more diverse and better, but not cheaper. In particular, the energy requirements for computers and digital equipment as a whole—not only at Goethe University—will continue to rise considerably. According to studies by the Fraunhofer Institute, the energy consumption of information and communication technology (ICT) in Germany in 2016 was 60 TWh per year. This corresponds to approx. 10.5% of the nationwide electricity consumption. The 50,000 computer centres in Germany alone account for more than 10 TWh, with a strong upward trend. This enormous, ever-increasing hunger for energy means that in Germany alone, four coal-fired power stations are needed to supply the data centres (Center for Scientific Computing n.d.). Goethe University's energy costs are around 20 million euros per year—mainly due to the ever more extensive and differentiated computer architecture and infrastructure. This also illustrates the need for attention to the topic. In this context, the fact that the new Goethe mainframe computer realised in 2019 is one of the most energy-efficient in the world is of great significance (cf. Sect. 3.2.8).

An overview of the special challenges of digitalisation at Goethe University includes the following points:

- Finding suitable structures for the size of the organisation
- Intangible objectives, no market pressure
- Implementation speed perceived as slow in networking, communication and establishing consistency in the system

- ‘Transformation costs’: People need to be convinced and must be trained
- Productivity paradox: Investment in digitalisation does not necessarily lead to immediate productivity gains
- Cost paradox: Despite the promise of increased efficiency, digitalisation investments do not lead to cost reductions but to cost increases
- Rising energy requirements for computers and digital devices
- Financing bottlenecks: no possibility to make profits from digitalisation
- Increasing data protection requirements and data safety risks for the infrastructure

In particular, neither the energy requirements for computers nor the new investments in relatively short time cycles resulting from depreciation or wear and tear have been adequately reflected in universities’ financing models in the past. Appropriate and incentive-compatible financing concepts must be developed with the involvement of the federal states, the federal government (e.g. within the framework of the Joint Science Conference, GWK) and third-party funding providers.

2 University of the Future

2.1 Vision and Organisational Culture

To bring a digital strategy to life, it is important that the entire university positions itself for digitalisation, provides the appropriate resources and establishes the necessary governance. This strategic positioning will vary considerably depending on the type of higher education institution; it is an opportunity for strategic differentiation. For Goethe University, it is clear that digitalisation does not mean turning away from the model of a lively, discourse-friendly **university with in-person, on-site teaching and research.** Not just tradition, but also the current self-image of most members of the Goethe community speaks for this view, which is the basis of the organisational culture. Goethe University remains a physical place of science and scholarship, where personal communication and networking are essential. Physical university campuses remain non-virtual meeting places for personal exchange. Yet, there will be a transformation: Lecture halls and offices will in future—at least in part—become collaborative spaces and ‘*experience centres.*’

Goethe University sees itself as a **learning organisation** (Goethe University 2014a), which it has also anchored in a general mission statement. In a learning organisation, the members share a vision and continually move beyond old thought patterns. Members communicate openly across hierarchical boundaries and put departmental and personal interests aside to achieve the shared vision of the organisation together (Senge 2006). To come closer to the ideal of a learning organisation, a strategy for innovation and change is crucial. One of the tasks of management is to influence the organisational structure and culture accordingly. The mission statement formulated in 2014 emphasises the vision of a cosmopolitan lab of the future in the centre of Europe, with an emphasis on interdisciplinarity and participation (Goethe University 2014a).

Digital leadership as a participative management style in university administrative structures encompasses aspects of responsibility, decision-making, results, information, goal setting and assessment, error and conflict management, and change (Van Dick et al. 2016, p. 22). As '*digital leaders,*' managers fill a wide range of roles. Depending on the situation and context, tasks are taken on permanently or temporarily; above all, however, the competencies of managers and employees are networked. This is in contrast to traditional leadership models, in which managers delimit responsibilities in their area by defining responsibilities and powers with the price tasks beyond the hierarchy and conflicting activities leading to friction and conflict. In digital leadership, decisions are increasingly made based on principles and processes rather than positions and hierarchies. This mode of leadership results in tension with the traditional organisation of state universities, where hierarchies are still effective.

Digital leadership creates the framework for high transparency and availability of information. Employees have an obligation to bring themselves up to date with the latest information. In goal setting and assessment, focus lies in equal measures on the employees and the team. There is continuous exchange and feedback that includes common goals and individual contributions, and cooperation and behaviour of individuals count along with the quality of the process results. In the case of mistakes and conflicts, managers ensure binding processes for learning from mistakes and for the productive resolution of conflicts. They support, moderate and create the framework in which results are made available. In contrast, in traditional leadership, errors and conflicts are suppressed by rules, consequences and control. As a result, the scope for creativity and originality tends to be limited. In the leadership model of digital leadership, independent action is strongly supported by a sustained high willingness and ability to change (Van Dick et al. 2016). Ehlers (2018) took stock of digital leadership at higher education institutions and found that little research has been done on digital change in educational organisations. Due to their limited terms of office and rotation, leadership at higher education institutions have only limited influence on the transformation of their organisations. They can make a difference through their personal style, their ability to establish relationships and build trust, their communication practices and the integration of their work team at their institution (Ehlers 2018). Ehlers sets out a framework for digital leadership with various dimensions relevant to the digital transformation of higher education institutions: first, positioning the institution; second, creating a vision and strategy; and third, a favourable culture. The **transformational leadership style** is described as a suitable approach for the digital transformation of higher education institutions, as it places particular focus on communication and participation, which are especially helpful for the development of a functional higher education organisational culture under the conditions of a digital world (Ehlers 2018, p. 27).

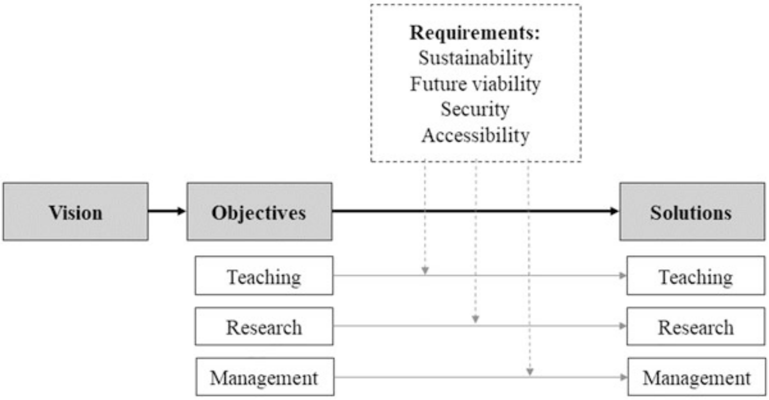


Fig. 1 Objectives, requirements and measures of digitalisation

2.2 *Aims, Measures and Requirements for Digitalisation*

At Goethe University, principles for teaching and studies (Goethe University 2014b), as well as a 'Leitbild Digitale Lehre' (Digital Teaching Mission Statement) (Goethe University 2018a) describing this vision have already been formulated, although it needs to be further specified and discussed. In this way, goals for the fields of teaching, research and administration/management are derived from the vision of the digitally supported university. These operational goals include, for example, a university-wide integration of data protection rules, appreciation and incentive concepts for the integration of digital media in teaching, qualification offers, and the establishment of various other support structures for the use of digital tools in teaching and research. These goals must be pursued—as a kind of secondary condition—while taking explicit requirements into account. Figure 1 illustrates this connection.

The requirements concern sustainability, future viability, security and accessibility and can be specified as follows.[2] Sustainability means that every digital service that is part of the university's standard provision is designed to be permanently supportable in terms of hardware, software and human resources. Necessary hardware upgrades are to be included in the purchase price, as are wear and tear and consumable materials. Updates of the software are to be ensured by contractual agreements with the providers or, if the software was developed internally, by sufficient capacities of the programmers. Personnel development is to provide for the permanent employment of qualified system administrators and developers after a probation period to avoid poaching offers from the industry and a resulting *brain drain*. A limiting factor in this respect is regularly found to be that in competition

[2] We would like to thank PD Dr. Jeannette Schmid for her contribution to this section.

with the private sector, the pay structures in the public sector are not sufficiently attractive to qualified personnel.

Digital developments are accelerating. For their future viability, it is important that the structure and operation of the systems used comply with binding conventions, so that migration to new systems is possible with as few losses as possible. The IT applications used are to be regularly evaluated to determine whether they still meet the current requirements of students, teachers, researchers and the administration. At the same time, new IT developments are to be continuously monitored and application scenarios tested. New applications can also be developed as '*in-house'* projects. Here, the criteria of security, sustainability, accessibility and future viability also apply.

The concept of security includes legal certainty in the acquisition and operation of digital applications. Data must be protected against theft, interference and loss by securing physical and virtual access. There must be a back-up for all data. It must be able to be restored in case of emergency. Services which are available for the entire university will be technically designed in such a way that a high number of users and a high number of simultaneous accesses as well as different authorisation structures are possible. Central services are always to be managed by several people who stand in for each other in case of holiday or illness.

Accessibility means that digital interfaces are designed to be accessible to users with special demands. Their presentation must consider the needs of blind and visually impaired people as well as those of the hearing impaired or deaf. This refers not only to the retrieval of information, but also to active interaction with the web interface. The access to the university's services is platform-independent on the client side. It is independent of the operating system and also possible with mobile devices. Accessibility is to be considered not only for centralised services, but also for decentralised services and materials. Support, also in the form of training courses, is provided for this purpose.

This conglomeration of goals and demands gives rise to concrete measures, instruments and projects at Goethe University; a selection of these is presented below.

3 Transition: Measures on the Path to a Vision

The following three sections describe exemplary measures in teaching, research and administration in more detail.

3.1 Teaching 4.0

To increase appreciation and motivation for the integration of digital media in the teaching/learning process, various formats exist at Goethe University, such as teaching awards and teaching sabbaticals. There are regular forums in which Goethe University members can exchange views on the integration of digitalisation in teaching of **e-learning** , such as the e-Learning Network Day, subject-specific dean's meetings and future workshops. Various support structures, continuing education programmes and workshop formats are available for the continuing education of teachers and staff at the central e-Learning facility *studiumdigitale*, the University Computing Centre (HRZ) and the Interdisciplinary College for University Didactics (IKH). Table 1 provides an overview of selected measures to support digital teaching. In each case, the problem to be solved, an idea for a solution, and its concrete implementation are presented. Explanations follow for measures in the table that are not self-explanatory.

A high proportion of the digitalisation measures required for Goethe University are carried out in close cooperation with the other Hessian higher education institutions and the state ministry in charge, to be able to make better use of the scale and network effects typical of digitalisation. Creating shared standards and platforms instead of isolated individual solutions will be a prerequisite to really benefit from digitalisation.

3.1.1 Digitally Supported Teaching and Learning in Hessen

The Hessian State Parliament decided to better exploit the opportunities offered by digitalisation for higher education teaching (motion 19/1796, Hessian State Parliament 2015). Therefore, Hessian universities want to jointly develop the didactic and technical prerequisites for the long-term use of new technologies in teaching and studies in the project 'Digitally Supported Teaching and Learning in Hessen' (2018).

In this project, prototypes for development and cooperation formats between the Hessian higher education institutions are being created, as well as a web portal with digital teaching-learning services. To this end, existing activities and structures within the higher education institutions will be developed into efficient service points for digital teaching, which will form the interfaces for both internal and general networking. Networking is coordinated by a newly established service facility jointly managed by the higher education institutions. In addition, subject-specific innovation forums ensure the further development of didactic concepts in the context of new technologies. Topics include needs assessment, web portal, didactic concepts, e-assessments, learning analytics, qualification of teachers and learners, effectiveness analysis and digital barrier-free study.

3.1.2 Digital Teaching as a Model

In March 2018, the principles of teaching and study at Goethe University (2014b) were supplemented by a mission statement on digital teaching in which Goethe University, as a university with in-person, on-site teaching and research, commits itself to a teaching and educational concept in which digital media are a natural part

Table 1 Measures to support digital teaching

Measure	Problem	Solution idea	Concrete content	Status
Digitally Supported Teaching and Learning in Hessen	Expensive development and implementation of didactic and technical teaching-learning concepts	Using synergies between Hessian universities	Joint platform and thematic innovation forums	Ongoing
Digital teaching as a model	Lack of a common vision of objectives	Participatory development of a mission statement	'Digitally supported university'	Implemented
DOIT	Technology fixation	Systematically holistic organisational approach	Networking of university didactics, infrastructure, lecturers and students	Ongoing
e-lectures	Fixing students to places and times	Virtual classroom	Two video systems	Ongoing
e-exams	Old-fashioned audit methods do not reflect modern working methods	Inclusion of new formats and tools (e.g. Excel)	Scanner exams, 'E-Plus' (e-exams)	Ongoing
DELTA	'Proliferation' of decentralised educational initiatives	Efficiency-oriented, science-based monitoring and roadmap	'Group Concept Mapping,' survey (method)	Ongoing
GInKo	Outdated campus management system, lack of transparency of student data, improvement of organisation in study management	Introduction of an effective and legally compliant integrated campus management with self-service function for departments	Complete implementation of the campus management system HisInOne	Ongoing
Online Study Choice Assistants (OSA)	In some cases, there is a poor match between prospective students and the chosen subject	Provide online information and self-tests for prospective students	Specialised OSA with information in the form of videos and text, self-reflective questions and test tasks	Ongoing
Study success in dialogue (SID)	Drop-outs are identified too late	Collection and use of subject-specific indicators for personal contact with students at risk	Visualisation of study progress data via a key data portal for use in individualised and targeted study guidance and quality management	Ongoing
Modifying incentives/ Appreciation	New teaching formats require a lot of time and are not	Teachers can apply for extra	'Time for Teaching' (reduction of teaching fees),	Ongoing

(continued)

Table 1 (continued)

Measure	Problem	Solution idea	Concrete content	Status
	always appreciated enough	developing time for innovative courses	e-learning support fund (also for students), 1822 teaching award	
Exchange formats	Innovative teaching ideas are not widely used	Exchange of best practice examples	e-Learning Network Day, dean's meetings	Ongoing
Teach the Teachers	Teachers need ongoing qualification for modern teaching	Didactic and technical training of teachers	e-learning certificate, didactic continuing education forums	Ongoing
Qualification of students	Writing counselling supports students in academic writing, but has no insight into individual progress	Software to visualise student writing	Visualisation of academic writing with VisaS	Aspired
Nationwide student ID	Lack of observability of educational pathways	Organisational boundaries should not be a limiting factor	Uniform ID	Aspired

of the course of study (Goethe University 2018a). The mission statement was developed in a participatory manner in several all-day workshops and future workshops with all status groups under the leadership of the Vice-President for Studies and Teaching.

The mission statement is aimed at teachers and students at Goethe University, because the dynamics of the development of digital processes and their design pose challenges to all university members—in technical and intellectual, legal and ethical terms. Participation in digital teaching is to be actively enabled for all university members; the university aims to achieve the highest possible qualification and support for teachers and to establish barrier-free systems and environments. Research into digital teaching is also a high priority for Goethe University.

3.1.3 DOIT Model

The DOIT model defines conditions for the success of digital teaching-learning processes (Horz and Schulze-Vorberg 2018). It addresses **didactics, organisation, the individual and technology (DOIT)**, which are not understood in isolation but in interaction with the other areas. Didactics emphasises the need to train teachers, while organisation is intended to ensure that all institutions involved are closely interlinked. At the level of the individual, the best possible acceptance for the use of digital media should be sought. Furthermore, the technical infrastructure should be adequate and adapted to the actual learning/teaching needs. The integrated perspective is important because traditional and frequently technology-driven programmes can lead to the creation of white elephants.

3.1.4 Digital Teaching/Learning Measures: E-Lectures and E-Exams

The services provided by three Goethe University institutions complement each other to successfully implement digital teaching. The HRZ provides the infrastructure, the IKH offers didactic courses, and *studiumdigitale* supports e-learning as an innovation driver. *Studiumdigitale* tracks down innovations in teaching, implements them on a trial basis and, if successful, transfers them into regular operation, e.g. in **teaching evaluation**, **e-lectures** and **e-exams, Inverted Classroom** and **Online Study Choice Assistants (OSA)**. Regular network meetings of the university's e-learning actors take place at Goethe University. An e-learning support fund for teachers and students has been in existence for ten years, and supports projects with digital media in learning and teaching processes in the departments. Student-only projects also benefit from this.

More than 100 rooms are technically connected to the two media technology control centres on the Westend and Riedberg campuses. There are 200 remote-controlled cameras in the lecture halls to record up to 25 lectures simultaneously as e-lectures. These recordings are in great demand among students and lecturers; all 16 departments make use of this option. Even before COVID-19, up to 90 lectures per week were being recorded in the lecture halls in the current semester, and since 2011 over 1.5 million hits have been registered. On this base, complemented with some fast adjustments and extensions, the University turned out to be well prepared for the requirements of COVID-19-compatible teaching start in March 2020.

Electronically supported examinations allow the simple creation of exam variants, save time during correction, and provide fast notification of results and evaluation objectivity. In the winter 2018/19 semester a total of more than 10,500 computer-based **e-examinations** were held at Goethe University. Overall, the use and number of participating departments is increasing. As an alternative to e-examinations, paper-based e-examinations have been available since 2010 to facilitate examination processing. Over 14,000 scanner exams were written in 2016, for example. In 2020, first experiments with remote e-exams followed, offering students an option to write exams at home under official 'exam-conditions.' Here, the challenges were posed by the German privacy requirements, not by technology.

3.1.5 DELTA

In cooperation with the Leibniz Institute for Educational Research and Educational Information (DIPF), Goethe University has launched the DELTA project (*Toward Digital Education with modern Learning Technologies and Assessment*

Approaches). The DELTA project is supported by the *Freunde und Förderer* (Friends and Sponsors) of Goethe University. The project—although not yet completed—is already providing information that influences the further processes of digital transformation. The DELTA project is focused on giving as many members of Goethe University as possible the opportunity to contribute to the strategy development process with their views, needs and concerns via semi-standardised interviews. One finding is that students expect increased flexibility and individualisation of their studies as well as stronger support for independent learning.

In addition to the opportunities offered by digitalisation, fears of increasing schooling, diminishing equality of opportunity due to expensive hardware and software, and inefficient or inconsistent use of digital possibilities were expressed on several occasions. In addition, a questionnaire was developed for students on the ideal use and expected use of data from digitised higher education **Learning Analytics** . The results available so far indicate that students generally have a positive attitude towards the use of Learning Analytics and also have a positive view of its use by Goethe University.

Parallel to this, a method called **Group Concept Mapping** (GCM) was used in the second phase. The aim of this method was to collect success factors for the design of digital higher education in an online brainstorming session open to all members of Goethe University.

Workshops will be held to identify achievable opportunities and long-term challenges for digitalisation at Goethe University. This will result in a *roadmap of* coordinated further measures which should meet with broad acceptance.

3.1.6 GInKo: Goethe University Information and Communication System

The aim of the GInKo project is to use a suitable **campus management system to** make the organisation of **study management** effective, transparent and legally compliant, and to support the strategic planning of the university with reliable and timely information. The project focuses on the introduction of an integrated campus management system to support the processes in study management. In a first project phase (2012 to 2014), the actual processes and requirements in study management were surveyed and analysed across the university, the decision for the iCMS software was made, the framework for the target concept was set and initial target processes were developed.

The situation analysis showed that study management is highly fragmented in organisational terms that some structures are redundant, and that interfaces between central and decentralised units are often poorly defined. The processes of course management, examination management, and the degree of utilisation of IT support are highly heterogeneous. Accordingly, the implementation of an integrated campus management system is mainly an organisational development project in which the focus is on basic, university-wide process and organisational development and harmonisation in the area of overall study management in addition to IT introduction.

Upon completion of the project, a high-performance, Bologna-compliant system geared to the requirements of Goethe University will be in place that will optimally support the processes of the university's study management and considerably simplify the reporting system. The implementation of the system will also be used to address organisational weaknesses in the process organisation of study management.

3.2 Research

Research also benefits from digitalisation. Digitalisation in research aims to store, organise and exchange data in the long term. Digitalisation can change the research system and also the incentive system. In this context, digitalisation means that the quality of publications takes precedence over quantity. At present, scientific data stocks are still frequently stored in a project-based, temporary and decentralised manner. Digitalisation can make these data systematically available, e.g. through a suitable **research data infrastructure**. There is a programme for this at national level: the National Research Data Infrastructure (NFDI). The NFDI aims to set standards in data management and, as a digital, regionally distributed and networked knowledge repository, to make research data sustainably secure and usable (Deutsche Forschungsgemeinschaft 2019).

Furthermore, digitalisation in research opens up new opportunities for **open science initiatives**. These include **open source**, **open research, open access**, **open data**, **open substances** and **open notebooks**. The Internet and the possibility of electronic publishing and the ensuing fast and easy dissemination of documents have made the issue of free access to scientific information within the framework of open access a topical one. At universities and non-university research institutions, the question of whether research results, which are largely financed by public funds, should be sold to publishers for publication, including rights of use, or made available to the general public with open access solutions is thus becoming increasingly important.

Similarly, in research support measures, Goethe University works closely with the other Hessian universities and the responsible state ministry in numerous collaborative projects to exploit the economies of scale and network effects already mentioned. Particularly in the case of database projects and information systems, it is important to proceed jointly at the planning and design stage. There are also significant synergies in operation. By adopting a coordinated approach, Hessian universities aim to exploit potential efficiency gains. Table 2 provides an overview of exemplary measures, followed by explanations.

Table 2 Measures to support digitalised research

Measure		Problem	Solution idea	Concrete content	Status
Joint projects of Hessian higher education institutions	HKHLR	Loss of efficiency in decentralised solutions	Cooperation at state level	Central provision of high-performance computing (HPC) capacity	Ongoing
	Hessenbox	No secure shared data storage	Hessen-wide legal and data protected 'drop box'	Cooperative sync-and-share infrastructure for Hessian higher education institutions	Ongoing
	HeFIS	Lack of transparency about research activities internally and externally	Common database, including doctoral student management	Documentation and database management	Ongoing
	HeFDI	Inefficient use of research data	Common research data infrastructure	Develop organisational and technological processes	Ongoing
	MobiDig	Various central systems for study services that have evolved historically	Bundling and mobile development of various services, app development	Connect student app and library research, at the same time linking digital resources from different universities	Ongoing
	LaVaH	Transience of digital media	Gradual development of an infrastructure for long-term availability	Conversion of data streams, archiving and risk management	Ongoing
	HeIDI	Multiple ID management systems	Merging of UB-ID management with university ID management	Create conditions for the use of DFN-AAI services	Aspired
CSC		Huge demand for high-performance computing capacity and research into green IT technology	Computers both as research objects and infrastructure for large-scale research projects	Energy-efficient mainframe computer	Ongoing
CMMS		No methodological basis for the use of highly informative data collected from novel imaging technologies	Data infrastructure makes scaling possible	Multi-scale modelling for new experimental approaches	Aspired

3.2.1 HKHLR: Hessian Competence Center for High Performance Computing

The continuation of the Hessian **Competence Center for High Performance Computing** (HKHLR) is intended to expand and consolidate the successfully initiated structure and working method for the provision of consulting and scientific services to enable more researchers in Hessen to make efficient use of high-performance computing. This will be done in implementation of the recommendations of the German Research Foundation (DFG) review commission. The HKHLR strengthens the governance structure with regard to a Hessian HPC state concept that considers both Tier 2 and Tier 3 systems when establishing national HPC centres. The project is led by the Technical University of Darmstadt.

3.2.2 Hessenbox: A Cooperative Sync-and-Share Infrastructure for Hessian Higher Education Institutions

The technical possibility for cooperation and the exchange of information it necessitates—especially in the form of files—are in urgent demand by all universities and scientific institutions. Indeed, access to files that are synchronised live from several devices, including mobile devices, is absolutely necessary for efficient IT support. Users have been accustomed to this for years through the use of their private systems and expect this functionality in their business environment as well. For this reason, many users use commercial **sync and share services** for their business purposes; however, this is not acceptable from the point of view of data protection and IT security. To offer an alternative that better complies with the applicable data protection laws and the principles of IT security, the state higher education institutions in Hessen are jointly setting up a sync-and-share service under the leadership of the University of Kassel.

3.2.3 HeFIS: Hessian Research Information System

The higher education institutions in Hessen want to jointly introduce a **research information system (Forschungsinformationssystem—FIS)** at each of their respective locations. Research information systems are web-based database applications that bring together information on research activities at higher education institutions including corresponding contextual information. Eleven strong research universities in Hessen have joined forces in the HeFIS network to procure, introduce, and further develop research information systems in a joint and coordinated manner.

3.2.4 HeFDI: Joint Hessian Research Data Infrastructures

The aim of the joint Hessian research data strategy is to establish a shared **research data infrastructure** at Hessian higher education institutions. This is intended to initiate, coordinate and finally establish the necessary organisational and technological processes for anchoring a **research data management (FDM) system** depending on the requirements of the higher education institution in question. Through coordinated and agreed-on procedures, resources are bundled, the available know-how is expanded, work processes are made more efficient, and the range of services is extended.

3.2.5 MobiDig: Mobile, Digital Access to University and Library

The project combines national or state-coordinated preliminary work on university-specific apps, library research systems and sustainable media distribution. It transfers these development strands into the coordinated development of a joint information service for the Hessian higher education institutions. To this end, MobiDig and the **Hessian Library Information System** (HeBIS), which is based at Goethe University, are working on a development partnership on the core components of a technical infrastructure (e.g. integration of campus management), in which their interfaces and site-specific functionalities are adapted to the needs of the higher education institutions.

The project partners are working together under the leadership of HeBIS and Goethe University to establish the necessary processes and technical infrastructures for the mobile presentation of information and content for students. This includes the introduction of various core components such as the integration of the campus management system with the corresponding user interfaces for different end devices and the adaptation of the app to university-specific requirements. In combination with the integration of librarian resources, this project will adapt the higher education institutions' information services to the needs of 'digital native' students.

3.2.6 LaVah: Long-Term Availability at Hessian Higher Education Institutions

How will digital objects be secured, archived and made available over the long term? This question is the focus of 'LaVah.' In this context, **long-term archiving** means controlled archiving and provision.

Supported by the Hessian higher education institutions, an infrastructure for the long-term availability of digital objects is to be gradually established. The solution to be developed should be able to accommodate digital objects of all formats, but is initially limited to formats with low and medium risks of obsolescence. The project under the leadership of HeBIS is aimed at establishing organisational measures for the development of an infrastructure for long-term availability. Curation of data flows, archiving and risk management, storage and access procedures will be implemented by several partners in the form of a pilot project with joint responsibility but clear lines of responsibility.

3.2.7 HeIDI: Hessian Identity Management Infrastructure

Hessen's higher education institutions are currently at very different levels in terms of their **digital identity management.** Some are still in the early stages; others, such as Goethe University, are already taking steps to introduce a uniform user ID together with the library, thus eliminating media discontinuity.

The project aims to enable all institutions of higher education to participate in the services of DFN-AAI (German Research Network Authentication and Authorisation Infrastructure) and thus create the prerequisites for the use of DFN-AAI services. In addition, an infrastructure will be established which enables the joint use of new services for students, teachers and employees across universities without the need for costly adjustments of the individual identity management systems. At the same time, the project, which is managed by Goethe University Frankfurt and the Rhine-Main University of Applied Sciences, will create the possibility of linking the library systems with the individual identity management systems of the universities.

3.2.8 CSC: High-Performance Computing at Goethe University

The GOETHE high-performance computer implements a green IT technology that is constantly being further researched and is at the same time indispensable for implementing research projects in the natural sciences, medicine, life sciences and economics. The development of the **LOEWE-CSC high-performance computer (*Hochleistungsrechner*—HLR)** in 2010 and the start of operations in January 2011 linked two different goals from the very beginning: the computer as a research object and the computer as a research infrastructure for large-scale research projects with a heterogeneous user profile. The Frankfurt high-performance computer LOEWE-CSC is one of the most energy-efficient large-scale computers in Europe. With a computing power of 299 TFlop/s, it was the second fastest supercomputer in Germany at the time of commissioning. The energy efficiency is 740 MFlop/s per watt. This means that LOEWE-CSC consumes only about a quarter of the energy of comparably fast computers, at investment costs of just under five million euros, which is about a third (Center for Scientific Computing n.d.). It was designed by two professors from Goethe University, Volker Lindenstruth and Hans Jürgen Lüdde. The particularly energy-efficient 'green' high-performance computer significantly reduces energy and operating costs through a special design principle. In 2014, the model came first in the world rankings of energy-efficient mainframe computers. In 2019 it was completely renewed and is now even more powerful and energy-efficient.

3.2.9 CMMS

LOEWE's focus on **CMMS (multi-scale modelling in the life sciences)** is intended to provide the methodological basis for the efficient use and translation of highly informative and heterogeneous data collected from novel imaging, microscopy and omics technologies into mechanistic and predictive models to enable a deeper understanding of biological systems. Such models will in turn foster the development of new experimental approaches. This integrative approach thus creates important conditions for innovative applications in many areas of basic biological research, medicine, biotechnology and pharmacology. This example in particular shows how the intelligent use of new technology leads to real expansions of research perspectives.

3.3 Administration and Management

Here, too, selected measures are first presented in an overview (see Table 3) and then explained. Some measures, e.g. GInKo, have already been mentioned because they also support research and/or teaching.

Table 3 Measures to support digitalised university management

Measure	Problem	Solution	Concrete contents	Status
GInKo (Campus Management)	Organisational weaknesses in study management process organisation	Improve processes and make them legally compliant	Complete implementation of HisInOne	Ongoing
HeIDI	Multiple ID management systems	Merging of UB-ID management with university ID management	Create conditions for the use of DFN-AAI services	Aspired
LMS PE/OE	High administrative effort, no digitally supported services, low user-friendliness in human resources development	Consolidation and transparency through the use of LMS	Complete online system compatible with GBS system	Ongoing
IT Security Policy	Data security	Handout (Website)	Detailed description in the form of an IT security directive	Implemented
IT Security Management Team	Data security	Combining expertise and responsibility	Organisation of the SM team	Implemented

3.3.1 GInKo

As already described in Sect. 3.1.6 **GInKo (Goethe-Universität Informations- und Kommunikationssystem—**Goethe University Information and Communication System) is intended, on the one hand, to map the organisation in study management effectively, transparently and in a legally compliant manner and, on the other hand, to support the strategic planning of the university by providing timely and reliable information. As an organisational development project, GInKo is of particular importance for administration and management, as the project aims to achieve fundamental and legally compliant university-wide process and organisational development and harmonisation of the entire study management system.

3.3.2 HeIDI

The state-wide HeIDI project of the Hessian higher education institutions was discussed above in Sect. 3.2.7. From the perspective of administration and management, this project is important because it deals with the development of an identity management infrastructure (IDM) beyond the borders of Goethe University with all state universities and higher education institutions in Hessen. The aim of the Hessen-wide project is that every member of a higher education institution can also register with services that are provided by other higher education institutions throughout Hessen with their own institution account.

3.3.3 Learning Management System (LMS) in Human Resources and Organisational Development (PE/OE)

To implement modern qualification formats for employees and managers, Goethe University relies on a learning management system solution. This enables various formats of lifelong learning, which follow a blended learning approach (analogue and digital), including collaborations, learning videos, e-learning and classroom courses. The **learning management system (LMS)** is adaptable and can react flexibly to changing and future needs. It is characterised by user-friendliness and features that make continuing education attractive, transparent and accessible. Learning content can thus be accessed independently and in a targeted manner.

3.3.4 IT Security Policy and IT Security Management Team (SMT)

Goethe University operates in accordance with an IT security guideline, which increases the security level of data processing systems and improves data protection as well as communication and data security. The **IT Security Management Team (SMT)** is the central decision-making and control body for **IT security** at Goethe

University. It collects and decides on the uniform IT security framework guidelines for the university and prepares an annual IT security report (Goethe University 2018b).

4 Summary and Outlook

The digital transformation process at Goethe University ties up a lot of attention and human and financial resources both in the coordination of the initiated projects, as well as in the implementation of the development measures in organisational and IT areas. The 2020 pandemic initially shifted priorities to secure more digital teaching. Not all the projects that have already been launched have been free of setbacks or surprises, but the Goethe community increasingly accepts—even more because of the coronavirus crisis—that investing in 'digitalisation' will make its own organisation fit for the future and that this transformation process will yield a payoff in the form of improved working and study conditions. However, the outlook also includes the insight that the implementation of the digitalisation strategy, in particular the anticipation of unforeseen (technical) developments, their integration into existing IT and organisational structures, the synchronisation of the services described, and the creation of corresponding interfaces will continue to require a very high level of resource input which will not decrease in the future. Not only do the IT infrastructure and process organisation have to be renewed at regular intervals, the training and continuing education needs of students and university employees in science and administration must also continue to grow significantly.

References

Center for Scientific Computing (n.d.) Forschung in Hessen am Hochleistungsrechner – von der elementaren Materie zum Klimamodell [Research in Hessen on high-performance computers – from elementary matter to climate model]. Goethe University, Frankfurt am Main, pp 2–7. https://csc.uni-frankfurt.de/wiki/lib/exe/fetch.php?media=public:loewe-csc2012final.pdf

Digital gestütztes Lehren und Lernen in Hessen (2018) [Digitally Supported Teaching and Learning in Hessen] – Antragstellende Hochschule: Philipps-Universität Marburg für die hessischen Hochschulen. https://www.hessen.de/pressearchiv/pressemitteilung/land-hessen-foerdert-digitales-lehren-und-lernen-mit-rund-zehn-millionen-euro-0

Dräger J, Müller-Eiselt R (2015) Die digitale Bildungsrevolution – Der radikale Wandel des Lernens und wie wir ihn gestalten können [The digital education revolution – the radical change in learning and how we can shape it], 2nd edn. Deutsche Verlags-Anstalt, Munich

Ehlers UD (2018) Digital leadership in Hochschulen. Synergy 6:24–27

Frankfurter Neue Presse (2018) Neuer "grüner" Hochleistungsrechner für Goethe-Universität [New "green" high-performance computer for Goethe University]. https://www.fnp.de/frankfurt/neuer-gruener-hochleistungsrechner-goethe-universitaet-10365874.html

German Research Foundation (2019) National Research Data Infrastructure. https://www.dfg.de/en/research_funding/programmes/nfdi/index.html

Goethe University (2014a) Mission statement of Goethe University. https://www.uni-frankfurt.de/67690212/SO_Leitbild_der_Goethe.pdf

Goethe University (2014b) Grundsätze zu Lehre und Studium an der Goethe-Universität [Principles of Teaching and_Studying at GoetheUniversity]. https://www.uni-frankfurt.de/51044043/Grundsätze_zur_Lehre_und_Studium_an_der_Goethe_Universität.pdf

Goethe University (2018a) Leitbild digitale Lehre an der Goethe-Universität [Mission statement digital teaching at Goethe University]. https://www.uni-frankfurt.de/72312239/RZ_Leitbild_dig_Lehre_A4_low.pdf

Goethe University (2018b) Die IT-Sicherheitsrichtlinie der Goethe-Universität Frankfurt [The IT Security Policy of Goethe University Frankfurt]. http://www.uni-frankfurt.de/67025588/Informationsveranstaltung_zur_ITSR_Web.pdf

Hessian Ministry of Higher Education, Research and the Arts (2018) Chancen der Digitalisierung für die weitere Verbesserung der Lehre an den Hochschulen Hessens. Bestandserhebung und Ist-Analyse. [Opportunities of digitalisation for the further improvement of teaching at the universities of Hessen. Survey and analysis of the current situation]. Institut für Hochschulentwicklung (HIS), Wiesbaden

Hessian State Parliament (2015) Antrag der Fraktionen der CDU und BÜNDNIS 90/DIE GRÜNEN betreffend Chancen der Digitalisierung für die weitere Verbesserung der Hochschullehre nutzen [Motion by the CDU and BÜNDNIS 90/DIE GRÜNEN parliamentary groups concerning the opportunities of digitalisation for the further improvement of higher education teaching]. http://starweb.hessen.de/cache/DRS/19/6/01796.pdf

Hochschulforum Digitalisierung (2016) The Digital Turn – Hochschulbildung im digitalen Zeitalter. Arbeitspapier Nr. 27. [The Digital Turn – Higher Education in the Digital Age. Working Paper No. 27]. Hochschulforum Digitalisierung, Berlin

Horz H, Schulze-Vorberg L (2018) Digitalisierung in der Hochschule. [Digitalisation in higher education]. In: Arnold N, Köhler T (eds) Digitale Gesellschaft – Gestaltungsräume. Konrad Adenauer Foundation, Berlin, pp 57–71

Kultusministerkonferenz (KMK) (2016) Bildung in der digitalen Welt – Strategie der Kultusministerkonferenz [Education in the digital world – Strategy of the Conference of Education Ministers]. KMK, Berlin

Koalitionsvertrag (Coalition agreement) (2018) Koalitionsvertrag zwischen CDU Hessen und Bündnis90/DIE GRÜNEN Hessen für die 20. Legislaturperiode. Aufbruch im Wandel durch Haltung, Orientierung und Zusammenarbeit [Coalition agreement between CDU Hessen and Bündnis90/DIE GRÜNEN Hessen for the 20th legislative period. Change through attitude, orientation and cooperation]. https://www.gruene-hessen.de/partei/files/2018/12/Koalitionsvertrag-CDU-GRÜNE-2018-Stand-20-12-2018-online.pdf

Opitz LK, Lommel M (2017) Zweite Universitätsweite Studierendenbefragung der Goethe-Universität – zentrale Ergebnisse [Second university-wide student survey of Goethe University – central results]. Department of Lehre und Qualitätssicherung (Teaching and Quality Assurance) at Goethe University, Frankfurt am Main

Picot A, Reichwald R (1994) Auflösung der Unternehmung? Vom Einfluß der IuK-Technik auf Organisationsstrukturen und Kooperationsformen [Dissolution of the company? The influence of information and communication technology on organisational structures and forms of cooperation], ZfB 64th edn., pp 547–570

Picot A, Ripperger T, Wolff B (1996) The fading boundaries of the firm: the role of information and communication technology. JITE 152:64–79

Senge PM (2006) The fifth discipline. The art and practice of the learning organization. Doubleday, New York

Van Dick R, Helfritz KH, Stickling E, Gross M, Holz F (2016) Digital Leadership – Die Zukunft der Führung in Unternehmen [Digital Leadership – The future of leadership in companies]. https://www.dgfp.de/fileadmin/user_upload/DGFP_e.V/Medien/Publikationen/2012-2016/Digital_Leadership_Studie.pdf

To See Humans in Vocational Education Differently: Vocational Pedagogy and Reflections on Discourse, Subject, and Education in Industry 4.0

Sabine Hering, Jacqueline Jaekel, and Tim Unger

Contents

1 Introduction

"Will robots take my job?"[1] This question represents a tangible and existential concern for many employees. In addition, it seems to express the insecurity of societal institutions in handling the challenges of Industry 4.0.[2] As an example new, or "smart" technologies circumventing human tasks in the future and making

[1] This question comes from a website, which informs on the substitution potential or risk of over 700 different jobs in reference to the results of a research study of Frey and Osborne (2013), and statistical surveys of the "Bureau of Labor Statistics" (online, URL: https://willrobotstakemyjob.com).

[2] The Economy and Science research group's communication promotion team defines the future project "Industry 4.0" as follows: "On the path to the **Internet of Things**, the production site Germany shall be guided to the new age via the merging of the virtual and the physical world into cyber-physical systems as well as the resulting potential convergence of technical with business processes" (2012, 8. Emphasis in the original).

S. Hering (✉) · J. Jaekel · T. Unger
Human Resource Education, RWTH Aachen University, Aachen, Germany
e-mail: sabine.hering@rwth-aachen.de; jacqueline.jaekel@rwth-aachen.de; tim.unger@rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_56

their jobs superfluous is a message often transmitted in **vocational pedagogy**. At the same time, they simplify and optimize work and business processes, and thereby create new **work and employment potential**. This is not a particularly surprising realization. Technologies have always changed professions and deeply interfered in people's way of life as well as in the structures of vocational education and training. The so-called fourth industrial revolution has been associated with radical implications for technical, social, and cultural dimensions of society. At times, a transfer of the number "four-point-o" to other areas of societal life such as work, economy, culture, education, or workforce 4.0 are inevitable.

However, the perspective of the above question is striking as it suggests an alleged **subordination of the human factor, or rather of the human workforce vis-à-vis the technology factor**. The question expresses an expectation, which addresses the aptitude of humans for Industry 4.0 and not, in turn, the configuration of the industry for humans. Of interest is this matter of course in the argument for a **"monodirectional relation"** between **humans and technology**: institutions of vocational education should aim at enabling qualified handling of new technologies. Vocational education then refers to:

> preparing trainees for future action requirements. Playing the ball where the player [the trainee, S.H.] will be aims at advocating expertise, which will be required in digitized business processes and job types in the future. (Gerholz and Dormann 2017, p. 18)

This limiting view of educational didactics puts the "digital" needs of worker and company organizations inevitably in the focus of vocational education, while the trainees' needs and any theoretical-systemic legitimation of vocational education remain, if at all, secondary. Put in a different way: vocational education could just as well be designed differently. Industry 4.0 could, for example, be conceptualized as the right to personal development and advancement in the sense of extending human freedoms in and through a profession. Accordingly, **vocational education** is connected to ethical dimensions and not merely with considerations of usability for technological developments (see Kutscha 2008; Unger 2014). This educational aspiration points to a constitutive problem of **vocational pedagogy**, which as a subfield of educational science has always existed in the tension between reference to the subject (and the idea of maturity), and reference to the system (and connected demands of a functionally sophisticated environment) (see Kutscha 2008, p. 4).

This article aims at widening the space for debate of the discipline of **vocational and economics pedagogy** and the current processes of transformation. The main objective is to show that scholars of vocational training should not only ask how **gainful employment and requirements of expertise** change in the context of Industry 4.0 and the learning for which people are being prepared, but rather how they can participate in shaping societal changes regarding didactics and subject-orientation within vocational training. The question of how vocational education as a discipline can critically influence the results of a "disruptive"[3] and with it the constructive working part of society in the core aspects of education and paid work, is therefore essential for the authors.

The article will stimulate reflections on two levels, on the **discursive and on the subject level**:

1. Which discursive practices in the context of Industry 4.0 currently determine the special discourse in vocational education? Which educational aims are being formulated and how are participants being addressed?
2. How can the subject relate to these expectations and requirements? Which resources do they have at their disposal when dealing with these challenges? Are there moments of liberty in dealing with Industry 4.0?

The foci are the interplay of the definitions of discourse, subject, and practice in vocational education. The authors develop discourse analysis and epistemic biographical tools to highlight the conditions, processes and results of education and employment that are of public urgency not only in the context of Industry 4.0 but also in the context of continuous **societal processes of change**.

2 Reflections on the Discourse Level: Questions of Discourse Analysis

Those familiar with the current discourses on vocational education on topics such as industry, vocational education, and Economy 4.0 will not be surprised by anything new: the main topic "digitization" has been dominating **debates on vocational pedagogy and politics of education** for years. While on one hand questions of education didactics, in particular the use of virtual learning tools and digital media in the learning environment of schools and companies dominate, there are regulative questions in focus dealing with researching practical and acting requirements and the connected adaptation and development of **qualification, competence, and professional profiles** on the other. They all share the seemingly natural bond with the institutional guiding logic of the **educational and employment system**. From where does this dominance come? Similar to other academic disciplines, vocational pedagogy derives orientation from reference fields and guiding constructs, which provide it with "disciplinary coherence" to the inside and outside (Arnold et al. 2016, p. 195).

[3]The term disruption is defined here in reference to Schumpeter (1943) as a profound "destructive" process of change, which regards various areas of societal life. Schumpeter has called this process "creative destruction" (see. ibid). "Schumpeter considers the process of creative destruction, where all goods and production procedures are continuously replaced by new ones, as a motor for economic development. The creative and inventive entrepreneur plays a central part as he continuously advances economic and technical progress with new ideas and the use of new production methods, techniques, and ways of manufacturing" (Bundeszentrale für politische Bildung 2016).

Niklas Luhmann underlines the central role of guiding terms for the self-conception of a field:

> Science uses terms as a sensor to adapt the system controlled by theory to reality; they serve to transpose vague complexity into defined complexity, which science can utilize. (Luhmann 1991, p. 13)

As a research tool, **discourse analysis** provides a useful approach for displaying such specific meanings and orientations of a discipline. According to Michel Foucault's studies on discourse theory and power dynamics, one should "treat discourses (. . .) as practices that systematically form the objects that they are talking of (1988, p. 74)." In their significance, discourses partake in determining what society recognizes as "true" and "valid," for example, in the various areas of practice of vocational education. In this perspective, truth refers to "valid and validity-claiming knowledge of the world," which agents use for orientation and with whose help social reality is constructed. In contrast to **biography analysis**, which mainly deals with the analysis of biographical self-articulations alongside biographical narrations (therefore stories), discourse analysis is concerned with reconstructing "discursively produced and transmitted" trans-individual representations on the micro-, meso-, and macro-levels (Bührmann and Schneider 2008, p. 101). It therefore decisively depends on the institutionalized and trans-individual practices, which knowledge dominates in the long-term. The object of discourse analysis consists in the analysis of its impact, that is to say its productive and reality-constructing influence, for example on the specific nature of guiding terms and related strategic concepts.

The quotidian behavior of agents of vocational education is likewise oriented alongside such practices. Which role the **human factor** plays in discourse has taken on starkly different interpretations in the tradition of discourse theory. In the research program of the **sociology of knowledge approach to discourse**,[4] Keller and Bosančić (2018) distinguish between various roles of social agents in **discursive processes**. Referring to the state of being human, they distinguish between **subject positions** ("as the implicit subjects constituted in discourse") and the actual **manners of subjectification** ("that is to say the way in which actually living, embodied and acting humans react, that subject positions address, react to these") (Keller and Bosančić 2018, p. 896). Therefore, discourses do not only generate model practices as particularly desired or undesired behavioral patterns, but rather also model types and model subjects. The latter are brought into the discourse as images of positive and/or negative existence of humans and have a conventionalizing effect on the acting and speaking practice of the agents (see ibid). Particularly in the transition from the twentieth to the twenty-first century, new forms of subjectification have formed among others through the erosion of societal core institutions such as work, family, and welfare state, and through the individualization of lifestyles and

[4]The sociology of knowledge approach to discourse (SKAD) according to Keller and Bosančić follows up on the theory of symbolic interactionism, the theory of sociology of knowledge and Michel Foucault's concept of discourse.

circumstances (see Beck 1986; Bührmann 2012). In the German-speaking literature, the social type of the worker-entrepreneur according to Voß and Pongratz (1998) and the entrepreneurial self (Bröckling 2007) are among the most widely known forms of subjectification. The message is: everyone is their own entrepreneur. Expanded freedom of subjective action and decision-making in the context of societal change open up not only new job chances, they rather also imply a pressure for action imposed by society to use the acquired freedom effectively and self-determined for the development of their own work capacity.[5]

The following section presents the model practices and subject positions around the topic of vocational education and Industry 4.0 that are processed within and with the help of vocational education discourse. The following questions are central in the context of the selected discourse analytical procedure, following the presented analytical concepts of discourse, power, and subject: which topics determine the current theoretical vocational education discourse on Industry 4.0? Which objectives have been formulated and how do they address the participating educational subjects? The question of how the subjects react, that is to say, which modes of subjectivation they develop would be the object of an empirical and biography-oriented empirical approach (see Sect. 3). The following **regulatory structures in special discourse**[6] of vocational training can be summarized regarding the focus topic "Industry 4.0." They appear dominant and structuring on the operating discourses.

1. ***Problem Detection***: Digitalization is the central key term in the dealing with the issues of the future of Industry 4.0. Macro- and meso-analytical considerations predominate regarding the digital transformation of the labor force, which primarily deal with the consequences on the areas of education, the economy, and the business world, as well as with the structural causes for existing job fit problems between the market for training and the job market.
2. ***Needs (for Action)***: Needs are deducted from various levels of societal reference systems and integrated into the discourse. Needs are addressed on the level of economic reference systems (the need for skilled workers, technical supply, qualifications, and competence), of educational reference systems (the need for integration, education and professionalization) and political reference systems (the need for reform and dialogue with the frame of current structures).
3. ***Guiding Objectives***: The communicated needs serve as a frame of orientation for the deduction and fixing of curricular goals, whose implementation should be the

[5] Note here that the mentioned subject positions of the work capacity-entrepreneur and the entrepreneurial self, characterize an ideal type that is not equal to the actual modes of subjectivation in reality.

[6] The analysis includes journal articles and contributions of academic discourse arenas of vocational education published between 2015 and 2018. Included in the detailed analysis were contributions from the following journals: Zeitschrift für Berufs- und Wirtschaftspädagogik, berufsbildung – Zeitschrift für Theorie-Praxis-Dialog, Berufsbildung in Wissenschaft und Praxis, bwp@ Berufs- und Wirtschaftspädagogik online. The selected bibliography was also the subject of the analysis "Industrie 4.0 – Wirtschaft 4.0 – Berufsbildung 4.0" (see Langenkamp and Linten 2018).

objective of vocational education (more narrowly of the politics of vocational education). In this context, the concept of competence and ability marks a central category of education theory with includes numerous subcategories. Such, in addition to the development of job-related professional, social, and decision-making competence, of late skills include the promotion of **digital, cross-section, programming, process, interface, software, Internet and technical competence**. The orientation of the **concept of capability** turns out to be similarly heterogenous as it links employability to the **individual adaptability, transformation ability, teachability, analytical skills, reflection capacity and organizational talent**, as well as the innovation capacity, sustainability, and with it **competitiveness** of the involved institutions such as **vocational schools and companies**. The objective of vocational education is therefore the development of a broad competence profile and its immediate societal usability. The question remains how such an oriented understanding of education, which considers education mainly as a "jumping board for a successful entry in the digitalization of the German economy" (Kunz 2015, p. 35), can live up to the expectation of maturity as a genuine educational aim of vocational education? A logic of adaptation and improvement are inherent to the educational aims, which discursively convey the scope of training, modes of working and job profiles, as well as the fit of the subject in training to the digital age.

4. ***Modeling Approaches and Practices***: For the implementation of the identified guiding objectives, primarily demand- and systems-oriented development strategies are favored. To manage **digital transformations**, the adjustment of competence and occupational profiles as well as professional spheres of activity (with integrated work and business processes) is oftentimes advocated. The focus is on securing the status quo, that is to say the future potential of the vocational training system. On the micro- and meso-levels measures are being addressed which aim at ensuring the attractiveness and continued existence of vocational training. A didactic convergence of the duality of vocational training (marked by **vocational schools and companies as learning environments** and their regulations). Consequently, model practices such as lifelong learning, digital learning, and "smart education" have been embedded into the institutional framework of vocational education, manufacturing and artisanry.
5. ***Subject Positions and Model Subjects***: Specialized discourse primarily takes the role of skilled workers and employees, of the companies and schools, and to a certain degree that of trainees and learners into consideration.[7] Dominant model practices on the institutional level are hereby transposed to the learning subject. An **education** type is being created, which is eager to adapt and technophile. The subject is being addressed along this attribution accordingly: "Create a fit! Develop new skills! Be digital, smart, and innovative!" The expectation is expressed that education participants' successful dealing with digital change is correlated to their networking and corresponding to the expectations of a digitized

[7] Students, adolescents, teachers and trainers play a secondary part in the discourse.

society. An orientation on the "outside" (the system environment) is maintained discursively. A turn towards the inside towards the interests, desires, and individual needs of the subjects remains secondary.

In conclusion: a guiding logic dominates on the level of educational, societal and economic systems, whose institutional reference categories and regulations are continuously brought into the discourse (see Hering 2017, p. 385). This entails also the dominance of macro- and meso-analytical examinations, which primarily refer to the structural changes of the system environments such as the development of professions, the job market and employment. When the significance of education and employment changes, as for example in the context of an increasing trend towards digitization and the **technologization**, it entails also an "upgrading" of education and employment (Hirsch-Kreinsen 2016). The discipline attempts to absorb the societal integration and employment risks by intensely turning towards the term of competence. It developed answers that are—according to the expectations of the field—targeted towards the securing of the structures of the educational system as well as its function of employment, allocation, and qualification (see Arnold et al. 1998, Sect. 3). This framing reveals an insecure and at times very rigid self-understanding of vocational education, which reacts to societal processes of change in a more conformist way than providing new directions. The discursive practices share reference to constructs that are embedded in the established frame of theory formation of vocational pedagogy, such as profession, work, technology, training, and qualification development (see Arnold et al. 2016). What they lack are approaches of distancing and reflection of their own disciplinary positions to address new societal challenges constructively and question existing figures of order accordingly.

Simultaneously, a certain counter-discourse has been forming vis-à-vis the analyzed special discourse (see above) in reference to it, which disturbs the aforementioned orientations on the system-level and partially reverses them. Not technology, but humans and machines become the origin of questions based in educational theory and practice. This becomes apparent, among other things, in contributions that primarily address the significance of humans in the interaction of humans and machines (see ZHAW/IAP 2017, p. 3) or that deal with approaches and critiques of good (that is to say, more humane) job designs (see Schröder and Urban 2016; Schröder 2014). Contributions that call for a "more humane and job adequate" vivid work (Brödner 2015, p. 19), where the "employee who is working at a machine is not hastily [degraded] as an automatable residual" (Pfeiffer and Suphan 2015, p. 24) correspond to this approach. Instead of the concept of competence Pfeiffer and Suphan put "sensual qualities" of day-to-day work such as intuition, gut feeling, and emotions, as well as connected dynamic practical knowledge for the dealing with complexity in the focus of their considerations in their contribution (ibid, p. 22). Even if this perspective gives the impression that not the human but rather their experience as resource for the handling of paid work is put to the foreground, it also accounts for the personal attitudes, stances and values as well as the potential that employees contributes to their work and thereby help shaping individually.

The following section deals with thinking from a different perspective the interplay in the development of subject and structure in changing business worlds. It will be shown that the discipline of vocational and business education can also address the human in a different way, by examining the degree of freedom in the dealing with individual biographies.

3 Reflections on the Subject Level: Questions of Biography-Theory

How can the subject be supported in dealing with the challenges of a disruptive labor force? The president of the Federal Office for Vocational Education (BIBB) Friedrich Hubert Esser, considers training subjects as foundational: "professional empowerment secures employability in the long-term. It is important to enable young people to master complex future tasks on the foundation of broad qualifications and competences." (Brandaktuell 2016, p. 1). On the one hand this view is future-oriented and aims to provide the individual with tools for mastering future challenges; however, this competence-oriented view risks to enter a "compulsory-continuing-professional-development-loop." Whenever professional skills can be carried out by computers or machines, the subject will be asked to take over different tasks: job requirements will continuously change—the subject will have to continuously advance and adapt. In this scenario, classical competences can no longer provide orientation and stability as they are located beyond the subject and on their turn too much influenced by increasing proliferation of technology and demand. It is likewise problematic when in a fast-track and **transformative society** particular expert knowledge is no longer relevant and therefore no longer in demand. Therefore, it is central that subjects dispose of other, identity-providing resources, which allow them to orient themselves and direct future actions. Their own biography can be such a resource. Biographization, dealing with one's own biography, has been described as the subject's achievement to organize lived life through meaning. Winfried Marotzkis' approach to ascribe biography substantial meaning as the origin of educational processes is relevant for the theory of vocational education:

> In order to claim that biographical development actually represents a learning process, it is not sufficient to demonstrate that the subject in question negates content of the previous development level (the semantic level of analysis), rather it must be possible to demonstrate that also the foundational structural principle of the previous level has been negated. One must demonstrate that the foundational principles of the production of meaning, that therefore the mode of experience processing and with it the foundational view of the world and of the self have changed. (Marotzki 1990, p. 219)

"Previous development levels" refer here to life segments that are connected to specific experiences that are to be reflected. In the sense of an educational process, specific experiences or situations are being rejected ("negated") and that not just those as such but also the previous subjective stance toward them. This "mode of experience processing" comprises the corresponding particular relations to world and self, that is to say the manner in which the subject understands to construe themselves and the world: "on the one hand regarding the relations that [people] develop with themselves (self-reference) and on the other hand regarding the relations that they develop with the world (world reference)" (ibid., p. 61). In other words, education has essentially to do with those chances where people learn to see themselves and the world differently. Such fundamental learning processes refer to the actually lived life and not on institutionally ordered processes of life. Education biographies are not to be equated with education careers (see Kade 2005).

The question arises to what degree education biographies as discussed above actually change in the context of the technologization and digitalization. Will people adapt their biographical attribution of meaning? Will we be even further dependent in the future on educational processes of occupational biographies than we currently are? Or will occupational biographies possibly be not so relevant anymore to make us grasp ourselves as individuals and to let us experience affiliation with our social world? With a glance at educational careers, therefore the institutionalized processes of qualification and vocational education, this would constitute a problem. Such a version would be the "muddling through" as situative reactivity in fast-track societies, according to Hartmut Rosa (see Rosa 2005). He also assumes that technological changes largely affect the job market, just like professions, professional structures and expectations profoundly change as a consequence, which leads to the necessity for the subject to adapt adequately. A situative reactivity would, however, demand such an elevated flexibility of the subject, that long-term biographical orientations would no longer be an effective means to deal with orientation crises. What is the state of the general goals of vocational training that orient and legitimize professional pedagogical action: **critical ability**, **self-determination aptitudes**, **democratic participation**, etc., are learning goals biographically invested in the long-term. What if the "muddling through" turns out to be a probate means of biographical design for people because they perceive it as more efficient, simple, smooth and legitimate? More than now, scholars of vocational training will need to deal with their own expectations for career processing in the future. They will need to consider the question who those people actually are, whom they are addressing and who they want them to be.

That much on the topic of educational careers. Contrary to educational careers, educational biographies do not mainly consist in institutional expectations providing direction for people. Contrarily, the educational biographical dimension of education is becoming increasingly important in times of indeterminacy of societal transformation processes and experiences of crises connected to them. In the approach of **transformative education**, Hand-Christoph Koller highlights the constitutive moment of irritations and experiences of crises for educational processes: "educational processes are triggered by the confrontation with problems previous world- and self-conceptions are no longer sufficient to master" (Koller 2010, p. 289). This leads to a disproportion of subjective **perception, thinking, and action schemes**, which were adequate in previous life stages but no longer endure under current challenges (see ibid.). The subject must find new ways for dealing with problematic situations. At this point the realization is decisive that solutions for such biographical crises are no longer institutionally secured. In professional initial education, the trainee can neither be promised biographically permanent continuance in the training profession nor secure knowledge on how to glue biographical breaches in times of life crises.

Regarding the attributions to people in the sense of “Create a fit! Develop new skills! Be digital, smart, and innovative!”, as addressed in Sect. 2, a transformatory theory of vocational education would search for liberty potentials in the way of dealing with people with such attributions. Educational processes can particularly consist in apprentices refusing to expose themselves to such expectations. While the educational career sets the socially expected continuation in the sense of “Develop new skills!” on the one side, education can lead to launching a very different direction of experience processing. Exemplary are the following motives of career changers in the **teaching profession**: “I want to quit the economy,” or “I want to work with people in a different way than I experienced it in the office” or “I don’t want to spend the whole day in front of a computer screen.” Whether such motives are right or wrong is not relevant here. From a vocational education perspective, the interesting aspect of such motives is that they are, in a very specific regard, cross with the institutional guiding logic for life in a digital age addressed in Sect. 2, that predominantly serve as orientation for vocational pedagogy. What if a career-changer carries out the break with the educational career, planned it in a highly reflected and educated manner, and arrived at the conclusion of wanting to lead a life that is not characterized by digitization and Industry 4.0? Presumably not a problem, it is just an isolated case. But what if ever more of the currently educated working population would behave in a similar manner? At what point would scholars of vocational education have to ask themselves what they can do to ensure that young people can actually carry out an educational career in the sense of the education profession?

A profound problem for vocational pedagogy is hiding behind such inconspicuous motives, when a break between educational career and educational biography occurs. On the one hand, from a theoretical vocational education perspective, the discipline can well legitimize that qualification and competence development enable participation in current employment. On the other hand, however, decisions of individuals to turn away from their employment are justified, since those were a product of educational processes.

How can vocational pedagogy confront this issue? How can it deal with its self-understanding reflexively?

4 Propositions for Designing Communication and Research Practice of Vocational and Economic Pedagogy

To conclude, the authors would like to provide concrete propositions for designing the disciplinary communication and research practice of vocational and economic pedagogy as a discipline.

To break open the presented discursive frame of orientation of vocational pedagogy and to allow for different perspectives, a critical examination of past disciplinary exchange and conference formats is necessary. The section conference of vocational and economic pedagogy and the university open days *Berufliche Bildung* (vocational education) have become established. In particular, the university open days have great resonance within the discipline due to their broad orientation on the actors from science, politics and practice. In addition to the respective university-internal organizational team, the study group *Berufliche Bildung e.V.* (Vocational Education e.V.) is significantly involved in organizing the conference, which guarantees a homogeneous and systematic exchange among all participants. Beyond that, however, approaches are necessary on the level of critically distanced academic communication, which provide space for systemic non-conformity and re-thinking, and which are adequate for temporarily muting disciplinary discursive mechanisms. Exchange formats are important that "disturb" the discourse and keep it open, and thereby provide the possibility of developing creative and new ideas. The format of BarCamps is such an example. "A BarCamp is a conference format where the organizer only defines the frame: time, location, and theme. The concrete content is then defined by the participants themselves at the beginning of the event" (Muuß-Merholz 2011). The proposed themes are treated in joint workshops as well as solutions developed and discussed. In particular, due to the open and interactive character regarding content and sequence, BarCamps are designated "un-conferences" or "non-conferences" (see ibid.). Their advantage is their consequent orientation along the participants' interests, questions, and needs. The practical realization, however, still necessitates some experience in the (conference) organization and corresponding general set-up, which could correspond to the freedoms and uncertainties of this still unusual conference format. A similar freedom is provided in *Future workshops* and scenario planning, which despite requiring an experienced moderator, provide significantly more space for opinion and difference formation compared to other formats without being removed from reality or uprooted. Moreover, events and collaborations that facilitate conversations between societal agents and experts from diverse practical areas are important. The education synod *BildungsKonzil Heldenberg* serves as such an example. The *BildungsKonzil* aims to be an "forger of innovative ideas and concepts" where representatives from education, economy, and civil society deal with current education issues to develop future-oriented concepts (BildungsKonzil 2019). In the context of a two-day exchange questions, ideas, and opinions are discussed in various "think rooms," which are presented to an interested and broad audience and up for discussion on the third day (see ibid.). The open format and the facilitation of shared leisure are the central aspects of this conference concept.

To provide a brief summary: innovative ideas require innovative formats of exchange. The discipline should be braver and more open in this regard! Moreover, a “steady education-theoretical identity” is decisive for the academic self-understanding of the discipline if it wants to be understood not only as “vicarious agent” of economy and technology but as a subdiscipline of science of education and also be recognized as such in discourse (Lisop 2009, p. 12). In that case, perspectives in vocational education are indispensable that address the subject’s empowerment in an education-theory regard as well as concerning a “wanting to be” and not exclusively a “should be” (see Bührmann 2012, p. 151). These subjective and collective want-to-be-requirements should also be part of empirical research when analyzing subject conditions in the contest of vocational social worlds. Consequently, of importance would be field studies on biography and interactions that deal with dynamic fitting of subject and structure and that inquire what determines action experience of skilled workers and managers in (super) intelligent work settings, which action problems occur in daily routine and which **coping strategies** actors develop when dealing with them. Subsequently, vocational education research oriented towards discourse analysis could gear towards the connections between micro-, meso-, and macro-levels in the context of a transformative process of change. Research questions would then refer to the examination of subjectivation processes in the modern work force, to the analysis of the relation between institutionally processed subject positions (of the type of modern gainful employment) and actual ways of subjectivation. “To retrace such processes means to problematize the automation of the being-turned-into-a-subject and the making-oneself-a-subject” (Alkemeyer et al. 2018, p. 29).

Contrarily to the initially posed closed question “Will robots take my job?”, open design questions that proactively face current and future processes of change in our society will be critical for the discipline.

References

Alkemeyer T, Bröckling U, Peter T (2018) Jenseits der Person. Zur Subjektivierung von Kollektiven. transcript, Bielefeld

Arnold R, Lipsmeier A, Ott B (1998) Berufspädagogik kompakt. Prüfungsvorbereitung auf den Punkt gebracht. Cornelsen, Berlin

Arnold R, Gonon P, Müller H-J (2016) Einführung in die Berufspädagogik, 2. Aufl. Verlag Barbara Budrich, Opladen

Beck U (1986) Risikogesellschaft. Auf dem Weg in eine andere Moderne. Suhrkamp Verlag, Frankfurt am Main

Bröckling U (2007) Das unternehmerische Selbst. Soziologie einer Subjektivierungsform. Suhrkamp, Frankfurt am Main

Brödner P (2015) Industrie 4.0 und Qualifikation: Vorschau durch Rückschau. In: Berufsbildung in Wissenschaft und Praxis. BWP 6:7–20

Bührmann A (2012) Das unternehmerische Selbst: Subjektivierungsform oder Subjektivierungsweise? In: Keller R, Schneider W, Viehöver W (Hrsg) Diskurs – Macht – Subjekt. Theorie und Empirie von Subjektivierung in der Diskursforschung. Springer VS, Wiesbaden, S 145–164

Bührmann A, Schneider W (2008) Vom Diskurs zum Dispositiv. Eine Einführung in die Dispositivanalyse. Transcript, Bielefeld

Foucault M (1988) Archäologie des Wissens. Suhrkamp, Frankfurt am Main

Kade J (2005) Erziehungswissenschaftliche Bildungsforschung im Spannungsfeld von Biographie, Karriere und Lebenslauf. In: bildungsforschung, Jahrgang 2, Ausgabe 2. 10 S

Keller R, Bosančić S (2018) Wissenssoziologische Diskursanalyse. Die Analyse sozialen Wandels am Beispiel öffentlicher Debatten über Abfall in Deutschland und Frankreich. In: Akremi L et al (Hrsg) Handbuch Interpretativ forschen. Weinheim, Beltz, S 886–916

Koller H-C (2010) Grundzüge einer Theorie transformatorischer Bildungsprozesse. In: Liesner A, Lohmann I (Hrsg) Gesellschaftliche Bedingungen von Bildung und Erziehung. Eine Einführung. Kohlhammer, Stuttgart, S 288–300

Kunz C (2015) Next generation competencies for a digital world – Erfahrungen aus dem SiemensProjekt "Industrie 4.0@SPE". In: Berufsbildung in Wissenschaft und Praxis. BWP 6: 33–35

Luhmann N (1991) Soziale Systeme. Grundriss einer allgemeinen Theorie. Suhrkamp, Frankfurt am Main

Marotzki W (1990) Entwurf einer strukturalen Bildungstheorie. Biographietheoretische Auslegung von Bildungsprozessen in hochkomplexen Gesellschaften. Deutscher Studienverlag, Weinheim

Pfeiffer S, Suphan A (2015) Erfahrung oder Routine? Ein anderer Blick auf das Verhältnis von Industrie 4.0 und Beschäftigung. In: Berufsbildung in Wissenschaft und Praxis. BWP 6:21–25

Rosa H (2005) Beschleunigung. Die Veränderung der Zeitstrukturen in der Moderne. Suhrkamp, Frankfurt am Main

Schröder L, Urban H-J (2016) Digitale Arbeitswelt – Trends und Anforderungen. Bund, Frankfurt am Main

Schumpeter JA (1943) Capitalism, socialism and democracy. George Allen & Unwin, London

Unger T (2014) Bildung und Erwerbsorientierung. Streifzüge und Wegmarken einer Theorie der strukturalen Erwerbsbildung. Auslesen, Kreuzau

Voß G, Pongratz HJ (1998) Der Arbeitskraftunternehmer. Eine neue Grundform der Ware Arbeitskraft? Kölner Zeitschrift für Soziologie und Sozialpsychologie 50(1):131–158

Online References

BildungsKonzil Heldenberg (2019). https://www.bildungskonzil-heldenberg.at/. Zugegriffen am 15.01.2019

Brandaktuell (2016) Wirtschaft 4.0 – Digitalisierung der Arbeit Herausforderung und Chance für eine zukunftsfähige Berufsausbildung. Newsletter04/2016. Europäische Union. http://www.lasa-brandenburg.de/brandaktuell/fileadmin/user_upload/MAIN-dateien/newsletter_2016/nl4-bibb_01.pdf. Zugegriffen am 22.01.2019

Bundeszentrale für politische Bildung (2016) Schöpferische Zerstörung. Duden Wirtschaft von A bis Z: Grundlagenwissen für Schule und Studium, Beruf und Alltag. 6. Aufl. Mannheim: Bibliographisches Institut. http://www.bpb.de/wissen/A83MMKDen. Zugegriffen am 20.01.2019

Frey CB, Osborne MA (2013) The future of employment: how susceptible are jobs to computerisation? https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf. Zugegriffen am 30.01.2019

Gerholz K-H, Dormann M (2017) Ausbildung 4.0: Didaktische Gestaltung der betrieblichberuflichen Ausbildung in Zeiten der digitalen Transformation. In: Berufs- und Wirtschaftspädagogik online, Nr. 32/Juni 2017. http://www.bwpat.de/ausgabe32/gerholz_dormann_bwpat32.pdf. Zugegriffen am 15.10.2017

Hering S (2017) Der Übergangsbegriff in der beruflichen Bildung – Eine Dispositivanalyse zur berufspädagogischen Formierung von Übergängen im Zentrum disziplinärer Wissensordnungen. https://publications.rwth-aachen.de/record/711485/files/711485.pdf. Zugegriffen am 23.12.2017

Hirsch-Kreinsen H (2016) Zum Verhältnis von Arbeit und Technik bei Industrie 4.0. Bundeszentrale für politische Bildung. http://www.bpb.de/apuz/225688/arbeit-und-technik-bei-industrie4-0?p=all. Zugegriffen am 29.4.2016

Kutscha G (2008) Bildung und Beruf. Leitfaden zum Modulthemenschwerpunkt "Grundlagen und Hauptströmungen der Berufspädagogik" (Teil 1). https://www.uni-due.de/imperia/md/content/berufspaedagogik/kutscha_leitfaden_bildung_und_beruf_17-04-2008.pdf. Zugegriffen am 10.01.2019

Langenkamp K, Linten M (2018) Industrie 4.0 – Wirtschaft 4.0 – Berufsbildung 4.0. Zusammenstellung aus: Literaturdatenbank Berufliche Bildung. https://www.bibb.de/dokumente/pdf/a1bud_auswahlbibliografie-industrie40.pdf. Zugegriffen am 01.11.2018

Lisop I (2009) Identität und Krisenanfälligkeit der Berufs- und Wirtschaftspädagogik im Spiegel der Kategorien Kompetenz und Employability. bwp@ Ausgabe Nr. 16. http://www.bwpat.de/content/uploads/media/lisop_bwpat16.pdf. Zugegriffen am 10.01.2019

Muuß-Merholz J (2011) Barcamp: Die "Un-Konferenz". Bundeszentrale für politische Bildung. http://www.bpb.de/lernen/digitale-bildung/web-2-0/71413/barcamp. Zugegriffen am 10.08.2011

Promotorengruppe Kommunikation der Forschungsunion Wirtschaft – Wissenschaft (Hrsg) (2012) Im Fokus: Das Zukunftsprojekt Industrie 4.0 – Handlungsempfehlungen zur Umsetzung. März 2012. https://www.bmbf.de/pub_hts/kommunikation_bericht_2012-1.pdf. Zugegriffen am 10.01.2019

Schröder L (2014) Digitale Arbeit bedarf der Humanisierung – einige Vorschläge für die Praxis. In: Gute Arbeit, Zeitschrift für Gesundheitsschutz und Arbeitsgestaltung, Nr. 1/2014, S 17–20. https://innovation-gute-arbeit.verdi.de/++file++53db57186f684449d10003cf/download/Zeitschrift_gute_arbeit_2014.pdf. Zugegriffen am 10.01.2019

Will robots take my job? (2019). https://willrobotstakemyjob.com. Zugegriffen am 30.01.2019

ZHAW (2017) Züricher Hochschule für Angewandte Wissenschaften, IAP – Institut für Angewandte Psychologie. IAP Studie 2017 – Teil 2. Der Mensch in der Arbeitswelt 4.0. Ergebnisse der qualitativen Interviews. Online, https://www.zhaw.ch/storage/psychologie/upload/iap/studie/IAP-Studie_Teil-2_Bericht.pdf. Zugegriffen am 10.01.2019

Innovation 4.0: The Agile Evolution of Innovation

Stefanie Paluch, Leif Grube, and Thomas Wittkop

Contents

1 Innovations in the Context of Digital Transformation

It seems that Peter Drucker's statement "the only constant in business is change" is more pertinent to today's business environment than ever before (Christensen 2004). In times of **globalization**, **hypercompetition** as well as digitalization, in which companies must constantly adapt and change, maintaining competitive advantage has become extremely difficult. In general, the world seems to be moving faster, both in the service and manufacturing sectors, where product life cycles are constantly shortening, so competition is continuously increasing due to easier penetration into dissimilar markets (Singh 2004). The real **challenge** for companies is thus to generate competitive advantages in these highly complex environments and to

S. Paluch (✉) · L. Grube · T. Wittkop
TIME, RWTH Aachen University, Aachen, Germany
e-mail: paluch@time.rwth-aachen.de; grube@time.rwth-aachen.de; thomas.wittkop@th-owl.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_57

maintain them as well. To achieve this, companies respond with innovative approaches. Within the last decades, there has been a growing awareness of understanding, integrating, and improving innovation as an essential tool to differentiate from the competition. Moreover, macroeconomic circumstances and various advancements in the technological and digital fields have increasingly forced organizations to think innovatively, as well as to establish this thinking within the respective company. Thus, innovation, both to improve internal processes and results and to maintain **business relationships** with customers and partners, has become an essential component in the business environment (Adelhelm 2012; Becker et al. 2017).

The modern world has undergone three significant changes in recent decades that have gradually but profoundly altered the way people work and live. These include the agricultural revolution, the industrial revolution, and the information revolution (Wang et al. 2015). Today, we are at the beginning of the fourth wave, the **innovation revolution**. This new revolution has become necessary and even possible because of several megatrends that are changing the world and thus the business environment as well as the ecosystem in which people work (Lee et al. 2010). One of these megatrends describes globalization, which influences the world today in any living and working environment. A second major trend is relentless progress. It is not only technologies that include information, communication, transportation, materials, etc., but also, and more importantly, the way these are synthesized to create new technologies (Duncan 2005). People around the world are connected and exchange information through the Internet for personal, social, scientific, and artistic purposes, which has and is constantly developing telecommunications, medicine, education, entertainment, and also social technologies with the help of the existing network economy. Today, more than 4 billion people worldwide use the Internet to share knowledge, conduct e-business, and create value, which is why information and **communication technologies** can be considered as the trigger for a digital world (Neumeier 2017). Due to the exemplified megatrends of globalization and information and communication technologies, purely product-oriented innovations no longer offer a sustainable competitive advantage. Of particular importance above all are the value-added services that a product delivers in the age of 3.0 or 4.0. Such service- and value-oriented innovations are necessary to secure a competitive advantage for a company in the future (Schatz and Bauernhansl 2016).

2 The Term Innovation

Originating from the Latin words "**novus**" (new) and "**innovare**" (renew) and further developed into the word "**innovatio**" (renewal, change), the term "**innovation**" can be traced back to changes caused by technical, social, or economic change (Horsch 2003). The "father" of the concept of innovation is considered to be the Austrian-American economist Joseph A. Schumpeter, who did not invent the term per se, but integrated it into economics and common usage (Snyder et al. 2016).

Since the introduction of the term, the concept of innovation has been widely defined, with the literature agreeing on the characteristic of novelty or renewal through a linkage of different products/processes that have already been there, or the conjunction of already existing and new aspects. This is because an innovation does not always have to exhibit a holistically novelty character, so that early in the literature the concept of innovation is described as a new idea that has its origin in the **recombination** of old innovations (Van de Ven 1986; Robertson 1967). According to this reading, an innovation is what is considered innovative by different stakeholders, customers, employees, etc. (Rogers and Shoemaker 1971; Rogers 1995), which often brings the subjective dimension of innovation into focus. Rogers (1995) thus emphasizes that the notion of innovation is dependent on the respective perceptions of the users and thus represents a social judgment. However, this in no way implies that an individual's subjective impression makes a novel product, technology, or service an innovation. This is only achieved when it is perceived and thus recognized as such by a larger social group.

Innovations can be tiered overall and thus distinguished from one another to a certain extent. For example, inventions such as Thomas Edison's carbon filament lamp or the placement of Apple's first personal computer certainly have a different **innovative character** than, for example, the use of software within a company that is new to that company. In this context, it should also be mentioned that the concept of innovation is viewed in a differentiated sense. A distinction is made in this context between the broad and the narrow concept, whereby the broad concept of innovation implies invention/discovery, the creation of an innovation as well as diffusion and imitation. The narrow concept of innovation, on the other hand, refers only to the phase in which an innovation is created (Fichter 2014; Ahmed and Shepherd 2010).

3 The Stage-Gate Model

Process models are an integral part of innovation management. A look at the literature reveals that various process models for innovation management form the basis of both textbook articles and specialist publications. For example, empirical studies use process models of innovation management to depict or illustrate observable activities. Large companies, on the other hand, develop process models to standardize their innovation activities (Verworn and Cornelius 2000).

Probably one of the most cited and used process models is Cooper's **stage-gate process** model (Cooper 1990). Within the stage-gate model, an **innovation concept**, or the development of a new product or service passes through several distinct steps, which are defined within the model as so-called stages. The individual stages are classified factually, whereby the model varies greatly in terms of the number of stages. In the automotive industry, for example, ten stages are not uncommon. The Cooper standard, however, defines only 4–6 stages through which an innovation passes until it is finally launched on the market (Cooper 1990). In contrast to other process models, the stages of the Stage-Gate model involve cross-functional

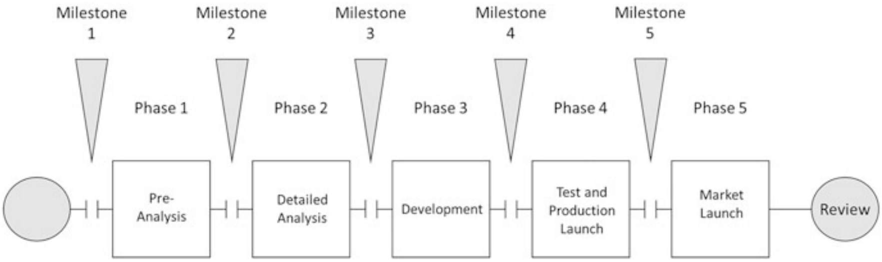


Fig. 1 The stage-gate model (own illustration)

activities from different departments of the company, ensuring interdisciplinary collaborative work with results to be achieved in each case. It also ensures that tasks and decisions are not driven by a single organizational unit alone (Cooper et al. 2002).

Following each stage, activities are measured against previously defined criteria called **milestones** (gate(s)). These milestones provide the basis for discarding or moving forward with ideas or projects. Figure 1 shows an exemplary stage-gate model in which the product to be developed passes through five stages before it is launched on the market.

4 The Evolution of Innovation

Innovation is a result of human creativity to fulfill a certain need. As mentioned, this conception of innovation refers to different areas (innovation for customers, innovation for companies, etc.). Thus, innovations are, in a way, part of the history of mankind, aiming to increase the corresponding quality of life. Within the last decades, the innovation has developed, among others, in the **business organization**.

4.1 Closed Innovation

According to the **closed innovation** model, research projects are launched starting from the scientific and technological base of the company. They undergo a development progress, whereby some projects are not pursued further throughout this process. Other successful projects, however, are selected to be brought to market (Chesbrough 2003a, b, 2012).

Despite the obvious terminology of a closed innovation, Berkhout et al. (2010), with reference to Chesbrough (2003a, b), describe some principles that can also be used to define the closed innovation:

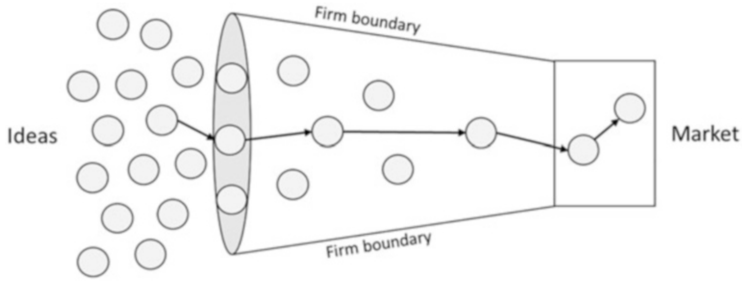


Fig. 2 Closed innovation (own illustration adapted from Chesbrough 2003a, b)

- "The smart people work for us"
- "To profit from R&D, we must discover, develop, produce and ship it ourselves"
- "If we discover it ourselves, we will get it to market first"
- "If we are the first to commercialize an innovation, we will win"
- "If we create the most and best ideas in the industry, we will win"
- "We should control our intellectual property so that our competitors do not profit from out ideas"

The principles of the closed innovation model clearly show that the goal of this approach only serves its own purpose. The focus is on the company's own position in the market and the resultant competitive advantages, so that cooperation or alliances between companies do not represent an alternative at this point. Another characteristic of the closed innovation process is the direction in which the innovation moves. It starts with the company itself and is subsequently brought to market. Figure 2 illustrates the idea of "closed" innovation.

4.2 "Open Innovation"

The concept of open innovation, by which the innovation process is interpreted as an interactive, distributed, and open innovation system, is mainly attributed to Chesbrough (2001, 2003a, b). Compared to and following the closed innovation model, the open innovation model follows a different guideline. The principles presented earlier are now paralleled with the following statements (Berkhout et al. 2010):

- "Not all of the smart people work for us so we must find and tap into the knowledge and expertise of bright individuals outside our company"
- "External R&D can create significant value; internal R&D is needed to claim some portion of that value"
- "We don't have to originate the research in order to profit from it"
- "Building a better business model is better than getting to market firsts"
- "If we make the best use of internal and external ideas, we will win"

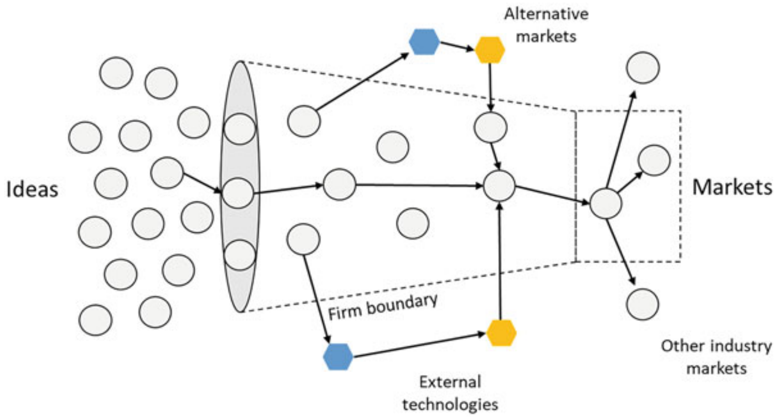


Fig. 3 Open innovation (own illustration adapted from Chesbrough 2003a, b)

- “We should profit from others’ use of our intellectual property and we should buy others’ intellectual property whenever it advances our own business model”

As is also evident from the terminology of the concept of open innovation, the focus of the companies is directed with regard to the innovations to be placed on the market. The term “open” refers on the one hand to the inclusion of external technologies and experiences of other companies, which are utilized in addition to the company’s own expertise in the creation of new products/services. On the other hand, it also refers to collaborative practices between companies. If ideas are promising, but cannot be pursued further due to various circumstances, such as the lack of technology or knowledge, or if the company’s own competencies within the organization would not be sufficient to achieve a sustainable competitive advantage, they can be passed on to other companies, for example via licenses.

In this way, the market share and competitiveness of a company in its own market is not necessarily further developed or expanded, but it is possible through such an approach to open up new markets together with other companies or to create alliances from which both parties participate. This is clearly illustrated in Fig. 3.

The concept of the closed innovation model in the understanding of self-sufficiency or self-sufficiency for global competition is increasingly becoming an impossible goal, which is why even world-leading companies within their respective industries are always on the lookout for cooperation partners to conceptualize an innovative value chain in which core competencies are combined (Tapscott 2006). The goal of the open innovation model is thus to unite a creative convergence of the company’s own with external expertise.

However, external expertise does not refer exclusively to other companies, but also to the customers to be reached or the market to be served.

4.3 The Customer in the Innovation Process

Not all innovations are adopted by the market, resulting in **flop rates** of over 50% depending on the industry in question (Cooper 1993). To avoid this, interaction between the company and the market to be reached is partly unavoidable, so that risks related to the target market are reduced as far as possible in the early phase of the innovation process. For this **risk minimization**, Thomke (2003) proposes two basic types of information.

In this context, "**need information**" refers to general information about the respective customer and market needs. Thus, this information includes the preferences, desires, satisfaction determinants, and buying motives of current and potential customers. "**Solution information**," on the other hand, describes the technological possibilities and potentials for effectively and efficiently channeling the respective customer needs into a specific service. Typically, these two types of information cannot be considered separately in a clear sequence and within a suitable process. This way a company tries to generate the need information on the market via different market research techniques and then to transfer these specifically and by application of the internally available solution information into a new product or a new service. Consequently, customers are seen as the source of needs information and companies as the locus of solutions information (Reichwald and Piller 2005).

Consequently, Schumpeter's (1942) image of a "lone innovative firm" gives way to a kind of network of diverse actors involved in an innovation process (Brown and Eisenhardt 1995; Freeman and Soete 1997; Laursen and Salter 2004; Piller 2003, 2004; Reichwald et al. 2004; von Hippel 1988). Consequently, the success of an innovation is largely based on the company's ability to involve different actors and a wide variety of perspectives of the respective stakeholders in the innovation process (Hirsch-Kreinsen 2004).

5 Innovation 4.0

In times of digitalization and the resulting improved information and communication technologies, which are strongly integrated into the everyday life of existing and potential customers, the innovation environment of companies is also changing.

The need to continuously innovate products and services is not only a task for start-ups, but also for established companies (Obermaier 2016). New technological insights and the increasing usage of data in the digital age, in addition to roles that are increasingly becoming important, such as the software developer and network operator, as well as the so-called data aggregators and platform operators, also create room for further development of every company (Lee et al. 2012). Technical possibilities that have originated and continue to emerge from Industry 4.0 always create further potential for innovative offerings. Thus, product and service innovations can be based on highly differentiated and customized products, synchronized product and service combinations, or value-added services. At the same time, dynamic configurations of the value chain, higher resource as well as equipment efficiency and thus cost reductions in a wide variety of areas are made possible. Examples include the possibility of flexible and dynamic reconfiguration of production capacities, shorter time-to-market, higher scalability, lower reject rates or preventive maintenance (ibid.).

However, Industry 4.0 and the opportunities it entails oppose the existing corporate culture at many established companies, which hinders the company's strategic flexibility, especially when it comes to innovative or disruptive solutions. As a result, a suitable organizational setup, including the project structure as well as the leadership and lived culture within a company, has been identified as a critical success factor for innovation projects (BMBF 2013; Kagermann et al. 2013; Rudtsch et al. 2014). To get around this, established companies, which have a strong and longstanding corporate culture, are advised to introduce particularly disruptive innovations in the 4.0 era in independent, small units.

Innovations additionally create value for the respective company and for existing, as well as potential, customers. On a customer basis, value creation through innovation consists of reinventing the concept of customer value (Hirsch-Kreinsen 2004). This is a particularly prolific area when it comes to co-creating value with customers. Traditional customer values, such as price, quality, speed, and customization are still essential aspects. Today, however, customers demand more than just experiences, emotional fulfillment, or public benefit. They want to be involved in the process of creating a product or service as well as have the opportunity to learn new things. Therefore, an expansion of the customer base should be created. E-customers (those customers who shop online), global customers (customers from abroad), customer communities that have a huge impact on the entire market, and non-customers who are potential future customers are also of particular importance (ibid.). Creating customer value for all customers, as well as differentiated value for specific customer groups, requires innovation.

From a business perspective, innovations affect the value chain, making it more efficient, which in turn reduces costs, improves quality and/or increases the speed of the process. Many process innovations such as Just-in-Time, TQM Six Sigma, Lean Manufacturing, etc. aim to make the value chain more efficient (Pan and Li 2016).

To increase value creation, through innovation on both the customer side and the entrepreneurial side, new business models as well as the integration of different technical possibilities created by Industry 4.0/digitization are relevant. Business models represent the approaches that strategically align the organization to produce and deliver its goods or services to the customer. The Internet has established business models such as e-banking, e-business, e-government, electronic markets, e-auctions, etc., thereby revolutionizing various models. Accordingly, the incorporation of digitization and the Internet results, for example, in an increase in the speed of transactions as well as an improvement in the exchange of information and the quality of service in the provision of new solutions to customer problems (Kim and Mauborgne 2005).

5.1 *Innovation 4.0 from a Process Perspective*

A **process** is a recurring sequence of activities in a predecessor-successor relationship, with a defined start and end time. Within the manufacturing industry, the goal of a process is to transform a specific, usually material, input into one or more outputs in a value-enhancing manner (Binner 2002). In the business environment, however, processes do not only refer to a specific output, so that a process can also be described as a path to a specific goal.

As shown, innovations are not always something entirely new, but can also be a novel combination of different innovations. In the 4.0 era and the associated new possibilities through digitization and the use of new technological conditions, it makes sense to take a closer look at processes that are practiced in established companies and, if necessary, to evaluate them critically. Processes that have always run in the same way can be made more efficient with the help of new (technological) possibilities for customers, as well as within the company itself, so that time and resource savings can be achieved. In addition, the quality of the products and/or services can also be increased.

5.1.1 Innovation 4.0: Customer Processes

Modern companies such as Amazon or Netflix recognized this early on and managed to simplify the processes of their (potential) customers in such a way that the quality of life is increased. Today, Netflix has matured into one of the largest online media libraries in the world (Ecco 2014). Prior to digitization and access to the new opportunities created by the Internet, Netflix offered online DVD rentals (Anderson 2007). Customers could order movies online, but DVDs and later Blu-rays, were shipped by mail. Netflix created an innovation within the industry by facilitating the process of DVD rental for customers. However, this facilitation was countered by the problem of customers having to wait for their orders. This and other reasons hardly brought any revenue to the company within the first few years (Stelter 2011). Nowadays, movies and series on Netflix are accessible at any time in almost any place, creating more value for customers by using new possibilities within the processes. Figures 4 and 5 illustrate the process change and resulting time savings for customers through the use of new opportunities on the part of Netflix:

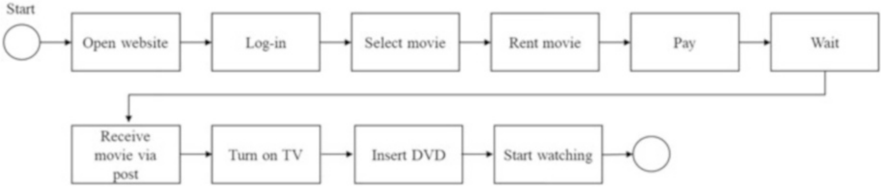


Fig. 4 Netflix business model at inception (own illustration)

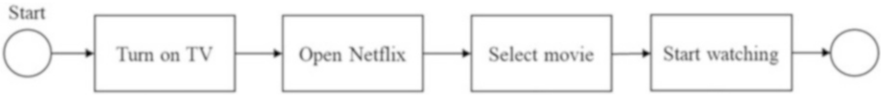


Fig. 5 Netflix's digital business model (own illustration)

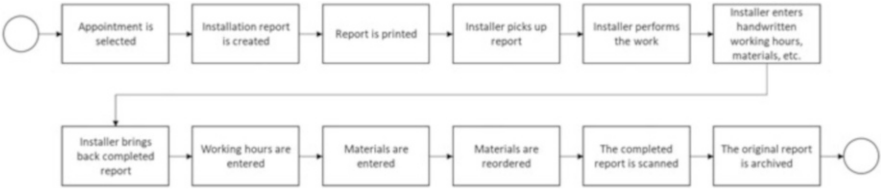


Fig. 6 Process of an installer in the plumbing industry, without the use of networked technologies (own illustration)

5.1.2 Innovation 4.0: Company Processes

The innovative approach outlined above for the example of Netflix, can also be considered for internal processes in several other industries. Thus, the use of new software or new technologies, as illustrated, can be described as an innovation. Innovation in this framework refers to a company for which the use of a new software or a new technology is innovative.

Innovation also incentivizes the creation of new business models with a focus on the new possibilities of information and communication technologies. However, this relates to the internal corporate context. Figure 6 shows an exemplary process of a fitter in the plumbing industry without the use of such new technologies.

The figure illustrates that tasks within the process take up time and money. For example, the fitter must make two trips every day to collect his work instructions and submit reports accordingly. In addition, the handwritten work report hast to be passed on to the HR department so that the time worked can be recorded digitally. Furthermore, the material used must be reordered manually. However, with the new possibilities offered by digitization and the resulting connection of different devices to the existing software within a company, it is possible to streamline this process. With the utilization of a tablet that provides access to the company's internal software, an innovation is introduced within the company, as Fig. 7 shows.

By linking existing processes and new technologies, Innovation 4.0 saves companies valuable time so that important human resources can be deployed elsewhere. In addition, Innovation 4.0 reduces costs, which can be used for further innovative approaches.

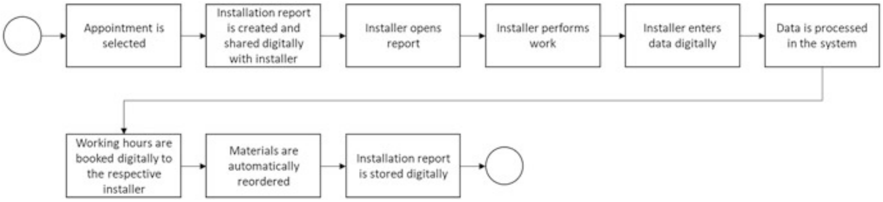


Fig. 7 Process of an installer in the plumbing industry with the use of networked technologies (own illustration)

5.2 *Agile Innovations*

The idea of 'agile innovations' originates from the domain of software development, which is characterized by digitization and a dynamic business environment. In contrast to traditional process models such as the stage-gate (Cooper 1986) approach, which is characterized by planned and controlled processes as well as systematic decisions, agile innovations are characterized by fast, short iteration cycles in which the product is systematically adapted to requirements through customer integration. The four basic ideas of agile innovation are summarized in the Agile Manifesto (Beck et al. 2001): (1) Individuals and interaction are central components of the agile development process and emphasize close collaboration and open communication within the team. (2) Working software is more important than following processes and documenting steps. (3) Customer integration encourages rapid feedback so that innovations can be better aligned with customer requirements. (4) Willingness to change is a basic prerequisite of the agile innovation process.

The Authors Paluch, Brettel, Hopp, Salge, Piller, and Wentzel (TIME Research Area) of RWTH Aachen University have studied the phenomenon of agile innovation from a management perspective and discuss in a Special Issue of the Journal of Business Research on "Innovation in the digital Age: From Stage-Gate to an Agile Development Paradigm?", whether traditional product development models such as Stage-Gate (Cooper 1986) are still pertinent in today's digital age or whether these models should be largely replaced by agile approaches even in more traditional contexts (e.g., product development). The basic idea of conventional models of how innovation is organized view innovation as a deterministic process that can be planned ex-ante and then executed and controlled. The innovation process is characterized by a series of predefined decision stages. At these points, a decision is then made to pursue or reject an idea based on criteria. In contrast, agile approaches are stochastic. They follow the understanding of an iterative planning cycle, where the results of a short stage of execution determine the design of the following stage, and so on. Uncertainties are discovered and continuously addressed during the execution process. Stage-gate aims to control uncertainty upfront to avoid changes later, while agile development focuses on adaptation and changes during the development processes. To do this, it is necessary to integrate the customer at every stage of the iterative development design and application process to respond as quickly as

Table 1 Comparison of stage-gate and agile innovation (Paluch et al. 2019)

	Stage-gate development	Agile development
Type	Macroplanning	Microplanning
Domain	Hardware development	Software development
Purpose	Investment model for sequential resource allocation	Tactical model for guiding largely self-managed teams
Focus	Risk and quality	Learning and speed
Logic	Deterministic	Stochastic
Directionality	Largely linear	Highly iterative
Scope	Idea to launch	Development and testing
Owner	Cross-functional team (R&D, marketing, sales, operations)	Technical team (software developers, engineers, project managers)
Customer involvement	Episodic	Continuous

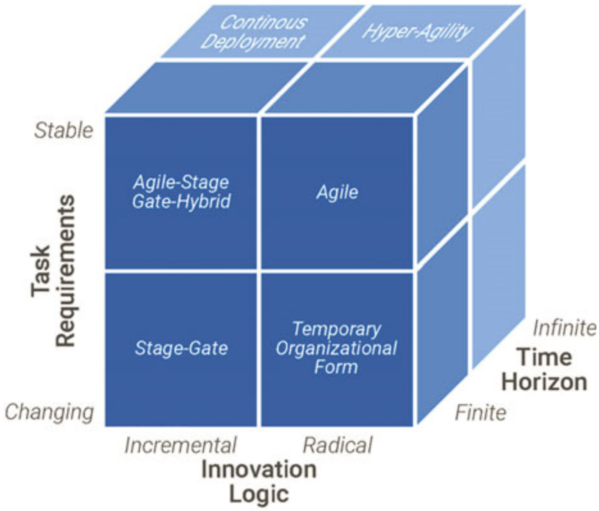


Fig. 8 Task, innovation, and time trade-offs (Paluch et al. 2019)

possible to new requirements and improvements (Paluch et al. 2019). Table 1 shows a comparison of the two approaches.

Companies are now faced with the question of how to respond to changes brought about by digitization and changing customer needs in such a way to reorganize their innovation process. To this end, the authors propose various dimensions that could support companies in deciding on the appropriate process model. This multidimensional model is shown in Fig. 8.

The initial task is to clarify whether the company is more likely to rely on incremental or radical innovations in the future. The nature of innovation determines to a large extent the innovation approach to be selected. The second decision relates to the task requirements. Are requirements rather known and stable and thus also the risk manageable, or is the company confronted with constantly changing and supposedly unknown tasks and thus also high risk? The third dimension comprises the time horizon (limited or unlimited) that companies plan for the implementation of their innovation. Classic stage-gate models seem best suited for incremental innovation in seemingly predictable contexts where customer and task requirements are well understood and largely stable (Bianchi et al. 2019). *Temporary organizational* forms are used when radical innovation is sought in relatively well-understood contexts with known requirements, provided that the knowledge is reintegrated into the organization (Fecher et al. 2019). The concept of continuous integration originated in software engineering and aims to enable organizations to continuously learn (in real time) from customer usage. This approach is now used by Amazon, Facebook, Google, and Netflix. Continuous integration offers several advantages over other existing software development methods, ranging from increased productivity to higher developer motivation, lower software risks and higher software quality. However, the use of the concept of continuous deployment can naturally come at the cost of a lack of control and predictability in the software lifecycle (Savor et al. 2016). *Agile stage-gate hybrids*, in turn, are a particularly attractive form for organizing incremental innovation in the presence of task uncertainty (Ghezzi and Cavallo 2019). Finally, Agile models in their pure form come into focus when radical innovations are in demand in dynamic contexts with unique and yet-to-be-discovered task requirements. Hyper-agile organizational structures, such as at Chinese appliance company Haier, require strategic and operational rethinking at all levels. Haier sees itself more as a platform and encourages employees to think and act independently as entrepreneurs. Everyone is expected to realize his or her potential and thus contribute to the continuous transformation of the company in his or her own way (Hamel and Zanini 2016).

6 Summary

Innovations in the age of the 4.0 hype are primarily aimed at the continuous integration of the customer and the general simplification of corporate processes. Many sequences that were previously carried out manually can now run automatically or be supported using these innovations thanks to modern, new types of information and communication technologies and various technical innovations. Innovations can thereby make life and work easier if they are used or developed in a purposeful manner. This article is meant to illustrate the importance of agile processes in the development and deployment of innovations.

Based on a clearly presented (modeled) overview of the processes within the company or from a customer's perspective, Innovation 4.0 can be developed strategically. In addition, the company can select the appropriate innovation approach based on the innovation logic, task requirements and time horizon and thus systematically promote innovations.

References

Adelhelm S (2012) Geschäftsmodellinnovationen. Open Innovation in Life Sciences. Gabler Verlag, pp 53–82

Ahmed PK, Shepherd C (2010) Innovation management: context, strategies, systems, and processes. Pearson Prentice Hall, Harlow, UK

Anderson N (2007) Netflix offers streaming movies to subscribers. In: Ars Technica

Beck K, Beedle M, Van Bennekum A, Cockburn A, Cunningham W, Fowler M, Grenning J, Highsmith J, Hunt A, Jeffries R, Kern J (2001) The agile manifesto. Agile Alliance

Becker W, Patrick U, Stradtmann M (2017) Geschäftsmodellinnovation als Wettbewerbsvorteil mittelständischer Unternehmen

Berkhout G, Hartmann D, Trott P (2010) Connecting technological capabilities with market needs using a cyclic innovation model. R&d Manag 40(5):474–490

Bianchi M, Marzi G, Guerini M (2019) Agile, stage-gate and their combination: exploring how they relate to performance in software development. J Bus Res (forthcoming). https://doi.org/10.1016/j.jbusres.2018.05.003

Binner HF (2002) Prozessorientierte TQM-Umsetzung, 2. Aufl., München/Wien, S. 77

BMBF (2013) Hightech-Strategie: Zukunftsbild Industrie 4.0, Bonn

Brown JS, Eisenhardt KM (1995) Product development: past research, present findings and future directions. Acad Manag Rev 20(2):343–378

Chesbrough H (2001) Open innovation: a new paradigm for managing technology, OECD Conference on New Business Strategies for R&D, 22 October 2001

Chesbrough H (2003a) Open innovation: the new imperative for creating and profiting from technology. Harvard Business School Press, Boston

Chesbrough H (2003b) The era of open innovation. Sloan Manag Rev 44(4):35–41

Chesbrough H (2012) Open innovation: where we've been and where we're going. Res-Technol Manag 55(4):20–27

Christensen KS (2004) A classification of the corporate entrepreneurship umbrella: labels and perspectives. Int J Manag Enterp Dev 1(4):1–15

Cooper RG (1986) Winning at new products. Addison-Wesley, Reading, MA, pp 1–14

Cooper RG (1990) Stage-gate systems: a new tool for managing new products. Bus Horiz 33(3): 44–54

Cooper RG (1993) Winning at new products: accelerating the process from idea to launch, 2nd edn. Perseus Books, Boston

Cooper RG, Edgett SJ, Kleinschmidt EJ (2002) Optimizing the stage-gate process: what best-practice companies do. I. Res-Technol Manag 45(5):21–27

Duncan DE (2005) Masterminds: genius, DNA, and the quest to rewrite life. Harper Perennial, New York

Ecco G (2014) Netflix – Ein Streamingdienst mischt die Unterhaltungsbranche auf. In: bluemind.tv. 2014

Fecher F, Winding J, Hutter K, Füller J (2019) Innovation labs from a participants' perspective. J Bus Res (forthcoming). https://doi.org/10.1016/j.jbusres.2018.05.039

Fichter K (2014) Grundlagen des Innovationsmanagements. Carl von Ossietzky Universität, Oldenburg

Freeman C, Soete L (1997) The economics of industrial innovation. Pinter, London

Ghezzi A, Cavallo A (2019) Agile business model innovation in digital entrepreneurship: lean startup approaches. J Bus Res (forthcoming). https://doi.org/10.1016/j.jbusres.2018.06.013

Hamel G, Zanini M (2016) The $3 trillion prize for busting bureaucracy (and how to claim it). Humanistic Management Network, Research Paper Series No. 28/16. Available at SSRN: https://ssrn.com/abstract=2748842 or https://doi.org/10.2139/ssrn.2748842

Hirsch-Kreinsen H (Hg.) (2004) Innovationsnetzwerke: ein anwendungsorientierter Leitfaden für das Netzwerkmanagement, VDI-Leitfaden, Düsseldorf: VDI Verlag

Horsch J (2003) Innovations-und Projektmanagement. Gabler, Wiesbaden

Kagermann, H./Wahlster, W./Helbig, J. (2013) Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Abschlussbericht des Arbeitskreises Industrie 4.0., Frankfurt/Main: Forschungsunion Wirtschaft-Wissenschaft/acatech
Kim, C. W., Mauborgne, R. (2005) Value innovation: a leap into the blue ocean. J Bus Strategy 26.4, 22-28
Laursen K, Salter A (2004) Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms, Working paper, Tanaka Business School, Imperial College London/Copenhagen Business School, presented at the AOM 2004 Meeting in New Orleans, LU, 2004
Lee SM, Olson DL, Trimi S (2010) The impact of convergence on organizational innovation. Organ Dyn 39(3):218–225
Lee SM, Olson DL, Trimi S (2012) Co-innovation: convergenomics, collaboration, and co-creation for organizational values. Manag Decis 50(5):817–831
Neumeier A (2017) Wert der Digitalisierung–Erfolgreiche Auswahl von Digitalisierungsprojekten. HMD Praxis der Wirtschaftsinformatik 54(3):338–350
Obermaier R (2016) Industrie 4.0 als unternehmerische Gestaltungsaufgabe – Betriebswirtschaftliche, technische und rechtliche Herausforderungen, Passau
Paluch S, Brettel M, Hopp C, Piller F, Salge O, Wentzel D (2019) Stage-gate and agile development in the digital age: promises, perils, and boundary conditions. J Bus Res forthcoming
Pan X, Li S (2016) Dynamic optimal control of process–product innovation with learning by doing. Eur J Oper Res 248(1):136–145
Piller F (2003) Von Open Source zu Open Innovation. Harv Bus Manager 25(12):114
Piller F (2004) Innovation and value co-creation, Habilitationsschrift an der Fakultät für Wirtschaftswissenschaften der Technischen Universität München 2004
Reichwald R, Piller F (2005) Open Innovation: Kunden als Partner im Innovationsprozess
Reichwald R, Ihl C, Seifert S (2004) Innovation durch Kundenintegration, Arbeitsbericht des Lehrstuhls für Allgemeine und Industrielle Betriebswirtschaftslehre an der Technischen Universität München, Nr. 40 (Juni 2004)
Robertson TS (1967) The process of innovation and the diffusion of innovation. J Market 31(1): 14–19
Rogers EM (1995) Diffusion of innovations: modifications of a model for telecommunications. Die diffusion von innovationen in der telekommunikation. Springer, Berlin, Heidelberg, pp 25–38
Rogers EM, Shoemaker FF (1971) Communication of innovations; a cross-cultural approach
Rudtsch V, Gausemeier J, Gesing J et al (2014) Pattern-based business model development for cyber-physical production systems. Procedia CIRP 25:313–319
Savor T, Douglas M, Gentili M, Williams L, Beck K, Stumm M (2016, May). Continuous deployment at Facebook and OANDA. In Software Engineering Companion (ICSE-C), IEEE/ ACM International Conference on (pp 21–30). IEEE
Schatz A, Bauernhansl T (2016) Geschäftsmodell-Innovation en: Profitabler wirtschaften mit hohem Vernetzungsgrad. Handbuch Industrie 4.0: Produktion, Automatisierung und Logistik (2016): 1–15
Schumpeter JA (1942) Capitalism, socialism and democracy. Harper, New York
Singh S (2004) Market orientation, corporate culture and business performance. Ashgate Publishing, Farnham
Snyder H, Witell L, Gustafsson A, Fombelle P, Kristensson P (2016) Identifying categories of service innovation: a review and synthesis of the literature. J Bus Res 69(7):2401–2408
Stelter B (2011) For Netflix, higher earnings and a milestone. In: The Now York Times
Tapscott D (2006) Winning with the Enterprise 2.0. New Paradigm Learning Corporation, pp 1–59
Thomke S (2003) R&D comes to services. Harv Bus Rev 81(4):70–79

Van de Ven AH (1986) Central problems in the management of innovation. Manag Sci 32(5): 590–607

Verworn B, Cornelius H (2000) Modelle des Innovationsprozesses. In: Leibniz-Informationszentrum Wirtschaft, Working Paper, Nr. 6

von Hippel E (1988) The sources of innovation. Oxford University Press, Oxford

Wang Y, Rolls ET, Howard N, Raskin V, Kisner W, Murtagh F, Shell DF (2015) Cognitive informatics and computational intelligence: from information revolution to intelligence revolution. Int J Softw Sci Comput Intell (IJSSCI) 7(2):50–69

Part VIII
Humanities and Social Sciences

Industry—Between Evolution and Revolution: A Historical Perspective

Paul Thomes

Contents

1 Prologue: Methods, Concepts, Intentions

Looking back, it seems that global social insecurity has never been as disconcerting as it is today. The reasons for this are to be found in acutely experienced socio-technical and political processes of change. Change has always been a continuum of human existence, but what is new is how it is now perceived to be changing humanity faster and more fundamentally than ever. After all, digital technologies are literally turning everything upside down; they are relativising the world and transforming with historically unprecedented quality all aspects of human existence. Mankind is facing an absolute novum and the most radical of all socio-technical upheavals (Ede 2019; Wolff and Göbel 2018).

As a last consequence, human beings might relinquish or, more precisely, lose their hitherto exclusive primate role of intelligent 'doers' to autonomous machinery that they themselves have ingeniously designed. It would happen now when they are no longer in a position to pull the proverbial plug (Grunwald 2018) and it would constitute a new dimension of control loss. From today's standpoint, the very

P. Thomes (✉)
RWTH Aachen University, Aachen, Germany
e-mail: thomes@wisotech.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_58

worst-case scenario. However, until that potentially happens, a fair length of time will have elapsed. Let us make good use of it!

It is true that for some time now, mankind has been caught up in an intoxicated state of doing everything that it is capable of doing. Humans are making the Earth subservient and, thanks to algorithm-driven technology, they are more flexible and mobile than ever, be it in the analogous or the digital sphere. In theory, pocket-sized (yet practically omnipotent!) digital personal assistants are now available to all 8 billion inhabitants of Earth, enabling consumers and producers alike to operate 'cost-free' in real time anywhere around the globe at any time. These possibilities are condensing the world tremendously; in a way, they are shrinking it into a village. This process brings with it opportunities, risks, hopes, and fears of loss. Production migrates effortlessly towards the best conditions; individuals do the same. Just as millions of Europeans with no prospects were drawn towards America in the eighteenth and nineteenth centuries, nowadays the more developed part of the world, which is stagnating demographically, attracts people from other parts of the world. They are eager to be part of this growing—but extremely unevenly distributed—global prosperity. And who would deny them the chance? (Zinn 2018).

A similar situation exists for capital and goods. The financial sector is moving capital around the world via digital platforms and in digital currencies in real time. Geographical distances are now irrelevant for even the cheapest of mass products, even by plane. The goods washed up along the North Sea coastline following the January 2019 disaster involving the container ship MSC Zoe (with its capacity of 19,500 containers, one of the world's largest transport vehicles) dramatically symbolised this situation. Highly efficient logistics systems, supported by a liberal economic order, enable flexible, worldwide-operating supply and distribution chains.

Due to its extensive interactive networking, the currently unfolding Internet of Things and Services—integral components of what digital speech coins 'Industry 4.0'—is opening up completely new production and consumption perspectives. Although it is not easy to foresee what their effects will be, what is certain is that they will existentially change humanity.

This would implicate that we have a definition of the overarching objective of the production concept 'Industry 4.0', namely that it should be understood as a subsystem of 'Society 4.0' (whatever the precise composition of this subsystem may be, as it is still in its early development phase). Any ideas about its future are correspondingly disparate, all the more since 4.0 has advanced in the meantime from a national label to a global one (Kagermann et al. 2011).

The gulf between praise and criticism is thus a wide one, but we regard the initiative in a positive light. On the one hand, it addresses the relevance of 4.0 for society as a whole; on the other hand, and differently from in the past, it signalises the will to consciously conceptualise change. This change will explicitly serve people's welfare, such as ambitiously formulated (and rightly so!) by the German federal government's High-tech Strategy 2025.

We are looking forward to the results of these endeavours—even more due to the rapid disruptive innovation from the Latin for 'break apart'/'shatter') of Industry

5.0—foreseen as the next stage—which promises us a mathematical revolution combined with true artificial intelligence, i.e. intelligent algorithms in the context of self-learning social machines (Lübbecke 2015; Manzei et al. 2016).

What remains to be noted is that mankind, aware of its own capabilities, is vacillating between horror and awe. It is precisely for this reason that we need to reflect upon developments.

One option would be to retrospectively interpret the actual state of things (which is why this article also has its place in this book). Five hundred years after the death of that ingenious Renaissance visionary, Leonardo da Vinci, this would at least provide a pleasant reminiscence.

How often have we learned from history that we have learned nothing from history! Yet it does not have to be that way; wrong conclusions are often simply the result of subjective anecdotal evidence which, for the most part, is of no use as a basis for decisions. What is needed is empirical evidence to sustainably transfer experiences and expectations to the future. As empirical evidence helps us to identify relevant parameters, events or developments more reliably and to assess them and their consequences more accurately, it thus helps us to generate a viable basis for producing a humane future (Radkau 2017).

A look at the wider picture is imperative; industrial change cannot be understood in isolation. A holistic approach is needed to interpret the object of study within socio-cultural contexts. The outcome is a more advanced understanding of complexity and of relevance. Considering the dramatically depleting resources and acute environmental changes, such an approach is more essential than ever. For this reason, we focus our research on causes and effects, i.e. on structural aspects, by using the STEPLE concept as our basis. Case studies will serve to illustrate (Peters and Thomes 2018; Timm 1964) (Fig. 1).

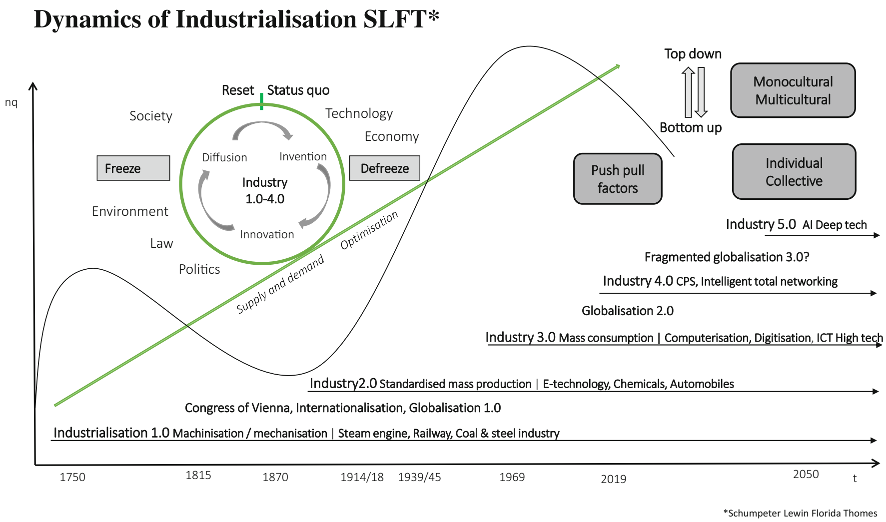


Fig. 1 STEPLE interdependency model (own illustration)

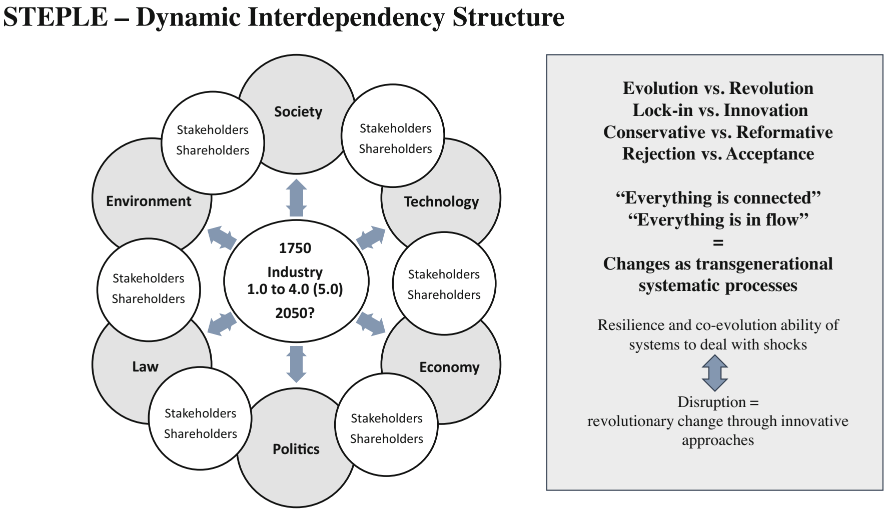


Fig. 2 Change model SLFT Schumpeter, Lewin, Florida, Thomes

The essence of change processes can be effectively broken down into three phases by a combination of two explanatory approaches to innovation processes (Schumpeter) and change processes (Lewin). In practice, processes of this kind often proceed on many different levels and at many different speeds: they are highly complex constellations offering hardly any really assessable opportunity and risk potential to either stakeholders or historians. The Great Reset concept, introduced by Richard Florida, will help us to relativise the generally negatively connotated effects of situations in crisis. In our specific case, the reset denotes the milestones of industrialisation (Florida 2010) (Fig. 2).

As most of us probably know, 'industry' is a loanword from the Latin noun *industria,* which may be defined, inter alia, as work, diligence and activity. At the end of the eighteenth century, it was Adam Smith who paved the way for this term to achieve global recognition in its modern sense. Periodisations always have an arbitrary aspect to them. There is constant discussion about meaningfulness and meaninglessness, as van der Pot has meticulously shown us. It is a fact that periodisation can optimise process understanding (Ede 2019; Hessler and Thorade 2019; van der Pot 1999).

What digital speech coins 'Industry 1.0' is the result of a long development process. This is signalised, inter alia, by the technical terms which periodise this phase: proto-industrialisation, early industrialisation, industrial revolution, peak industrialisation. In the broadest sense, they cover a time span of at least 400 years (van Zanden 2009). The industrialisation phase, which is alternatively known as the factory industry or the machine age, began about mid-eighteenth century, whereby the steam engine was the initial technological impetus.

This is not to say that there had been no mechanical processes in existence beforehand; there were, for instance, water mills and windmills as well as machines driven by muscle power. However, the steam engine was the first device capable of performing mechanical work in a constant, controllable and effortless way. Put simply, mechanical energy complemented natural and human energy (Paulinyi 1997). In addition, there were socio-economic changes taking place. The term 'Industrial Revolution' is itself indicative of their significance. This particular designation came into being in the first half of the nineteenth century, analogously to the political term 'French Revolution'—which itself drew upon the Copernican Turn and England's Glorious Revolution of 1688.

The term 'Second Industrial Revolution' was presumably first coined by the French philosopher and sociologist G. Friedmann in 1936. He addressed man-machine problems and identified electrification, which had been diffusing since the 1870s, as the initial technological impetus behind this particular revolution (van der Pot 1999).

Friedmann is also said to have started to extend his concept by adding a third phase to it in the 1960s, as even more rapid development post-1945 was taking place. This extension regarded automatisation and peaceful utilisation of atomic energy to be the initial driving forces. Van der Pot argues that energy began to replace the human brain—a process that led smoothly onto the fourth phase (van der Pot 1999).

For the sake of simplicity, we will in the following apply the most used terminology: 1.0 to 4.0 (Kagerman et al. 2011).

In reality, the transition from one period to another is a seamless process. For instance, the steam engine asserted its relevance well into the twentieth century; the electric engine and the combustion engine were already making their mark in the last quarter of the nineteenth century, while the traditional leading industries of textiles, railways as well as coal and steel, all contributed to the growth of the future key sectors of electrotechnology chemicals and mechanical engineering. Except for the steam engine, all of them are still highly relevant factors in the 4.0 phase.

It is an undeniable fact: periodisation makes sense. However, whether four or more phases are necessary is another matter. We will return to this point later. We would like to highlight here, though, an interesting qualitative aspect. The first three periodisations emerged more or less observationally and descriptively ex post out of different experiences; the fourth period, in contrast, was proclaimed ex ante out of acute experience. The latter period is also seen as a consciously proactive concept that needs to be designed, and this book concerns itself very intensively with that process. The following chapters will take a detailed look at the pathways leading us there.

2 Industrialisation

Because of its revolutionary impacts, industrialisation is one of the most heavily researched themes, the research covering all possible facets. The contribution at hand should, then, only be understood as an interpretive and inspiring synopsis on Industry 4.0. It makes no claim to completeness. Likewise, the bibliography represents only a small selection of authors. The following three works are intended as a compact but stimulating impression of the multitude of ideas on the theme. In 1969, D.S. Landes positively subsumed the technological aspects of industrialisation under the label of *The Unbound Prometheus* (Landes 1969). From an optimistically critical viewpoint, S. Pollard regarded industrialisation to be a *Peaceful Conquest* in 1981 (Pollard 1981). C. Roser summarised the manufacturing productivity of industrialisation in his dynamic message: *Faster, Better, Cheaper* (Roser 2016).

2.1 Industry 1.0 (the First Industrial Revolution)

Much simplified, the transition narrative is the following: An accumulation of innovations led to significant increases in productivity that gave rise to the overcoming of the agrarian subsistence economy, culminating in an unprecedented growth phase. It enabled a simultaneous increase in population and in prosperity. Behind this simplified outline exists a complex interplay of cause and effect which will be analysed in the following.

The availability and mobility of the production factors ‘labour’ (in its two dimensions ‘physical’ and ‘cognitive’), ‘soil’ and ‘capital’ are prerequisites for dynamisation. In our specific case, mobility refers to the possibility of the stakeholders to make their own decisions. A further precondition is that of absorptive markets.

If such a constellation occurs, significant socio-technological change follows. Whether such change is actually desirable remains open to question. The word ‘progress’ itself means nothing more than a neutral movement forward, yet we see diverse historical developments which are all interpreted differently (Rössner 2017). What researchers do concur on is that England—or Great Britain—delivered the blueprint for industrialisation. Clearly in the eighteenth century, all the factors relevant for transformation first came together in that country in the required intensity (Berg and Hudson 1992; Buchheim 1994; Roser 2016).

A vital impulse had been provided back in the sixteenth century: the accelerating expansion of systematic knowledge within the ‘Gutenberg galaxy’ through the invention of letterpress printing (Burke 2000/2012; Castells 2001; Dean 1982; Ede 2019). Together with a parliamentarily and constitutionally protected legal and political framework, enlightened rationality and scientific principles formed the institutional basis. They guaranteed, inter alia, the liberty of the individual and private ownership, i.e. the ‘mobility’ of the production factors. This momentum was flanked by the Protestant faith ethics that were prevailing in the sixteenth century during the European Reformation movement. Put another way: here we have a re-occurrence of the Copernican Turn, which was presented by the thinker himself in his controversial work *De revolutionibus orbium coelestium*—published in Nuremberg in 1543 and defining his heliocentric concept of the universe.

The founding of the Royal Society as a scientific academy in 1660 and the Glorious Revolution of 1688 exemplify our thesis, although it was to be another

100 years before Continental Europe, prompted by the French Revolution, experienced the end of its change-resistant, archaic feudal system. Scholars such as Locke, Hume or Newton—and later Smith and Ricardo—laid the theoretical foundation for change.

It is no coincidence that in around 1700, Britain was already deemed to be a modern, widely commercialised society. Although the estimation by the chronicler Gregory King at the end of the seventeenth century that one third of the population consisted of shop owners, craftsmen and merchants seems rather doubtful, it nevertheless does reflect the direction in which things were going.

In 1776, Smith's pioneering macroeconomic analysis on the wealth of nations ultimately defined the potential of rational economic behaviour. By deliberately exaggerating the negative aspects of the mercantilist system—so dubbed by Smith himself—he distanced himself from it. No longer was the focus on a zero-sum game, where government-related enrichment was at the expense of others (the beggar-your-neighbour principle); the focus was now on productive labour. With his 'economic man', Smith created an ideally rational, utility-maximising individual whose relevance still holds today. The state needed to do no more than to ensure that compatible guidelines were set. Thus, Smith's work laid the foundation of economic theory for industrial development. Incidentally, this happened in the year that the American colonies declared their independence and Watt's steam engine met with success. The division of labour and standardised parts and processes were the keys to productivity (Ede 2019; Roser 2016). Ricardo in 1817 practically only needed to slot in an international theoretical framework together with his thoughts on comparative advantage and free trade (Ede 2019; Roser 2016).

The favourable transport conditions of the British Isles were reflected in all of this. No town is farther than 70 km away from the coast, and there is no lack of waterways. This was an important point because in the pre-industrial age, the by far most efficient method for transporting goods was that of water transport. Although not faster than by road, this method was less costly because it was more energy efficient and possessed more volume potential, particularly for bulk goods such as coal and ore. It is no coincidence that the canal networks of Britain expanded between 1760 and 1830 from 1400 km to 3900 km (Rössner 2017). Road construction was also advanced during that time. As of 1830, the railways started in quite a disruptive manner to fill in any overland gaps. They also redefined the role of road transport, i.e. rendering it a supplier for the proverbial last mile. This particular role was to remain firmly in place until mass motorisation started to appear post-World War Two (Ede 2019).

Just as importantly, Britain had the necessary natural resources of iron ore and coal. The extreme shortage of timber accelerated the transition from an economy based on renewable resources to a fossil-fuelled one. The burgeoning population supplied the necessary labour. The country's imperial and global trading networks accelerated the essential formation of capital, fostered a *laissez-faire* liberality and, not least, enabled the procurance of external capital and raw materials in the form of cotton, precious metals and colonial produce. Out of this constellation arose a culture of progress, which ultimately gave birth to the key technologies of industrialisation

(Buchheim 1994; van Zanden 2009). The transformation could only be successful in a reciprocal dynamisation of agricultural and commercial progress together with a growing demand for products. By around 1800, the seemingly unsolvable dilemma of agricultural shortages had been solved; this was indeed a revolution—at least in a qualitative sense. The key to it was a combination of factor mobility, increased areas of cultivated land, enhanced land productivity through more efficient tools, greater crop diversity, and knowledge. Agricultural intensification was a necessary prerequisite for industrialisation. In Continental Europe, the formation of this positive set of changes did not occur until the first half of the nineteenth century.

The scope of redistribution arising from the increase in supply at lower costs helped to ward off an impending population disaster. In fact, it actually enabled urbanisation and provided an increasing stability of income, although the adaptation processes lasted around three generations and acquired a pejorative connotation through the term 'pauperism'. Although the phenomena of poverty, exploitation and child labour had also been common occurrences in the agrarian society and had actually helped to trigger the rural exodus, they became both visibly and dramatically more ostensible in the newly emerging urban-industrial society (Mathis 2015).

Moreover, work gained a new quality. The machines that were erected in a central location known as a factory decoupled work life from homelife and set the pace for the workers. Work became routinised and monotonous; working hours became longer. Twelve-hour shifts were standard.

Nightwork doubled both the demand for labour and the time available for production; a limitation was set, however by the unhealthy working environment and the disadvantaged living conditions of the workers that offered them no respite.

It is not surprising that employers, who were dependent on a sustained labour force, were already attempting in the eighteenth century to humanise what was later to become known as Manchester capitalism. Shrewd business considerations paired with social responsibility merged together to form what might—at best—be seen as philanthropic paternalism (Bonin and Thomes 2013; Roser 2016). Legal entitlement to any form of employee compensation was still a thing of the distant future. Karl Marx and Friedrich Engels quite rightly denounced the exploitation, calling in their Manifesto of the Communist Party for the proletariat to unite. That being said, industrialisation was not the cause of the social crisis but the solution to it. Incidentally, industrial work did not displace traditional handicrafts. It was more of a case of the latter redefining themselves in many ways as sub-contractors and customised service providers, particularly so because many stages of production were not yet mechanisable. Concomitant with this development, a productivity-enhancing process of specialisation took place. Handicrafts struck gold in the era of industry, too. Industrialisation did not destroy jobs at all; it mass-produced them.

The bottle neck *par excellence* of commercial production proved to be the limited availability of energy. Traditionally, regenerative energy was available in the form of biomass, sunlight, water, wind, and human or animal labour. However, it was afflicted by the well-known deficiencies with regard to availability, continuity and capacity. The solution to the problem was coal. Simultaneously, coal heralded in the transition to a fossil energy economy while the steam engine, powered by coal, became the motor of industrialisation. For the first time, not only did the steam engine—as the first heat engine—provide in its stationary form mechanically-produced—and thus effortless—energy in any individually controllable quantity; this energy could be more or less produced and used anywhere. In this way, the steam engine revolutionised the hitherto cast-in-stone location structures and had a massive impact. The steam engine marked the birth of a leading heavy industrial sector, which would comprise fossil energy production, iron production, mechanical production and toolmaking. Growing at the same rate were the typical urban areas clustered around the ore and coal deposits across the country.

Multi-functioning since 1829 as an embodiment of new technologies demanding a wide variety of products and services and as a transporting medium, the steam railway pioneered the synthesising of the individual components. Its superior efficiency completely redefined overland transport and thus the location factors. Its integrated markets opened up new economic regions and created competition. It also gave individual mobility a new quality, rendering travel available to the person on the street. Lastly, it created secure jobs on an inconceivable scale. This also applied to traditional manual labour, such as in the building industry. Furthermore, it provided work for whole regions, liberating them from poverty. Within a very short time span, it had become the largest industrial employer. With its manifold linkage effects, the railway was *the* driving force behind economic growth. At the same time, it was the spider in the web of modernisation. Evidence of its economic relevance is documented by the fact that the share of net investments in Germany's railway since the 1850s was significantly higher than 10% (Fremdling 1985). Well into the twentieth century, its disproportionate growth dynamics forged a sort of infrastructural link between the first and the second phase of industrialisation (Ede 2019; König 1997).

Iron and steel became the basic materials of industrialisation. A prerequisite was that ore could be smelted with coal in large quantities. The coal option removed the growth-constraining bottlenecks for energy and materials that were related to timber (Ede 2019; Roser 2016). Here, too, the breakthrough came at the beginning of the eighteenth century in the form of a coke-fuelled blast furnace process. Blowers powered by steam engines optimised productivity, whilst the application of Henry Cort's 1784 puddling process—likewise powered by steam—created the prerequisite for low-cost wrought iron. Thus, several bottle necks were removed. Already between 1750 and 1790, pig iron production trebled. Whereas prior to 1800 Great Britain had relied substantially on imports, mainly from Sweden, in the 1830s British producers were exporting more pig iron than the whole of Continental Europe was producing. Without British iron, early industrialisation in Continental Europe would have been slower moving. The substitution process commenced in the 1850s with the industrial 'take-off' (Rostow 1962), which had been preceded by a corresponding transfer of knowledge. Following the Congress of Vienna, the much-travelled William Cockerill, who had been born in Lancashire, built the largest single machine-building complex of the time in Seraing, near to Liège, Belgium. Below the green fields on which the complex stood lay coal and ore. Cockerill promised that any product innovation would be quickly copied at top quality—a milestone for the heavy industrialisation of Continental Europe. The earliest technological stimuli came from spinning frames and looms, which at the time were still hand operated. Textile manufacture played a pivotal role in driving industrialisation—here, too, induced by growing demand. The burgeoning population needed to be clothed. Since 1767, the start-up of mechanisation had begun with the famous Spinning Jenny. Initially hand-powered but quickly becoming a mechanical spinning wheel, it could be equipped with up to 100 spindles and needed to be manned by one person only (Paulinyi 1997). It put an end to a bottle-neck situation caused by the need for up to eight spinning wheels to supply looms that had been common since the 1730s and had flying shuttles which provided threefold weaving capacity. Power looms followed in the mid-1780s, generating—in combination with further optimised spinning wheels—yet another productivity increase. The spinning costs for one pound of thread sank between 1780 and 1830 to one twenty-fifth; the time needed dropped fifteenfold. The productivity effects corresponded to those of the railway.

A further aspect is significant here: mechanical manufacturing was initially only possible with finely grained cotton, which needed to be imported, marking the beginning of a globally efficient supply chain in a mass market. Without it, the manufacture of cotton would not have advanced to be the pacemaker of industrialisation. Simultaneously, its mechanisation created impulses in sheep farming with regard to the quality of wool: a further example of the trans-sectoral and international linkage effects stemming from industrialisation innovations.

In the wake of the Napoleonic Wars, the diffusion process was fostered by the liberalism that had been consciously initiated in a peaceful, balanced Europe by the Congress of Vienna.

The development culminated in 1861 in the ground-breaking Cobden-Chevalier Treaty in which France and Great Britain granted each other most-favoured-nation status, thus setting off a free-trade chain reaction across Europe. In parallel, gold became the basis of the international payments systems. The so-called gold standard, for which David Hume had laid the foundation in the eighteenth century with his price-specie flow mechanism theory, significantly lowered transaction costs—inclusive of positive growth effects. Until 1914, the gold standard provided a reliable basis for trade.

This framework enabled Great Britain to initially supply products and capital to a Continental European market which, in the aftermath of the Napoleonic Wars, was lagging behind and needed to catch up. This was an important precondition for Britain to secure itself the role of 'Workshop of the World' and that of a dominant economic power.

As all the foundations for industrialisation in the whole of Western and Central Europe had been laid by the start of the 1860s, it started to gather speed. Via the 'take-off' phase of W. W. Rostow (Rostow 1962), accelerated growth culminated in self-sustaining dynamic growth. In other words, growth became institutionalised and peak industrialisation ensued. Thus, the regenerative era ended. At the time, very few people gave thought to the adverse environmental impacts. Although emissions were intense and perceived as oppressive, people tolerated them.

The setting outlined here essentially applies for all processes of industrialisation. It differs only in the values of individual parameters and the corresponding rate of change, which increased due to the advantages of the catching-up economies. Here, the concept of leapfrogging (Mathis 2015) comes to mind with its steep learning curve and optimised change management.

2.2 Industry 2.0 (the Second Industrial Revolution)

Until the outbreak of World War One in 1914, the mechanisms outlined in the previous section remained uninterruptedly dynamic, even gaining speed yet again. Mechanised economic activity was the defining socio-economic characteristic. An affinity for progress and a growth ideology took a firm hold, as did a culture of entrepreneurship and efficiency. By 1914, Great Britain, the USA and Germany found themselves at roughly the same stage of industrialisation (Buchheim 1994; Roser 2016). The latter two countries had passed through Britain's development in record time. But not just that: they had mutated from imitators to innovators. It was in the USA and Germany where the new leading industries of electrotechnology, chemicals, and mechanical and automotive engineering were forming (Ede 2019). The term 'peak industrialisation' (German: *Hochindustrialisierung*) describes the processes rather well. Great emphasis was placed on scientific education. Humanistic content was extended by the practicality of the *Realschule* school form and by polytechnics, particularly in Germany. The larger industrial companies soon all had their own research and development departments where—almost like a factory manufacturing process—theoretical knowledge was transposed into applied knowledge. Here we have a symbiosis of knowledge and technology. Germany played a dominant role in these transfer processes, as evidenced by numerous Nobel Prize awards. German became the language of science, too. In 1911, Taylor coined the term of 'scientific management'. Rationalisation became synonymous with efficiency (König 1997; Roser 2016).

The heavy engineering industry and the railway were able to consolidate their strong position of productivity drivers because they continued to improve their quality and performance. This was the only way, for instance, that the North American frontier was able to be closed by 1890 or that from the Asian coast opposite, the Trans-Siberian Railway was able to lead as of 1916 directly to Europe. Around the world, the rail network grew between 1830 and 1910 to encompass one million kilometres, only to accelerate again over the next 20 years, adding a further 500,000 km. The linkage effects that we touched on earlier remain just as relevant and can hardly be overestimated.

With regard to materials, new technologies—such as the Thomas-Gilchrist process—accelerated the mass production of steel. Productivity increased twenty-five-fold. By-products, such as gases and fertiliser intensified the multiplier effects. In 1913, there were more than 400,000 workers engaged in the steel industry of Wilhelmine Germany. The metal processing industry was expanding even faster.

Steel engineering opened up completely new structural dimensions. An impressive example was the Eiffel Tower—a 300-metre-high wrought-iron construction created in 1889. No less ground-breaking was the combination of steel and concrete, patented in France in 1867. In the building industry—hitherto characterised by craftsmanship—workers increased threefold. This increase was as rapid as that within the mining industry, which was driven by the practically exploding need for energy. Here, the number of workers increased threefold between 1873 and 1913; production increased fivefold and productivity by almost 70%. Germany's Ruhr mining area alone employed 450,000 individuals (Henning 2008).

It is a fact that the optimised technologies of the First Industrial Revolution were still dominating the world's industrial economy well into the twentieth century. Complementary to these were the new leading industries of electrotechnology and chemicals as well as mechanical and plant engineering with their own fully disruptive impact—the full extent of which would be a feature of the entire twentieth century. These industries changed the technology of production, they impacted on other sectors, and they dynamised each other—as demonstrated by the emergence of the electrochemical industry.

The highly compact electric engine together with the Otto and the Diesel engine expanded the diversity of power drive concepts to a significant degree (Kirsch 2000; Mom 2004). Each of them began to successively displace the steam engine. With regard to automobile manufacture, they totally redefined mobility and location. Not only did it lead to a 'renaissance of the road', it created individual mechanical automobility and led to the opening-up of airspace. The vehicles of Daimler and Benz made their appearance in 1885/86. However, the pioneer of application was the USA once more, where Ford's standardised mass production, made the breakthrough. This paved the way for mass automobilism, enabling the automotive industry to become the most important industrial sector after World War Two for practically all of the world's developed economies (Thomes 2003).

In 1903 in the USA, the Wright brothers successfully made the first steered aircraft flight. In 1919, civil aviation began in Germany. The alternative concept of 'lighter than air' in the form of a Zeppelin airship after a promising start proved to be a failure (Ede 2019; König 1997).

With regard to energetic aspects, motorisation put a new perspective on oil and gas as energy sources. They took the pressure off coal as an energy source which, like oil and gas, meanwhile had acquired a central role in the chemicals industry—an industry that had been growing into a major industry since mid-nineteenth century. The foundation was formed by the unique accumulation of systematic knowledge to date. In 1914 in Leverkusen, the Bayer company (established in 1863) had about 600 chemists among its employees. The key innovation was the synthetic production of natural substances in any quantity and quality. On this basis, completely new applications opened up, affecting all economic and social areas in an often revolutionary manner. The spectrum ranged across aniline dyes, synthetic fibres and materials, photo technology and film technology, detergents, artificial fertilisers, pesticides and pharmaceutical products. The latter, interacting with the equally rapid medical progress being made, revolutionised the fields of health and hygiene through antiseptic or pain management in the form of the ‘all-purpose’ drug Aspirin, which Bayer brought to market at the turn of the century.

Metallurgy could no longer manage without systematic chemical analysis and products either. The coal industry—and later the petrochemical industry—established their own research and business fields, as did medical technology. In around 1900, the German chemical industry was dominating the world market. In 1913, 28% of the world’s chemical exports stemmed from German companies. Established in 1865, the Badische Anilin und Sodafabrik company to this day continues to incorporate the innovative products of its early years in its name: BASF. All this activity had resulted by 1913 in a significant increase in the quality of life and life expectancy together with favourable living costs, factors which themselves resulted in demand and linkage effects (Braun and Kaiser 1997; König 1997).

The industry to demonstrate the most rapid growth was the electricity industry in the sectors of lighting, power and heating; its effects were in no way inferior to those of the chemical industry. The first viable battery, invented by Volta in about 1800, and the discovery of electromagnetism quickly led to practical applications of this completely new form of energy. For instance, in 1838, the first boat with an electric motor was launched and—not much later—Morse brought the disruptive information technology of telegraphy to practical maturity. After the dynamo had started to revolutionise electricity generation in 1866, product innovations followed one after the other: the telephone in 1871, electric bulb and the first electric tramway in Berlin in 1879, the transformer in 1882, long-distance electric power transmission lines in 1891, wireless information transmission in 1896, cathode ray tubes in 1897, vacuum tubes for radio in 1905, etc. Edison and Siemens stand exemplarily as inventors and entrepreneurs (Ede 2019; Feldenkirchen 1997; Roser 2016).

Already in 1908, wireless radio communications had started up between Europe and the USA. In parallel, the world of 1914 was linked together by more than 500,000 km of cable—often near the railway lines. Electricity was a game changer in this field, too, dematerialising the transmission of information and thus enabling global real-time communication and information, as it were, as well as the further networked integration of time, space and markets. The electric motor inspired tool manufacturing (turning, milling, drilling) and the mechanisation of handwork and agriculture. It combined several advantages: compact design, robustness, affordability, low noise level, energy efficiency and local zero emissions. However, it was not able to succeed as an automotive drive (Thomes 2018).

Water and steam turbines were further innovations that filled the supply needs. In 1900, one turbine produced 1000 kilowatts; in 1916 it was 50,000 (König 1997). The world was being electrified at full throttle and, as a logical consequence, the electronic age was about to be ushered in. Although the renewable generation potential of electricity did not play a decisive role in those days, its universality did (Chandler 2001; Roser 2016).

With regard to production and organisational aspects, custom-made production became almost fully substituted by market production in those years. Factories—places where the standardisation of parts and the division of labour facilitated efficient mass production—took on new dimensions. Already by 1900, 66% of manufacturing workers were employed in factories.

Krupp—the German coal and steel giant and plant constructor—employed more than 80,000 people in 1913, whereas the figure had been only 21,000 in 1887. Siemens also had about 80,000 employees, 24,000 of whom were working outside Germany—reflecting the company's position as one of the earlier global players (Feldenkirchen 1997). US Steel, founded in 1901 as the result of a merger by J.P. Morgan, was the world's largest and in terms of capital the first billion-dollar corporation. These corporations increasingly adopted vertical and horizontal organisational structures to cover the whole supply chain and product spectrum by dominating the market (König 1997; Roser 2016). The fast rate of growth could only be realised in the form of publicly traded (joint stock) companies.

Because of productivity, the scope for distribution of wealth increased, as did the population and purchasing power. Despite rising real wages, the working hours fell. In Germany Bismarck's social reforms during the 1880s established the work force's statutory right to social insurance. In principle, these reforms defined the global standard industrial employment relationship that is still relevant today.

Consequently, life expectancy leapt from 38 to 50 years and the population grew from 41 to 68 million between 1871 and 1913. However, 3 million people emigrated during this period, mostly to the USA, due to a lack of prospects. Urbanisation accelerated. Cities and conurbations mutated into highly complex technical systems (König 1997; Mathis 2015). The entire know-how of the age was incorporated into the ocean liners, yet the sinking of the 'unsinkable' Titanic in 1912, which resulted in almost 1500 deaths, evidenced the vulnerability of that know-how (Weber 1997).

In barbaric fashion, the two World Wars manifested the dark side of technology in all its destructive potential. The interim years between the two wars spawned the term 'technocracy' and, after a brief interlude of consumption frenzy, destroyed in the form of the first industrial world economic crisis the belief that progress is immune to any setback. More than that: democratic liberalism quickly turned into dictatorial coercion. Whether in warfare or in the elimination of unfavoured minorities, the Nazi terror regime had no scruples about instrumentalising technological against humanity in a hitherto unprecedented way. The inhuman Holocaust bears witness. It was the dropping of the atomic bomb—democracy's final devastating heave against Japan's autocratic monarchy—which horrifically put an end to the horrors of a war in which some 60 million people lost their lives (Ede 2019; Roser 2016; Tooze 2006).

After the world had paused for a moment's breath, it returned more quickly and indiscriminately than ever to its technological growth path. The production technology—militarily optimised during the war through the compulsion for efficiency—became 're-civilised'. It laid the first steppingstone on the proven path toward normalisation—a 'miraculous' process of making up for lost time. The world economic crisis was forgotten, perhaps because the now peaceful usage of atomic energy appeared to have solved the energy and growth problems for ever (Braun and Kaiser 1997; Wehler 2008). It was the stagnating years of the 1970s, with their high numbers of unemployed and rising energy prices, that first brought to light both the limitations of the chosen technology and growth pathway and the need for a change of direction (Meadows 1973; Wehler 2008).

2.3 Industry 3.0 (the Third Industrial Revolution)

We specifically begin by looking at the seventeenth century to refresh our awareness of the long-term nature of developments, even in terms of digital technologies. In 1673, G. W. Leibniz demonstrated his ingenious calculating machine before London's scientific academy, the Royal Society. His objectives are best reflected in the following citation which is attributed to him: 'For it is unworthy of excellent men to lose hours like slaves in the labour of calculation which would safely be relegated to anyone else if machines were used' (https://math.dartmouth.edu/~m3cod/LeibnizWheelBig.htm). The phrase symbolically sums up the general goals of automatisation. Consequently, Leibniz—around the turn of the eighteenth century—invented the binary arithmetic system (*l'Arithmétique Binaire)*, creating the basis for digital technology. Further earlier milestones were the Jacquard looms, which had been operated by punch cards since 1808, or the 'Analytical Engine' designed by Babbage in 1837 (Bruderer 2015; Roser 2016; Troitzsch 1997).

The momentum took off when statistics as well as administrative and managerial processes, which had been growing exponentially to the population and size of companies, could no longer be effectively handled in analogue form. In 1890, Hollerith's punched-card-based electric sorting and tabulating machine significantly shortened the processing of the U.S. census. The tech giant IBM was to emerge out of the same company. In 1906, the first fully automatic electric calculator—Autarith—was presented in the USA. In 1938, the German pioneer Zuse notably called his Z1 computer the 'Mechanical Brain'. In 1951, the universal automatic calculator UNIVAC1 marked the birth of commercial electronic data processing computers, (Chandler 2001; Braun and Kaiser 1997).

Already in 1950, the Turing test was opening the doors to artificial intelligence. In 1953, the University of Cambridge offered the world's first taught course in computer science; the USA followed in 1962; the Federal Republic of Germany in 1967 with 'informatics'; the German Democratic Republic two years later. The importance of education moved even further to the forefront and universities were required to open themselves up to society.

It was not long before industrial applications began to appear. The first programmable manipulator was patented in 1954. A short time later, removal robots, welding robots and machine tools followed. For the first time, numerically controlled tasks could be performed effortlessly and efficiently: a highly appealing phenomenon considering the manpower shortage at the time. In 1966, work began on developing a mobile robot. In 1969, the electronic calculator made its appearance on the market; mankind flew to the moon on computing power and the CAD-constructed jumbo jet Boeing 747 started its take-off into a new age of aviation, remaining to this day an impressive mix of analogue and digital high-tech. Intel introduced the first microprocessor in 1971. In 1976, Apple 1—the first personal computer—democratised digital technology (Ichbiah 2005; Roser 2016).

There seemed to be no limits to digital efficiency. The so-called Moore's law of 1965 observed that the number of transistors in a defined integrated circuit doubles every two years. Effects of scale and leaps in development allowed costs to drop dramatically and performance to increase. Digitalisation became unstoppable. Information technology became standard in organisations.

What was lacking in the meantime was the networking of stationary applications. In 1969, ARPANET was established. In 1991, the first website appeared (Castells 2001). No more than 30 years later, more than 4 billion people now use the Internet. Whereas the telephone needed 75 years to connect about 100 million users, the Internet managed it in only seven years. In the last 50 years, digitalisation has completely redefined production and consumption by informatising value creation (Manzei et al. 2016). Early pioneers had been adversely affected by this phenomenon ever since the 1960s, not because the products themselves were no longer called for but because they could be produced much more cost effectively in other countries.

Simple analogous innovations such as the pallet and the container in combination with efficient logistics demolished the costs of transport (Levinson 2008; Preuß 2010). Those productions which remained where they were on account of the know-how involved nevertheless rationalised. The production factor 'labour' found itself in a stranglehold, as it were, and escaped from a central European view via Southeastern Europe and then to developing Asian economies. In Asia meanwhile Japan advanced to become the global high-tech champion. Its superiority within the magic triangle of price, performance and precision led the augurs to predict Japanese world economic dominance for 2000.

Industrial production nowadays would in fact be inconceivable without just in time manufacturing, the kaizen concept, etc. (Chandler 2001; Braun and Kaiser 1997). Amongst other countries, South Korea and China followed suit. Since the 1980s they have—as production locations and markets—gone through a remarkable development to become innovators. Theoretically, this was not possible with a rigidly organised authoritarian system such as that of the People's Republic of China. The key has been that country's self-prescribed ability to adapt. A case of déjà vu that has us sitting up and taking notice. Is history repeating itself?

3 Taking Stock and Looking Ahead: Back to the Future?

Doubtlessly, there are many other aspects that could have been and perhaps should have been addressed here. However, what has become clear is that industry should not be understood in isolation as a commercial processing of raw materials into goods. Industry is much rather a reciprocal interaction between several parameters (STEPLE). Networking and intensification are fundamental continua in all phases. Thus, this has led, particularly in the last 200 years, to an unparalleled socio-technological transformation. One positive aspect of this process: every minute, about 100 individuals break free of the poverty spiral in Africa or Asia. Never have so many people lived in freedom and prosperity and achieved longevity (Norberg 2016).

It was in Great Britain in the eighteenth century that all the relevant factors for a productive break with archaic feudal agricultural structures initially came together: natural resources, manpower, capital, demand, rationality, knowledge, entrepreneurial spirit, mobility of goods and people, international networking.

It was that initial technological impulse—the steam engine—that was able to perform for the first time controllable mechanical work in any location. In railway operation, it was able to open up markets in a transport-efficient manner, creating multiple disruptive linkage effects in all commercial fields. The steam engine also gave new significance to the value of proto-industrial innovations (manufacturing, division of labour, standardisation of parts and processes). Further integral aspects were the transition to a fossil energy basis and to new raw materials beyond timber, i.e. iron and steel. Production and consumption liberated themselves from any natural limitations. In a sort of collective incubator, supply and demand fired each on to ever-more dynamism, their spark jumping over to Continental Europe and North America thanks to a liberal foreign trade regime. Elements of revolution and evolution merged to produce something entirely new.

The electro-technological industry and the chemicals industry were the next impulses; their applications even more disruptive. They propelled the existing leading industries onwards, just as they dynamised each other. The automobile—a further basis innovation—would not have run without electricity, lubricants, rubber, etc. At the same time, electrified communications were connecting the world in real time via wire or 'wireless'.

The arrival of the industrial laboratory dismantled further limitations and allowed mankind to feel like an omnipotent creator. Artificial fertiliser, plastics, synthetic dyes and large-scale industrial systems are just a handful of the achievements. In so far, referring to this period as 'Industry 2.0' is justifiable, whereby here again, we see evolutionary elements merging with revolutionary ones.

Although the 3.0 stage was a logical consequence. it was also completely novel and elevated mankind to a new level, the digital age. No longer was it necessary for human beings to operate machines; they programmed algorithms that did the work for them. Machines control machines and—as a next step—they are starting to take on human form. We humans are beginning to make optimised duplicates of ourselves. Technology is no longer merely replacing human strength; more than ever, it is replacing the human brain—as anticipated by Zuse's 'Mechanical Brain'.

And now, since the Internet of Things and Services is concerning itself with a comprehensive virtual network of machines, products, people and requirements in cyber-physical systems, a radical change is once again imminent: Industry 4.0. Some notable aspects of our long-term comparison demonstrate the need for explicit political action. Society is going back to the future particularly in the energy and resources sectors, having succumbed to the temptation of productivity for far too long without questioning it. There is certainly no alternative to finding regenerative and sustainable solutions, either in production or in consumption. One such solution might be that of redefining ‘subsistence’.

Moreover, society is currently moving towards digital capitalism, with individuals surrendering their hard-won ownership rights. The maxim of ‘using rather than owning’—which is admirable from a sustainability perspective—might lead to the coming of neo-feudalism, where the feudal lords would be global corporations that—up to now—have not really been controllable. Adaptive production—in the form of mass customisation—is bringing the original principle of ‘customised production’ to the fore again, the watchword being ‘lot size 1’. If we think this further, mass customisation has the potential to make consumers into autonomous prosumers who are perhaps organised into cooperatives and generating their own renewable energy through decentralised networks.

The role of human beings as the production factor ‘labour’ is a problematic one. From a production perspective, democratic consensus on prosperity made labour more expensive during the last century, while robotisation is increasingly replacing not just simple but qualified human work. The industrial standard employment contract is gradually disappearing. The number of fixed-term contracts are increasing exponentially. From a historical perspective, on-demand workers are a replication of the day labourers in the feudal system or of the proto-industrial cottage workers. Employing such workers always involves an externalisation of costs. Even highly qualified research and development can be performed in low-wage locations through digital dialogue. Meanwhile, the tertiary sector has disappointingly become a low-wage sector (Fourastié 1949). Even here, androids are crowding out humans. It appears that a downward spiral of precarity is gaining momentum, leading us back in many ways to a pre-industrial era. If this is the case, then the periodisation of Industry 4.0 should be that of a post-industrial age.

On the other hand, humans are gaining value as data producers. Data are becoming a new raw material. Data mining excavates preferences. Work is dematerialising itself. Up to now, the provision of data information by making information services available has been inadequately remunerated. A digital tax or a value-added tax scheme, for instance, would be a way of globally securing individual income, particularly as the wage share of the national income is already declining.

More than ever, it is imperative that we act globally in a considered and concerted way rather than lapsing into neo-mercantilism. Furthermore, Industry 4.0 as a sustainable system requires global democratic institutions and binding global regulations. These must be based on inclusive capitalism in the sense of peaceful coexistence and global convergence (Schmidt et al. 2016), adhering to the motto of ‘Put the world first’. Almost 250 years after Adam Smith, mankind needs a new concept, one that is in line with the (freely adapted) call of the French Revolution: Liberty, Equality, Humanity. Theoretically, all that is necessary are the appropriate algorithms.

References

Berg M, Hudson P (1992) Rehabilitating the industrial revolution. Econ Hist Rev 45(1):24–50
Bonin H, Thomes P (eds) (2013) Old paternalism, new paternalism, post-paternalism (19th-21st centuries). Peter Lang, Brussels
Braun HF, Kaiser W (1997) Energiewirtschaft Automatisierung, Information seit 1914 (Propyläen Technikgeschichte 5). Propyläen, Berlin
Bruderer H (2015) Meilensteine der Rechentechnik. Zur Geschichte der Mathematik und der Informatik. De Gruyter, Berlin
Buchheim C (1994) Industrielle Revolutionen. dtv, München
Burke P (2000/2012) A social history of knowledge 2 vols. Polity Press, Cambridge
Castells M (2001) Die Internet Galaxie. Internet, Wirtschaft und Gesellschaft. Springer, Berlin
Chandler AD Jr (2001) Inventing the electronic century. Harvard University Press, Cambridge
Dean P (1982) The first industrial revolution, 2nd edn. University Press, Cambridge
Ede A (2019) Technology and society. A world history. University Press, Cambridge
Feldenkirchen W (1997) Siemens. Von der Werkstatt zum Weltunternehmen. Piper, München
Florida R (2010) The great reset. How new ways of living and working drive post-crash prosperity. Harper Collins, New York
Fourastié J (1949) Le Grand Espoir du XXe siècle. Progrès technique, progrès économique, progrès social. Presses Universitaires de France, Paris
Fremdling R (1985) Industrialisierung und Eisenbahn. Zug der Zeit – Zeit der Züge. Deutsche Eisenbahn 1835-1985 1: 120–133. Siedler, Berlin
Grunwald A (2018) Der unterlegene Mensch. Die Zukunft der Menschen im Angesicht von Algorithmen, künstlicher Intelligenz und Robotern. Riva, München
Henning F-W (2008) Die Industrialisierung in Deutschland 1800 bis 1914. 9. A. UTB, Stuttgart
Hessler M, Thorade N (2019) Die Vierteilung der Vergangenheit. Eine Kritik des Begriffs Industrie 4.0. Sozialgeschichte 86:153–170
https://math.dartmouth.edu/~m3cod/LeibnizWheelBig.htm, retrieved on 08.01.2021
Ichbiah D (2005) Roboter. Geschichte _Technik _Entwicklung. Knesebeck, München
Kagermann H, Lukas W-D, Wahlster W (2011) Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution. VDI-Nachrichten. 4 2011: 13
Kirsch DA (2000) The electric vehicle and the burden of history. Rutgers University Press, New Brunswick et al
König W (1997) Netzwerke, Stahl und Strom 1840 bis 1914 (Propyläen Technikgeschichte 4). Propyläen, Berlin
Landes DS (1969) The Unbound Prometheus. Technological change and industrial development in Western Europe from 1750 to the present. University Press, Cambridge
Levinson M (2008) The Box – how the Shipping Container made the world smaller and the economy bigger. Princeton University Press
Lübbecke M (2015). https://mluebbecke.wordpress.com/2015/12/16/industrie-5-0/
Manzei C et al (eds) (2016) Industrie 4.0 im internationalen Kontext. Kernkonzepte, Ergebnisse. Trends. VDE, Berlin
Mathis F (2015) Mit der Großstadt aus der Armut. Industrialisierung im globalen Vergleich. University Press, Innsbruck
Meadows D (1973) Die Grenzen des Wachstums. Bericht des Club of Rome zur Lage der Menschheit. Rowohlt, Hamburg
Mom G (2004) The electric vehicle. Technology and expectations in the automobile age. The John Hopkins University Press, Baltimore u. a.
Norberg J (2016) Progress. Ten reasons to look forward to the future. One World Publications, London
Paulinyi A (1997) Mechanisierung und Maschinisierung 1600 bis 1840 (Propyläen Technikgeschichte 3). Propyläen, Berlin

Peters R, Thomes P (2018) Ein Modell zur strukturierten Analyse von Veränderung als Plädoyer für eine integrierte Vergangenheits-, Gegenwarts- und Zukunftsperspektive. Scripta Mercaturae 47: 1–30
Pollard S (1981) Peaceful conquest, the industrialization of Europe 1760- 1970. University Press, Oxford
Preuß O (2010) Eine Kiste erobert die Welt. Der Siegeszug einer einfachen Erfindung. Murmann, Hamburg
Radkau J (2017) Geschichte der Zukunft. Prognosen, Visionen, Irrungen in Deutschland von 1945 bis heute. Hanser, München
Roser C (2016) “Faster, better, cheaper” in the history of manufacturing: from the stone age to lean manufacturing and beyond. CRC Press, Boca Raton
Rössner PR (2017) Wirtschaftsgeschichte neu denken. Mit einer Darstellung der Ursprünge moderner ökonomischer Theorien. Schäffer-Poeschl, Stuttgart
Rostow WW (1962) The stages of economic growth. University Press, Cambridge
Schmidt RH, Seibel HD, Thomes P (2016) From microfinance to inclusive banking. Why local banking works. Wiley, Weinheim
Thomes P (2003) Searching for identity. Ford Motor Company in the German market. In: Bonin H et al (eds) Ford, 1903–2003: The European History 2: 151–195. P.L.A.G.E., Paris
Thomes P (2018) Elektromobilität –eine historisch basierte Analyse. In: Kampker A et al (Hrsg) Elektromobilität. Grundlagen einer Zukunftstechnologie, 8–17. Springer, Berlin
Timm A (1964) Kleine Geschichte der Technologie. Kohlhammer, Stuttgart
Tooze A (2006) Wages of destruction: the making and breaking of the Nazi economy. Penguin, London
Troitzsch U (1997) Mechanisierung und Maschinisierung 1600 bis 1840 (Propyläen Technikgeschichte 3). Propyläen, Berlin
van der Pot JHJ (1999) Sinndeutung und Periodisierung in der Geschichte. Brill, Leiden
van Zanden JL (2009) The long road to industrial revolution. The European economy in a global perspective. Brill, Leiden
Weber W (1997) Netzwerke, Stahl und Strom 1840 bis 1914 (Propyläen Technikgeschichte 4). Propyläen, Berlin
Wehler H-U (2008) Deutsche Gesellschaftsgeschichte, 5. Beck, München
Wolff D, Göbel R (Hrsg) (2018) Digitalisierung Segen oder Fluch? Wie die Digitalisierung unsere Lebens- und Arbeitswelt verändert. Springer, Berlin
Zinn KG (2018) Malthus, Marx und Keynes zur Übervölkerungsthese. Ein Essay zur Aktualität theoriegeschichtlicher Hinterlassenschaft. In: Gischer H et al (eds) Bewegungsgesetze des Kapitalismus. Fs. für F. Helmedag. Metropolis, Marburg, pp 251–269

Ethics of Digitalization in Healthcare

Arne Manzeschke and Alexander Brink

Contents

1 Ethical Issues in Healthcare:[1] An Introduction

Health is, for most of us, one of the most important things, if not the most important: "The main thing is health!" If health is regarded as a "conditional good" (Kersting 2012), i.e., a condition for leading one's own, self-determined life, ethical questions are not far away—especially if institutional frameworks are essential for health or healthcare (Callahan 1998; Gruskin et al. 2005), and the expectation of the

[1] We focus on the German Healthcare System a lot. Thus, we have referred to so many German literature. Although Ethics should go for more or less universal statements, it nonetheless must consider the very specific situations where moral questions are at stake.

A. Manzeschke (✉)
Lutheran University of Applied Sciences - Evangelische Hochschule Nürnberg, Nürnberg, Germany
e-mail: arne.manzeschke@evhn.de

A. Brink
Bayreuth University, Bayreuth, Germany
e-mail: alexander.brink@uni-bayreuth.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_59

population for a technology-based healthcare system are enormously high (e.g., Reiser 2009, 1978; Savulescu 2007; Niederlag et al. 2008).

Considering that, in economic theories, the healthcare market is seen as a major driver of innovation and growth, and that needs and basic innovations in the field of healthcare, wellness and holistic lifestyle form the basis for a sixth Kondratieff cycle (Nefiodow and Nefiodow 2014), which once again underlines the importance of this sector. Health-related data collected digitally, such as vital parameters, mobility, or consumer data, is a lucrative business model for health insurance companies and other organizations (Waschinski 2019).

Ambiguities must be taken seriously: the sensitivity of personal health data and the chances of "personalized medicine" based on this data (Niederlag et al. 2010a, b), questions about the support of nursing activities through robotics, the increased nursing care rate in the face of a decreasing number of skilled workers (due to **demographic changes**) and the support of medical treatment decisions through artificial intelligence (AI) all reinforce the impression that the need for ethical orientation in this field increases along with the technical possibilities. As the British philosopher Stephen Toulmin put it: "Modern medicine has saved ethic's life" (Toulmin 1982). Modern medicine is associated with modern technology, and it is not only technical possibilities that have opened up new treatment options and made decisions necessary where "fate" previously ruled. The existential dimension of the topic of health and its economic significance make a substantial contribution to the fact that digitalization and the networking of technical systems are making inroads here, while traditional forms of interaction, hierarchies and moral orientations are being questioned.

In the following, we will illustrate these developments with a few examples, and interweave ethical reflections that owe a great deal to our fundamental considerations (see also Manzeschke and Brink "Ethics of Digitalization in Industry").

The healthcare system is a social institution that is connected to moral demands, because an individual and a society's way of life depends on the rules and presuppositions that apply here. These assumptions include ideas about what and how a person is or has to be, or what is considered "healthy," "sick," "disabled," or "restricted." To help people in need is generally regarded as morally necessary, which is why help in the health sector is well understood, morally. In a modern, pluralistic, and differentiated society, which measures should be given or accepted as help are disputable (the current controversy about compulsory vaccination) and how seemingly scarce resources should be distributed. Questions of **justice** (justice of distribution, participation, and performance justice) are thus called for. Technology, and not least digital technology, appears to be an effective means of allocating resources more efficiently, using effective means of diagnostics and therapy and increasing the well-being of the individual. Digital technology in the healthcare sector is encountered in simple, stand-alone solutions, such as a drinking quantity measuring devices for individual users to complex digital worlds in operating theatres, with augmented, virtual reality and extensive robotic systems that can be operated tele-medically over long distances. Digital technology is also the basis of modern stem cell research and genetics, which could not be done without high-performance computers and the corresponding algorithms. This is not the place for detailing this. Rather, our aim is to use select topics to illustrate fundamental ethical questions. The contribution is based on the conviction that, at the current stage of discussion, the first thing is to ask the questions more precisely. However, this is not without a certain normative orientation, which we have already pointed out in our fundamental article (see also Manzeschke and Brink "Ethics of Digitalization in Industry").

The former EU Commissioner for Competition, Neelie Kroes, was keen at pointing out where a starting point lies for ethical questions in the digitalization of the healthcare system:

> "In the digital age, we face a dilemma. We know that technology must move forward and that we enjoy the benefits and productivity of new devices and services every day, but questions remain in our hearts." She continues, "ICT and telecommunications have their own worth as industries, but their real significance lies in their contribution to the rest of the economy. As the source of more than half of the growth in productivity, they ensure competitiveness and ensure that the internet economy contributes more to growth today than any other industry." (Kroes 2015, p. 96, 98)

The dilemma that she details in an article for the Frankfurter Allgemeine Zeitung (FAZ)—part of an ongoing debate in that paper—is that, in her view, we citizens of the European Union do not want to and cannot do without the benefits of digital technologies, which create jobs and value. At the same time, however, we would be massively irritated by the use of these technologies in our moral orientation, and in what is often called the "European order of values" (European Union 2009; Bindé 2004). The questions related to "in our hearts" are to be addressed in this article, with regard to healthcare. These questions are, among other things, attached to health-related functions. Frank Schirrmacher, then editor of the Frankfurter Allgemeine Zeitung, criticized Kroes for her digital agenda:

> One of the most powerful women in Europe addresses the audience in the style of a New Year's speech, but it is not what she says that is electrifying, rather what she does: after a few seconds of her speech about 'health in the wallet', she points to her wrist, on which she is wearing one of the new electric bracelets that measure movement, fitness and other physical functions. Almost exclusively, she talks about the undeniable advantages of such a bracelet in times of demographic ageing and defective health care systems. [. . .]. One does not even think that politics is not limited to the description and use of a control system – there are engineers for that – but must ask and answer questions about social consequences. (Schirrmacher 2015, p. 62)

If one regards politics as the space in which questions about social consequences must be asked and answered, then political debate about ethical reflection cannot be avoided, since all political actors—consciously or unconsciously—connect moral attitudes with their interests. In the field of health policy, Gernot Böhme's sentence on ethics as the reflection of serious moral questions is immediately obvious: "A moral question in the area of ethics concerned with the formation of an individual mode of living is a question by which it is decided how a person regards himself or herself, and who that person is; a moral question in the field of the public discourse devoted to establishing social norms is a question by which a society regards itself and what it becomes. In each case these are questions in which matters become serious for the individual person or for the society." (Böhme 2001, p. 9)

In the following, it is a matter of pursuing these questions, "with which it becomes serious for everyone at some point" (ibid.). This applies to questions of health and illness, which always have something to do with one's own vulnerability and—in the last analysis—one's own mortality. The subject of health is particularly suitable for offering technology against these "offenses" of human existence, and to connect them with promises (such as overcoming pain and illness, and even prolonging life) (Kirkwood 1977; Kurzweil 2005). The technical potential to continuously improve states or processes now extends to man himself, as the material of this bio-technical access, with the option of leaving the old bio-corporeal man behind (De Carolis 2004; Weiß 2009; Braidotti 2013; Spreen et al. 2018). In this article, we will limit ourselves to some topics from the field of digitalization, artificial intelligence, and robotics, in terms of their recognizable and expected effects, from an ethical-anthropological perspective.

After this brief introduction to a dilemma which is becoming increasingly virulent in the digital age, especially with regard to health issues, developments will be addressed based on **datafication** (Cukier and Mayer-Schöenberger 2013), **dissolution of sectoral boundaries**, and **surveillance** that are of particular relevance to the healthcare system from an ethical perspective. On health and the specific constitution of the human being (**vulnerability**, corporeality, sociality), ethically sensitive points are touched upon here, which must be considered in the research, development, distribution, and use of digital health technologies (Sect. 2). This consideration is rounded off by considerations regarding the use of robots (Sect. 3) and assistance systems in medicine and nursing (Sect. 4), because they can be used to illustrate the transcendental character of this technology, which must also lead to an ethical-anthropological redefinition, insofar as this technology goes beyond the purely instrumental character of previous tools, machines, and automats. A short conclusion follows these considerations (Sect. 5).

2 Data Validation, Sectoral Delimitation and Surveillance

2.1 The Importance of Data in Healthcare

In the healthcare system, data appears to be the basis of meaningful diagnosis and therapy. Even if data, information, and knowledge can be epistemologically distinguished even more precisely, the idea is widely held that the more data is available, the more precisely individual and social processes can be analyzed, understood, controlled, and managed (Bartens 2018 is critical of this). Big Data applications—despite the vagueness of the term (Kolany-Raiser et al. 2018)—mean that increasing health-related phenomena are measured and, in almost real time, processed and applied. This applies to the differential diagnosis of the individual patient with non-specific upper abdominal pain in the family doctor's practice, as well as the victim of a traffic accident, for whom data transfer from the ambulance can, under certain circumstances, be lifesaving. This also applies to epidemiological studies on the spread of a wave of influenza, or the analysis of which chemotherapeutic agent can be effectively used for which patient group, due to their genetic disposition. In all cases, data forms the basis for health-related decisions and interventions. The more data is available in digital form, the easier it can be exchanged and linked with other data and processed to further information (Weichert 2018).

Digitalized and data-based healthcare is driven by high expectations for the alleviation of individual suffering and global epidemiological challenges. On the other hand, there is a skeptical view that the collection, possession, and further use of data is associated with considerable intrusions into the privacy of the individual. For the health of the individual, far-reaching—and in some cases justifiable—forms of surveillance can be installed, which are highly problematic for society as a whole (Lanier 2014; Hofstetter 2014; Grafenstein et al. 2018). Furthermore, it is possible that the objective impression of technical data will have an impact on the individual's self-image, and relationship to themselves and others. However, it seems very serious for a correlative evaluation of an individual to take place through this data, marking their status in society, and thus defining their chance of realization, and ultimately tilting the idea of a human dignity that cannot be lost or offset in favor of algorithmic calculations or highly significant statements (Grafenstein et al. 2018). These points concern us all, because they are important for all of us now (and not at some point in the future) (Grunwald 2019). The concerns articulated here apply to all the undeniable advantages and opportunities of digital technology in the healthcare system.

2.2 Data Validation and Big Data in Healthcare

In the measurement and digitization of increased expressions of life and social executions, as well as the networking of the corresponding devices, very large data sets are being built up, in which algorithms are used to search for patterns of known variables (a certain tumor, a genetic marker or a correlation of a disease pattern and its social determinants). However, patterns of an unknown kind are then also searched for, i.e., significances, deviations, and correlations that were not known about (Pasquinelli 2017). In general, the larger the amount of data, the greater the probability of a hit. Learning algorithms are increasingly becoming accurate with the amount of data and their learning times, assuming a correspondingly high data quality.

This is accompanied by an "epistemological shift" **from causality to correlation**, or, as Rheinberger (2007, p. 123) puts it, from hypothesis-based to data-based research. Foucault had already pointed out the "conversion from truth to functionality" in the governmentalization of modern states in the late 1970s (Foucault 2007). This means, above all, the departure from a strong ontology and the turn to the question of what "works." What "works" is that which is worthwhile and statistically valid in “political arithmetic.” Here, semantics and the mathematical proximity of status, state and statistics are significant, which are the precursors of a data and algorithm-based control of health policy, on a small and large scale (Pankoke 1984, esp. p. 998 ff.).

Evidence-based medicine has been working with statistical methods for some time now, arriving at statements and justifying decisions. However, up to now, the claim of proving causality between two or more factors remains decisive. This claim may be superseded by “evidence of correlation.” To the extent that artificial intelligence is used in the healthcare system, the correlative moment could come to the fore, because technical systems work based on statistical methods. Whether and to what extent this will be positive or negative for concrete treatment decisions remains unknown. This is ethically relevant, in the sense of the condition of the possibility of ethical judgements, because the basis for these judgements is something which is increasingly eroding. Digitization, networking of digital end devices and the basing of their processes on **Big Data** mean, epistemologically speaking, that what is relevant is that which (due to high data rates and statistical patterns) gives the most probable case. However, humans are not equipped for such a “reference to reality,” which is why it could become increasingly difficult for them to make ethical decisions within the framework of such a relationship to reality – even against the statistical evidence of machines (cf. epistemological and theoretical problems for humans versus artificial intelligence: see also Manzeschke and Brink, Ethics of digitalization in industry and Frenz, Who owns the data?).

In recent years, efforts have been made to make more information available in digital form, or to collect relevant data in a digital form in the first place. A certain “collecting mania” can also be observed, in which data is “dug up” as a potential raw material (**data mining**), without having a concrete area of application (Ashwinkumar and Anandakumar 2010; Bou Rjeily et al. 2019).

At first, the collection of data, its digitization and its integration into a digital infrastructure is primarily a technical problem. Legally it is one, because no data can be collected without a defined reason. Beyond this, however, one should not overlook the fact that digitization is epistemologically a prerequisite. Measuring errors of systematic and occasional form, arithmetic and programming errors and erroneous assumptions must not be disregarded here, because existential decisions may depend on them.

A Big Data basis will give digitalized healthcare a different and broader information base, which should lead to better insights, a larger stock of knowledge and thus to better orientation with regards to decisions on health. This applies from the level of individual patient care (micro level) to the more efficient management of a healthcare organization based on the comprehensive and meaningful process data (meso level), to the health policy control of regional, national, and even global tasks (macro level). As much as this is to be welcomed, and should create advantages for healthcare at all levels and in all areas, it also has its downsides. A comprehensively networked digital infrastructure in the healthcare system also means that it is vulnerable “from inside and outside,” and is particularly vulnerable. Hacks, viruses, Trojans, and other manipulation can cause great damage, which in many cases affects people. With the degree of networking, the range of possible damage also increases, which is why approaches are being developed to protect the physical and electronic infrastructure and its processes (cf. SAFECARE 2018).

As a rule, health-related data is extremely sensitive, very personal, and thus closely linked to freedom (Article 2 of the Basic Law), which requires a high degree of care and safeguards against misuse by actors processing this data. In addition to this legal aspect, it must be emphasized that people must be respected, and their **dignity** and **integrity** protected. This requirement applies, to a greater extent, to those who, due to their **vulnerability**, cannot or can only partially stand up for their integrity and rights.

2.3 Cross-Sector Supply

Digitalization means that technical equipment functions at the level of this digital code and is networked with it via a digital information line, which should lead to an increase in range and network effects in healthcare, meaning the efficiency and bandwidth of care services should be greatly increased (keywords include: e-health, telemedicine, tele-care). The care of patients beyond the borders of the respective organizations (hospital, rehabilitation institution, nursing home, family doctor, etc.) and thus also beyond the borders of the respective social code books have been required for decades. This could be promoted by digitization, as networking and overcoming borders is one of the central features of digitization, its “functional logic.”

The cross-sectoral care of patients is objectively and economically reasonable and ethically necessary because it is oriented towards their well-being. However, this is not without side effects. The **dissolution of sectoral boundaries** with digitization goes beyond the structures of the healthcare system. “**Dissolution of sectoral boundaries**” is the term we use to describe the fact that the large amount of data from very different sectors (consumption, living habits, nutrition, exercise, income, education, etc.) makes it possible to draw health-relevant conclusions. Social medicine has been doing this in an analogous way for quite some time but based on the big-data applications, the whole process gains new quantities and qualities. A simple variation is the self-tracking of people with fitness bracelets, delivering their data to a health insurance company (they may receive a bonus for this) (Selke 2016). This data becomes an essential factor for the evaluation of a person, and ultimately their pricing for insurance companies and other healthcare institutions. This applies to specific personal data, but it also applies, to an increasing extent, to statistical conclusions about the individual based on meta and mass data.

In the networking of many data sets, health-related data (in the narrowest sense) is included in the concrete care for the individual or in health research and care for society, and data from other areas of life. Whether and to what extent this is recognizable for the individual will be difficult (if not impossible) to understand. The data will increasingly become a point of interpretation and intervention in healthcare, but the question is to what extent the patient can understand this interpretation and intervention, because to some extent they cannot see or understand this data. The **autonomy** of the patient, who is given high priority in all health-related decisions, is, at least for the time being, opposed by a lack of **sovereignty** in the rights of access and disposal of the data, which is decisive for health-related decisions. The **vulnerability** resulting from illness or disability is now, in a sense, doubled again on the data side. The prospect of being able to behave "sovereignly" here, as envisaged by the German Ethics Council (2018), makes high technical, legal and health policy demands, to consider the ethical good of patient autonomy. The design of the **electronic patient record**, with its access rights and the transparency of data stored elsewhere or used for calculation purposes, will be a central factor in establishing the trust and acceptance of such a system (IEEE 2018; Ozair et al. 2015; Zuckerman 2014; Kopala and Mitchell 2011). As in other industries, demands for transparency become difficult here, because the algorithmic processing of the data is often removed from public view, and is a trade secret.

A healthcare system based on citizens' data, which regulates access, pre-structures decisions and carries out evaluations, could increasingly lose solidarity and develop into a merit-based system, in which individual performance and willingness to adapt become the benchmark. In the view of solidarity and justice, health policy debates will have to be held on the shape we want to give a digitalized healthcare system. This will also change our understanding of health and illness. Health will be able to—and will have to—be conceptualized in a much broader sense, based on data sets and algorithms. This need not be bad or wrong per se, but one should be aware of its possibilities and inherent tendencies to shape such developments from proven normative considerations.

2.4 Surveillance in Everyday Life

The collection of health-related data is accompanied by a specific form of monitoring, surveillance, control, and regulation of people. What this means in political and democratic theory can only be hinted at here (Christl 2017; Zeh 2012). At this point, it is more about forms of self-government (because one knows that one is observing oneself) and regulation by means of approaches that may have their starting point in the healthcare sector, but which go beyond it, grasping and shaping the individual in a comprehensive sense. This is where the "digital double" plays a role, the data shadow that we all create based on our own data traces or in the data of others, to whom we are related. As already mentioned, a physical person has only very limited control over the collection, combination, and evaluation of this data. Neither can they track which conclusions are drawn from this data by other parties. The fact is that people are affected by these evaluations less due to their physical presence than to their data (which is increasingly the starting point for evaluations, interpretations, and interventions). This will bring about far-reaching changes in concepts of autonomy, self-determination, action, and responsibility.

The policy approach to improving healthcare based on health-related data will not stop at the boundaries of healthcare but will be much more far-reaching. Health is a particularly sensitive area in which data is at stake, which makes it a morally sensitive issue, and calls for the utmost care. At the same time, the willingness of people in the healthcare sector is particularly great (and perhaps also exploitable), renouncing some rights and freedoms for the sake of health so that far-reaching standardization and regulation is conceivable. This will not be limited to compliance but will make a much more comprehensive government of the body possible, in the sense of Foucault. Human rights and privacy should not only be something for healthy people, while sick people have to do without due to their digital support.

Some central ethical and political questions are: how can we use the opportunities of data-based healthcare without falling into dystopian monitoring scenarios? What price are we willing to pay for health and well-being? How do we ensure that the currency in which we pay is not our democratic rights and freedoms?

3 Robotics in Healthcare

It seems obvious that robots will also be used in the healthcare sector, considering the increasing need for care and the decreasing number of skilled workers, especially in nursing. Robots are not robots after all, and so, although many things are still in the development and testing stage, we will have to differentiate more precisely between different fields of application and robot types (Manzeschke 2014a, b; Manzeschke and Petersen 2020). Robots will replace and support human activities and ideally improve effectiveness, efficiency, and precision. This is already urgently needed, and even desirable in some respects. On the other hand, we will also have to pay a price for these gains, and it makes sense to agree on this price ex ante.

In the field of care, robots are far from being ready for use, as public debates on "**care robots**" sometimes suggest. In the domestic sector, however, it is not only the high purchase prices but the living conditions which go against this. In homes, this problem is less present due to building structures, but even here, robotics is not yet a very common occurrence:

> Technology that supports the independence of the residents and at the same time provides a working aid for carers in the provision of services close to the body (washing, support when going to the toilet or getting dressed) is hardly ever used - electronic height adjustment, adjustable toilet seats, washstands and cupboards equipped with a lifting and lowering mechanism for the clothes rail are virtually unrecognizable, although they have been on the market for years. (Isfort et al. 2018, p. 63)

Even if these problems are now being addressed in research policy terms by care innovation and care practice centers (which have the task of systematically developing digital care technologies with the care sector and testing their suitability for practical application), it will also depend very much on the acceptance of the technology in the profession and the population—and there is still a considerable need to catch up here (Technikradar 2018, p. 50 ff.). It is of great importance that technical support can only substitute or compensate for certain tasks but cannot replace the interpersonal care provided by a caregiver. Even though there may currently be an empirical lack of good interpersonal care, this does not go against the ethical requirement to design socio-technical arrangements in such a way that this is made possible, and indeed guaranteed.

Once again, the situation is different for the use of digital technology in the field of medicine, especially clinical medicine. Here, the use of technology has a long tradition, and is met with widespread approval. But here, too, ethical questions arise, which result from support provided by AI systems (Manzeschke 2014c, 2018), and from changing hierarchies and forms of interaction (Kluge 2017; Steil et al. 2019). The background to this is the observation that robots clearly exceed the status of instrumental technology, and "appear" to be new interactive partners, whose status has not yet been clarified epistemologically, legally, morally, or socially. The German Ministry of Education and Research (BMBF) programmatically formulated this in 2013:

> Technical systems are increasingly being developed from pure tools to cooperative, interactive partners. This opens up a variety of opportunities in different areas of life. They will increasingly support people at work or in everyday situations, and make an important contribution to strengthening their productivity, social participation, health or everyday competence. (Bundesministerium für Bildung und Forschung (BMBF) 2013)

Ethically, it is necessary to ask more precisely about the relation, interaction, cooperation or collaboration between man and machine in each case. However, this is not only an empirical question, but has normative implications that need to be discussed and implemented in regulations. To this end, the EU has presented drafts for the legal status of electronic persons ("ever more sophisticated robots, bots, androids, and other manifestations of artificial intelligence") (see EU 2015; Manzeschke and Brink 2022). An important question here is the roles the "interaction partners" are assigned: if a machine remains under the control and ultimate responsibility of humans, which is an ethical demand that has long been made in dealing with robotics (Sturma 2004). Given the performance of machines and the incomprehensibility of their "decisions," does this still hold for humans at all? Do we humans have to accept the fact that we have developed machines that are vastly superior to us in terms of rationality, precision, and speed, and should we therefore trust their decisions? What possibilities does a treatment team in a clinic have of deviating from the recommendations of the expert system (databases and algorithm) for a certain therapy? How should such a deviation be justified if the "data situation" is against it? How should treatment teams justify deviating ideas if their own experience is increasingly accompanied and guided by data-based decision support systems? This problem becomes more acute as data-based evidence becomes the standard, and defense against recourse underpins compliance with standards.

Here, it is clear that technical structures and procedures will have an impact on humans in general and on professionalism. In a broad sense, this concerns our conception of human beings as moral subjects, who set their goals based on free will and who select and use certain means for this in a well-founded manner. It may be that confrontation with the systems we have developed ourselves will bring up some questions about this conception of freedom, rationality, and morality. How will human **autonomy** merge with that of artificial intelligence, robots, or intelligent environments in the future? How will we conceive decisions and actions in the future, when self-learning technical systems will play an important role, but will increasingly elude us humans?

4 Age-Appropriate Assistance Systems

An important "right of way" rule in the healthcare system is "outpatient before stationary": health measures should be provided, as far as possible, at home or in passing, to reduce costs and to keep inpatient stays in clinics or nursing homes as short as possible. In the realm of assistance and care, the use of so-called **age-appropriate assistance systems** plays a special role in consistently pursuing these goals and keeping the intended stay something for the affected people to decide. Assistance systems increasingly converge to form an intelligent environment in which medical, nursing, social and logistical aspects are interlinked, and which is intended to provide the affected person with a high degree of safety, comfort, self-determination, and social participation (Manzeschke 2019). The concept of demographic change and society's human-rights assurances towards people with cognitive, sensory, or motor disabilities, who are often age-associated, "age-appropriate assistance systems" have a certain degree of justification (United Nations 2006). At the same time, it is becoming increasingly clear that such limitations not only correlate with age, but can also occur independently, at any stage of life. Moreover, assistance is not only a necessary support in limitations, but can also be a comfort factor. Many assistance systems cannot thus be assigned to a medical or nursing purpose in the strict sense, e.g., window and door locks, power, water, and heating controls. The ethical argument changes immediately with regard to the use of such assistance systems if it is not about an extension of comfort, but about necessary use in the case of illness, disability, or old age. In addition to these general assistance systems, specific systems can also be identified which are important in healthcare contexts, such as sensorimotor systems for pressure sore prophylaxis, systems for route planning and documentation in outpatient care, and avatars to support cognitively impaired people. It is impossible to present the full range of systems and application contexts in this article, which is why some basic coordinates are marked here, and reference is made to further literature.

On assistance itself, a distinction must be made between primary assistance (which is provided to the person in need of support) and secondary assistance (which supports the supporting person). Furthermore, the share the technical assistance has in the overall arrangement has to be clarified. The duration, frequency and communication modes of the assistance can be specified, as well as legal conditions and cultural backgrounds. All these factors must be considered when asking about the ethical implications of technical assistance. Here, too, attention must be paid to the way in which data management and decision-making procedures must be made transparent to the people supported, and in which way the socio-technical arrangement influences relationships, sociality, and interpersonal interactions. In our opinion, the "how" is measured by the human condition, specifically by the opening of self-determined freedom with real relationships and a dependence on other people. This moment of humanity must not be lost, despite all the technical substitution and compensation of human activity. Empirically, one could dismiss this as naive and idealistic, with reference to the current situation. From a normative perspective, however, it must be countered by the fact that we are dealing with the question of what should be. How are we to design and use technical assistance systems to enable good healthcare, and not to control central moments of interpersonal relationships?

This shows once again that the ethical assessment of technology is not concerned with isolated technical artifacts, but the context of their use. The model for the ethical evaluation of socio-technical arrangement (**MEESTAR**) has been developed for such a consideration and has been in use for several years (Manzeschke et al. 2015; Weber 2015; Manzeschke 2015). It aims to integrate ethical and social issues at the early stages of research, development, construction, and distribution of technical assistance systems, to strengthen the ethical judgement of designers and users, and to discuss the responsible design of digital applications in a broad social debate.

5 Conclusion

As in other sectors of society, digitalization in the healthcare sector opens up opportunities for improving healthcare and maintaining current standards of care under more difficult conditions. As in other sectors, similar questions arise about the use of digital technology: replacement of jobs, changes in the hierarchies and processes of work, changes in relationships. In the healthcare system, these changes are given special emphasis, because encounters and relationships are constitutive for the process of diagnosis, therapy, and care, and it is therefore of particular ethical relevance that these encounters and relationships are not corrupted by the use of technical systems. It is essential that socio-technical arrangements are designed in such a way that the necessary ethical questions are identified and evaluated at an early stage. Only then there is a chance for these considerations to become effective in the technical design of arrangements. Otherwise, ethical considerations will always come too late because the technical structure and infrastructure has already been developed.

Here, the direction in which ethical problems are expected to go has been outlined. The shaping of digitalization is a political task, which has at its core the recognition of people as citizens of a community, and the need to address them with regard to their rights and duties.

References

Ashwinkumar UM, Anandakumar KR (2010) Ethical and legal issues for medical data mining. Int J Comput Appl 1(28). https://doi.org/10.5120/520-841

Bartens W (2018) Dr. Data. In: Süddeutsche Zeitung vom 18. März 2018. https://www.sueddeutsche.de/wissen/medizin-dr-data-1.3908072. Accessed 28 Jan 2021

Bindé J (2004) The future of values: 21st century talks. Berghahn Books, Incorporated

Böhme G (2001) Ethics in context. The Art of Dealing with Serious Questions. Polity Press, Cambridge in Association with Blackwell Publishers Ltd., Oxford

Bou Rjeily C, Badr G, El H, Hassani A, Andres E (2019) Medical data mining for heart diseases and the future of sequential mining in medical field. In: Tsihrintzis G, Sotiropoulos D, Jain L (eds) Machine learning paradigms. Intelligent systems reference library, Bd 149. Springer, Berlin, pp 71–99

Braidotti R (2013) The Posthuman. Polity Press, Cambridge

Bundesministerium für Bildung und Forschung (BMBF) (2013). https://www.bmbf.de/foerderungen/bekanntmachung-908.html. Accessed 28 Jan 2021

Callahan D (1998) False hopes. Simon and Schuster, New York

Christl W (2017) Corporate surveillance in everyday life. How companies collect, combine, analyze, trade, and use personal data on billions. Open Society Foundations, Wien

Cukier K, Mayer-Schönberger V (2013) The rise of big data. Foreign Aff:28–40

De Carolis M (2004) La vita nell' epoca della sua riproducibilita tecnica. Bollati Boringhieri, Torina

European Union (2009) Treaty of Lisbon. https://www.europarl.europa.eu/factsheets/en/sheet/5/the-treaty-of-lisbon. Accessed 28 Jan 2021

European Union (2015) European Parliament on rules for robotics, 2015/2103(INL) – 16/02/2017 Text adopted by Parliament, single reading. http://www.europarl.europa.eu/oeil/popups/summary.do?id=1477231&t=d&l=en. Accessed 28 Jan 2021

Foucault M (2007) Security, territory, population. Lectures at the College de France. Palgrave Macmillan, New York, pp 1977–1978

German Ethics Council (2018) Big Data und Health – Data Sovereignty as the Shaping of informational Freedom. Deutscher Ethikrat, Berlin

Grafenstein MV, Hölzel J, Irgmaier F, Pohle J (2018) Nudging. Regulierung durch Big Data und Verhaltenswissenschaften. Gutachten im Rahmen des BMBF-Projekts Assessing Big Data. http://www.abida.de/sites/default/files/ABIDAGutachten_Nudging.pdf. Accessed 28 Jan 2021

Grunwald A (2019) Der unterlegene Mensch. Die Zukunft der Menschheit im Angesicht von Algorithmen, künstlicher Intelligenz und Robotern. Riva, München

Gruskin S, Grodin MA, Annas GJ, Marks SP (eds) (2005) Perspectives on health and human rights. Routledge, New York

Hofstetter Y (2014) Sie wissen alles. Wie intelligente Maschinen in unser Leben eindringen und warum wir für unsere Freiheit kämpfen müssen. Bertelsmann, München

IEEE | Institute of Electrical and Electronics Engineers (2018) Ethically aligned design. A vision for prioritizing human well-being with autonomous and intelligent systems. https://standards.ieee.org/industry-connections/ec/ead-v1.html. Accessed 28 Jan 2021

Isfort M, Rottländer R, Weidner F, Gehlen D, Hylla J, Tucman D (2018) Pflege-Thermometer 2018. Eine bundesweite Befragung von Leitungskräften zur Situation der Pflege und Patientenversorgung in der stationären Langzeitpflege in Deutschland. In: Deutsches Institut für angewandte Pflegeforschung e.V. (DIP), Köln. https://www.dip.de/fileadmin/data/pdf/projekte/Pflege_Thermometer_2018.pdf. Accessed 28 Jan 2021

Kersting W (2012) Wie gerecht ist der Markt? Ethische Perspektiven der sozialen Marktwirtschaft. Murrmann, Hamburg

Kirkwood T (1977) Understanding the odd science of aging. Cell 120(2):437–447

Kluge EW (2017) Health information professionals in a global eHealth world: ethical and legal arguments for the international certification and accreditation of health information professional. Int J Med Inform 97:261–265

Kolany-Raiser B, Heil R, Orwat C, Hoeren T (Hrsg) (2018) Big Data und Gesellschaft. Eine multidisziplinäre Annäherung. Springer-VS, Wiesbaden
Kopala B, Mitchell ME (2011) Use of digital health records raises ethics concerns. J Nurs Adm Healthc Law Ethics Regul 13(3):84–89
Kroes N (2015) Ich bin nicht naiv, und Europa darf es auch nicht sein. In: Schirrmacher F (Hrsg) Technologischer Totalitarismus. Eine Debatte. Suhrkamp, Berlin, S 96–106
Kurzweil R (2005) The Singularity is near: when humans transcend biology. Penguin, London
Lanier J (2014) Who owns the future? Simon & Schuster, New York
Manzeschke A (2014a) Heikel ist, wenn die Technik die Fürsorge übernimmt. Ethische Kernfragen. Thema Fokus 115:7–11
Manzeschke A (2014b) Roboter in Pflege und Betreuung: hilfreich oder schädlich? Ethische Fallbesprechung. Thema Fokus 15(116):20–21
Manzeschke A (2014c) Digitales Operieren und Ethik. In: Niederlag W, Lemke HU, Strauß G, Feußner H (Hrsg) Der digitale Operationssaal. de Gruyter, Berlin, S 227–249
Manzeschke A (2015) MEESTAR: Ein Modell angewandter Ethik im Bereich assistiver Technologien. In: Weber K, Frommeld D, Manzeschke A, Fangerau H (Hrsg) Technisierung des Alltags. Beitrag für ein gutes Leben? Franz Steiner, Stuttgart, S 263–283
Manzeschke A (2018) Auswirkungen neuer Technisierung auf das Arzt-Patient-Verhältnis am Beispiel digitaler Operationstechniken. In: Klinke S, Kadmon M (Hrsg) Ärztliche Tätigkeit im 21. Jahrhundert. Profession oder Dienstleistung. Springer, Berlin, S 289–310
Manzeschke A (2019) Technische Assistenzsysteme. In: Fuchs M (Hrsg) Handbuch Alter und Altern. Anthropologie – Kultur – Ethik. J. B. Metzler, Stuttgart/Weimar (coming soon)
Manzeschke A, Brink A (2022) Ethics of digitization in industry. In: Frenz W (Ed) Handbook industry 4.0: law, technology, society. Springer, Berlin/Heidelberg
Manzeschke A, Petersen J (2020) Digitalisierung und Roboterisierung in der Pflege. Ethisch-anthropologische Überlegungen. In: Dibelius O, Piechotta-Henze G (eds) Plädoyer für eine menschenrechtsbasierte Pflege. Hogrefe, Bern, pp 161–173
Manzeschke A, Weber K, Rother E, Fangerau H (2015) Ethical Questions in the area of age appropriate assisting systems. VDI/VDE, Berlin
Nefiodow LA, Nefiodow S (2014) The sixth Kondratieff: a new long wave in the global economy. St. Augustin
Niederlag W, Lemke HU, Nagel E, Dössel O (2008) Gesundheitswesen 2025. Implikationen, Konzepte, Visionen, Bd 12. Health Academy, Dresden
Niederlag W, Lemke HU, Golubitschnaja O, Rienhoff O (Hrsg) (2010a) Personalisierte Medizin. Sind wir auf dem Weg zu einer individualisierten Gesundheitsversorgung? Bd 14. Health Academy, Dresden
Niederlag W, Lemke HU, Rienhoff O (Hrsg) (2010b) Personalisierte Medizin & Informationstechnologie. Innovative Konzepte, realisierte Anwendungen, gesellschaftliche Aspekte, Bd 15. Health Academy, Dresden
Ozair FF, Jamshed N, Sharma A, Aggarwal P (2015) Ethical issues in electronic health records: a general overview. Perspect Clin Res 6:73–76
Pankoke E (1984) Soziologie. In: von Brunner O, Conze W, Koselleck R (Hrsg) Geschichtliche Grundbegriffe. Historisches Lexikon zur politisch-sozialen Sprache in Deutschland. Klett-Cotta, Stuttgart, S 997–1032
Pasquinelli M (2017) Arcana Mathematica Imperii: the evolution of western computational norms. In: Hlavajova M et al (eds) Former West: art and the contemporary after 1989. MIT Press, Cambridge
Reiser SJ (1978) Medicine and the reign of technology. Cambridge University Press, Cambridge
Reiser SJ (2009) Technological medicine. The changing world of doctors and patients. Cambridge University Press, Cambridge
Rheinberger J (2007) Wie werden aus Spuren Daten, und wie verhalten sich Daten zu Fakten? In: Nach Feierabend. Zürcher Jahrbuch für Wissensgeschichte, Nr, 3, 117–125

SAFECARE (2018) Safeguard of Critical Health Infrastructure. EU 2020 Horizon Projekt. https://www.safecare-project.eu/. Accessed 28 Jan 2021
Savulescu J (2007) Genetic Interventions and the Ethics of Enhancement. In: Steinbock B (ed) The Oxford handbook of bioethics. Oxford University Press, Oxford, pp 516–535
Schirrmacher F (2015) Das Armband der Neelie Kroes. In: Schirrmacher F (Hrsg) Technologischer Totalitarismus. Suhrkamp, Berlin, S 62–69
Selke S (2016) Lifelogging. Digital self-tracking and Lifelogging – between disruptive technology and cultural transformation. Springer VS, Wiesbaden
Spreen D, Flessner B, Hurka HM, Rüster J (2018) Kritik des Transhumanismus. Über eine Ideologie der Optimierungsgesellschaft. Transcript, Bielefeld
Steil J, Finas D, Beck S, Manzeschke A, Haux A (2019) Robotic systems in operating theatres: new forms of team-machine interaction in health care. On challenges for health information systems on adequately considering hybrid action of humans and machines. In: Methods of information in medicine, vol 58, pp e14–e25
Sturma D (2004) Ersetzbarkeit des Menschen? Robotik und menschliche Lebensform. Jahrb Wissensch Ethik 9:141–162
Technikradar (2018) Was die Deutschen über Technik denken. Schwerpunkt: Digitalisierung. Hrsg. von acatech, München, und Körber-Stiftung, Hamburg
Toulmin S (1982) How medicine saved the life of ethics. Perspect Biol Med 25(4):736–750
United Nations (2006) Convention on the Rights of Persons with Disabilities. https://www.un.org/disabilities/documents/convention/convoptprot-e.pdf. Accessed 28 Jan 2021
Waschinski G (2019) Google, Apple und Amazon werden zur Gefahr für deutsche Krankenversicherer. In: Handelsblatt vom 3. März 2019. https://www.handelsblatt.com/unternehmen/it-me-dien/sxsw-2019/gesundheitspolitik-google-apple-und-amazonwerden-zur-gefahr-fuer-deut-sche-krankenversicherer/24056698.html?ticket=ST-102566-cgGc6VIBIevpTThGEtWY-ap3. Accessed 28 Jan 2021
Weber K (2015) MEESTAR: Ein Modell zur ethischen Evaluierung sozio-technischer Arrangements in der Pflege- und Gesundheitsversorgung. In: Weber K, Frommeld D, Manzeschke A, Fangerau H (Hrsg) Technisierung des Alltags. Beitrag für ein gutes Leben? Franz Steiner, Stuttgart, S 247–262
Weichert T (2018) Big Data im Gesundheitswesen. Gutachten im Rahmen des BMBF-Projekts Assessing Big Data. http://www.abida.de/de/blog-item/gutachten-big-data-im-gesundheitsbereich. Accessed 28 Jan 2021
Weiß MG (2009) Bios und Zoë. Die menschliche Natur im Zeitalter ihrer technischen Reproduzierbarkeit. Suhrkamp, Frankfurt
Zeh J (2012) The method. Random House, London
Zuckerman EL (2014) Electronic Health records raise new ethical concerns. The National Psychologist. https://nationalpsychologist.com/2014/01/electronic-health-records-raise-new-ethi-cal-concerns/102429.html. Accessed 28 Jan 2021

Ethics of Digitilization in the Financial Sector Using the Example of Financial Services

Arne Manzeschke and Alexander Brink

Contents

1 Introduction

The digitalized financial world finds its expression in the many screens in front of which brokers sit and complete their trades—often, still by telephone. However, this is an insignificant part of financial trading. By far, the greater part is now handled by high-speed computers, which make buying and selling decisions in fractions of a millisecond, accumulating minimal differences in value to enormous profits and significantly influencing the financial sector and, from there, labor and goods markets (Gsell 2010; Gresser 2016, 2018). Financial economists such as Bernard Lietaer assume that a key moment in the financial crisis of 2007, which began with the collapse of the Lehman Brothers investment bank, was due less to the greed of managers and speculators than to the hyper-efficiency of **high frequency trading**: "Natural river systems become sustainable because nature does not strive for maximum efficiency, but for an optimal balance between efficiency and resilience" (Lietaer 2009, p. 157). According to Lietaer, the systems were too strongly trimmed for efficiency, and were therefore not sufficiently resilient. The collapse of the banks led to remarkable social problems behind the technical processes. Banks lost

---

A. Manzeschke (✉) · A. Brink
Lutheran University of Applied Sciences - Evangelische Hochschule Nürnberg, Nürnberg, Germany

Bayreuth University, Bayreuth, Germany
e-mail: arne.manzeschke@evhn.de; alexander.brink@uni-bayreuth.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_60

confidence and stopped lending money. The credit crunch also threatened to bring social life to a standstill, and could only be averted by massive state guarantees and debt absorption by the taxpayer. Once again, this example shows that technical infrastructures have a strong influence on the world, determining how people experience the possibilities they have to realize their own lives, and how coexistence is shaped. This calls for a serious moral question, which according to Böhme (1997, p. 17, see also Manzeschke and Brink 2022) is reflected in ethics.

Individuals can leverage and convert enormous amounts of money with little effort, which has repeatedly led to major cases of fraud and put the banks concerned in a considerable predicament. The financial crisis of 2007 et seqq. has shown that a financial crisis can easily turn into an economic crisis, and then into a state crisis. The underlying problem is a much bigger one, which is only made the more glaring by digitization. The vulnerability of social structures (due to network effects, their global reach and the enormous speed at which information spreads) leads to a potential destabilization of state, especially democratic structures (Vogl 2010, 2015).

Further analyses which can fairly assess the connections with the necessary distance are pending. Nevertheless, the financial sector is experiencing a basic problem of digitalization and networking: the vulnerability of the systems increases with the degree of networking and the speed of their performance. For humans, algorithmic instructions for buying and selling being executed by high-speed computers have long since become incomprehensible (see also Manzeschke and Brink 2022). From an ethical perspective, this is a problem of attributing and assuming responsibility. It also becomes apparent how technical acceleration is creating or continuing a competition of its own between the market players, which consists of gaining an advantage over the other players through increasingly powerful computers (Anthony 2016).

Such phenomena are, above all, not a genuinely technical problem, but more precisely one based on digitilization. They illustrate a profoundly human problem, which is intensified by the digital infrastructure in and through which money is increasingly being created, traded and allocated.

> Our brain reacts to money in a similar way to cocaine, morphine or sex. This does not increase our chances of dealing rationally with monetary quantities, and presumably this irrationality (intoxication) is systematically increased once again by the virtual indispensability of profit opened up by financial capitalism. The virtuality, which opens an infinite horizon and turns the money sphere into an alchemy project, justifies the fact that the financial markets represent an ethically relevant area. (Schramm 2015, p. 124)

Based on our approach to the ethics of digitilization, which we have presented in a separate chapter (see also Manzeschke and Brink 2022), we want to outline some ethical questions for the financial services sector in this section, without (of course) being able to provide any answers here: the field is still too new and uncharted. Nevertheless, it will be possible to formulate some normative orientations.

2 The Example of Customer Advice

If one follows the first paragraph of the German Banking Act, credit institutions are defined as enterprises that conduct banking business on a commercial basis. In Germany, these are private banks, cooperative banks, and public law banks. They form the three pillars of the German banking system. If a company provides financial services that are not covered by the above-mentioned banking business (according to the first paragraph of the German Banking Act), it is called a financial service institution (according to the second paragraph of the same act). Financial services are then understood to be a group of financially marketable services. Providers are the **financial intermediaries**: usually companies that mediate between supply and demand on financial markets. Financial instruments and financing instruments are offered, along with asset management, portfolio management and financial advice. The **financial intermediaries** which offer services include credit institutions, insurance companies, building societies and credit brokers. As a rule, other financial intermediaries, companies and natural persons act as customers.

Customer advisory services are an important factor in the value creation of the banking business and financial services. In the course of the digitalization of banks and financial service providers, the use of artificial intelligence to support **customer advisory** services is discussed. In the context of a low interest rate environment, regulations and a loss of trust, this results in enormous potential.

In the case of savings and cooperative banks, not only is the public welfare principle anchored in the statutes but also the **honorable merchant**, for example in the IHK law. According to the Chambers of Industry and Commerce, they have to act for "the election of decency and custom of the honorable merchant" (§ 1 IHK Law). From this practice, fundamental questions arise such as: does a respectable banker or an artificial intelligence decide on this? How is artificial intelligence compatible with the model of the **honorable merchant**, which real people have to face? Can respectability be programmed into artificial intelligence? Can artificial intelligence learn respectability?

In the case of complex products and services which require intensive advice, the client advisor provides assistance, as they have an advantage over the client in terms of information. The client advisor traditionally advises the client by telephone, online or face-to-face. Digitalization has had a strong impact on the **advisor-client relationship**. Banks (both savings banks and cooperative banks) and insurance companies are increasingly supported by technical assistance for **customer advisory**. The worldwide rise of the call centre was one of the first developments in this direction. The customer no longer had to visit in person: distance selling was made possible over the phone. Most banks operate their own call centers for customer service. Increasingly, the advisory process between customer advisor and customer is now supported by the use of algorithms and artificial intelligence (see the Digital Banking Expert Survey of GFT Technologies, SE 2017). AI based solutions dominate **banking platforms**. There are also the first **socio-technical arrangements** for large insurance companies (where the customer is advised via chatbots) or banks (where the trading of securities is fully automated, also called algorithmic trading).

Customer advice is based on a moral claim that can be associated with Ricœur's Relata "**I, Others and Institutions**" (Ricœur 2005). The relationship between "I and Others" is expressed precisely in the bilateral structure of the client-consultant relationship. The institutions include both state actors (such as the banking supervisory authority BaFin) and regulations that banks and financial service providers place upon themselves (e.g., a code of conduct or ethics). Recently, and in the aftermath of the financial crisis, numerous regulations have been enacted which have had an impact on the client-consultant relationship. Cooperative banks are supported by the so-called "Genossenschaftliche Beratung" concept, savings banks by the so-called "S-Finanzkonzept." This results in **socio-technical arrangements** which must now be subjected to a normative examination. Here, it is not yet clear how this combination of moral normative claims and algorithmic-technical efficiency can be brought together, along with which role humans play on each side of the arrangement, or how they can follow and determine the negotiation processes.

We would encourage the supportive use of the machine in the customer relationship, which is consistently oriented towards the human being (in their various roles and positions, for example as customers or employees). This is about **the "good life"** of the customer, with regard to the financing of planned studies, trips, home, or insurance against risks such as illness, accident, unemployment, or death. In a closely considered "**I-Other" constellation**, it is also about **the good life** of the consultant. They should ultimately be relieved in their work, have a good income, be satisfied with their work, and see it as meaningful. It has been shown that through an increase in customer contact (e.g., due to a rising number of customer inquiries or complaints, also made possible by digitalization), the employee in the **customer advisory** service is strongly challenged. The relief should mainly take place during routine work, ideally as a frequently presented argument. This will allow the client advisor to work more efficiently and better, in terms of addressing clients and focusing on their needs. For example, artificial intelligence is currently used for knowledge management, but also for fully automated mail processing, identifying customer discontent at early stages. A customer can also be advised completely digitally if they wish, if it corresponds to the product, the service and their personal life. Banks and other financial service providers use algorithms to investigate the creditworthiness of their customers. In addition, initial success has been achieved in the implementation of compliance and the detection of moral misconduct, e.g., in fraud detection or money laundering prevention. Artificial intelligence can analyze documents for conspicuous passages and mark them accordingly. In this case, however, measures sometimes conflict with customers' privacy claims, and with inconsistent international regulations. In addition, fiduciary duties and legal aspects must be guaranteed, such as risk assessment in investment advice or information on indebtedness when granting loans.

If the **automation** of work processes in the granting of credit leads to savings in transaction costs and can identify customer needs individually and quickly, there is a risk of **discrimination** by AI. For example, a customer's creditworthiness can be checked based on a relatively small amount of personal data (such as place of residence, gender, occupation) in combination with pattern recognition (e.g., the level of loan repayment rates in this quarter), which statistically achieves quite high success rates. This may mean that it is not the situation of a borrower that is used for the decision, but their profile, which, when compared with many other profiles, appears to be statistically significant enough to delegate such decisions to an AI. If one considers that the "decision" made by a machine may not even be comprehensible to the lender, it must be asked from an ethical perspective whether organizations are able to keep their professional promises (honesty, trust, fairness, and personal relationship) in this way. The question also arises as to whether **justice** can still be given to the individual. It is questionable whether AI can be trained without discrimination, so that it does not select according to factors such as ethnicity or religion.

The normative approach is therefore to weigh up certain efficiency gains achieved through **automation** against the "promises" that companies have made to their customers (Brink 2019). Some banks, for example, are putting people back at the centre of their activities, such as GLS Bank (with its concept of transparency) or Teambank (with its principle of fairness). Both are examples of a promise to customers that is also referred to in **customer advisory** services. In this context, the concept of **Corporate Digital Responsibility** is also discussed (Brink and Esselmann 2016). Additionally, the consequences of the automated calculation of credit applications for the credit system should be considered. What effect does this have on innovations, such as on the granting of credit to niche businesses or start-ups that may not be properly covered by the usual parameters? What effects does this have on the credit structure as a whole? Will there be individualized loans, each with their own risk surcharges? Is this still covered by our ideas of justice? Which measures must banks take when purchasing algorithms to be able to guarantee (at least in principle) the accountability of machine-made decisions? How can ART criteria (Accountability, Responsibility, Transparency) be considered and presented to the customer (and also to the employee)?

3 Conclusion

Overall, special attention is always paid to individual cases. Even within a socio-technical arrangement, conflicts of norms, ambivalences and dilemmas are possible, perhaps even the norm. Ultimately, human judgement decides on the application and assessment of different standards, in the context of the situation. The decisive factor is still that the human being retains control over the decision-making process, and that human competences remain available, even if machines intervene to support them (e.g., still knowing how to open an account, how to make a transfer, how to select an investment product, which products have which consequences on other markets, production conditions). The customer advisor must remain sensitive to moral questions: their power of judgement must be trained so that they can differentiate at any time, and evaluate individual cases independently. For example, investment advice for a wealthy business customer may be assessed differently to credit advice for an indebted young person.

In this context, it seems important that the **"I-Other" relationship**, which we have developed from Ricœur's ethics, is not only a customer-consultant relationship, but ultimately for other stakeholders who have a relationship with the company too, such as employees or suppliers, along with local and natural environments. Since business effectiveness (the handprint for the organization) cannot be achieved in such a short time, and the impact dimension cannot be quantified so well, very little attention is currently paid to this aspect, unfortunately. This is precisely where institutions need to be well-designed to ensure that people can live well together.

References

Anthony S (2016) The secret world of microwave networks. https://arstechnica.com/information-technology/2016/11/privatemicrowave-networks-financial-hft/2/. Accessed 1 Jan 2022

Böhme G (1997) Ethik im Kontext. Über den Umgang mit ernsten Fragen. Suhrkamp, Frankfurt

Brink A (2019) Value Promise. Das Halten von Versprechen als Legitimation unternehmerischer Tätigkeit. In: Graup S, Ötsch WO, Rommel F (eds) Spielräume des Denkens. Festschrift für Karl-Heinz Brodbeck, Metropolis, Marburg, pp 399–417

Brink A, Esselmann F (2016) Corporate Digital Responsibility. Den digitalen Wandel von Unternehmen und Gesellschaft erfolgreich gestalten. Spektrum: Universität Bayreuth. Wissenschaftsmagazin 12(1):38–41

Digital Banking Expert Survey der GFT Technologies SE (2017). https://blog.gft.com/blog/2017/10/25/gfts-digital-bankingexpert-survey-2017-whats-moving-the-financial-services-world/. Accessed 1 Jan 2022

Gresser U (2016) Praxishandbuch Hochfrequenzhandel, Bd 1. Springer Gabler, Wiesbaden

Gresser U (2018) Praxishandbuch Hochfrequenzhandel, Bd 2. Springer Gabler, Wiesbaden

Gsell M (2010) Essays on algorithm trading. Columbia University Press, New York

IHK-Gesetz. https://www.gesetze-im-internet.de/ihkg/. Accessed 1 Jan 2022

Lietaer B (2009) Erhöhte Unfallgefahr. brand eins 11(1):154–161

Manzeschke A, Brink A (2022) Ethics of digitization in industry. In: Frenz W (Ed) Handbook industry 4.0: law, technology, society. Springer, Berlin/Heidelberg

Ricoeur P (2005) Ethik und Moral (1990). In: Ricoeur P (ed) Vom Text zur Person. Hermeneutische Aufsätze (1979–1999). Meiner, Hamburg, pp 251–268

Schramm M (2015) Die menschliche Natur "schubsen". Moralökonomisches Anreizmanagement in der Behavioral Business Ethics. In: Neck R (ed) Wirtschaftsethik nach der Finanz- und Wirtschaftskrise. Wirtschaftsethische Perspektiven X. Humblot & Duncker, Berlin, pp 95–128

Vogl J (2010) Das Gespenst des Kapitals. Diaphanes, Zürich

Vogl J (2015) Der Souveränitätseffekt. Diaphanes, Zürich

Digitalization and Global Responsibility

Hartmut Sangmeister and Bernd Villhauer

Contents

Within a relatively short amount of time, disruptive digital technologies have fundamentally changed many aspects of life. Digital transformations of production processes and consumer habits are happening so rapidly and radically that they are often referred to as the Fourth Industrial Revolution. The revolutionary essence of web-based digitalization is the incredibly fast and systematic fusion of technologies that break down the barriers between the physical and digital worlds. Digital networks, which are interconnected, communicate with each other, collect and exchange information, analyze huge amounts of data, and use it as a basis for decisions (*Big Data)*, enable the mobilization of gains in efficiency, effectiveness, and innovation around the globe with their enormous performance potential. In theory, these gains could benefit everyone around the world—but in reality, not everyone will benefit equally from the "digital dividend" expected from this digital revolution (Hilser 2018, p. 92). Although three-quarters of the world's population can communicate using cell phones, around two billion people lack access to digital information and communication technologies (ICT); more than 40% of the world population is still offline while 4.66 billion people use the Internet (Datareportal, Digital 2020 July Global Statshot Report). Digital change does not necessarily mean

H. Sangmeister (✉)
Ruprecht-Karls-Universität Heidelberg, Alfred-Weber-Institut für Wirtschaftswissenschaften, Heidelberg, Deutschland
e-mail: sangmeister@uni-hd.de

B. Villhauer
Weltethos-Institut an der Universität Tübingen, Tübingen, Deutschland
e-mail: villhauer@weltethos-institut.org

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_61

a win-win situation for all societies and their members; digitalization has not yet guaranteed prosperity for everyone.

Taking advantage of the opportunities offered by digitalization is unequally distributed, both within each society and between societies. This unequal distribution of opportunities for the productive use of digital technologies leads to a "digital divide" between the states of the world, but also within individual nations. Counteracting the exclusion of certain social groups from participating in digital change is the task of national policies. To avert the "digital divide" at the global level, international development cooperation can promote digitalization in the countries of the Global South. For example, the "Digital Agenda" of the Federal Ministry for Economic Cooperation and Development (Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, BMZ) includes the use of digital innovations as one of the strategic goals of German development policy (BMZ 2017a, p. 10). The project "Digital Change in Development Cooperation" from 2017 to 2019, which aimed to integrate approaches and methods of digitalization for sustainable development into German development cooperation also served this purpose. Numerous programs and projects support, for example, local digital solutions to local challenges and promote digital centers and technology start-ups. (BMZ 2019, p. 14).[1]

Development cooperation can also help reduce asymmetries that are created or reinforced by digitalization and thus hinder autonomous development in partner countries. This becomes particularly clear when one considers the two levels of digitalization, formal and content-related. Digitalization of development cooperation itself is taking place with new forms of communication and information, *Big Data* applications, use of Artificial Intelligence (AI), etc., but also digitalization strategies in partner countries as an element of their specific development paths.

To ensure that as many people as possible can benefit from the "digital dividend," digitalization must be shaped by establishing global framework conditions and national regulations. This is not a matter of halting the far-reaching structural change that digitalization is initiating and accelerating; however, societal processes of training and standardization of law and morals are needed for the further development of digitalization in the economy and society, as well as ethics for the digital century (Hemel 2020). This can be achieved even more successfully as awareness grows, in that digitalization is not purely a technological process, but also has cultural, psychological, sociological, political, and economic dimensions. In particular, the ethical aspects must also be considered. Thus, new digitalized possibilities for shaping the world of life and work require new forms of responsibility.

[1] During the COVID-19 pandemic, digital technologies made it possible to continue the support and evaluation of development cooperation projects and programs together with the responsible persons in the partner countries even without extensive travel and on-site visits.

1 Digitalization Opportunities

It is no coincidence that digitalization opportunities are most likely to be recognized and exploited at the level of profit-oriented companies. Bureaucratic routine processes can be handled more cost-effectively and faster by using digital technology and periodic reporting obligations can also be mostly fulfilled automatically by using software based on AI. The digital transformation of standardized processes mobilizes efficiency gains with decreasing opportunity costs. International business relationships can be made more agile and environmentally friendly through digitalization. People in many countries can communicate and work together in a networked manner via the Internet without incurring travel costs and emitting environmentally harmful greenhouse gases such as CO_2. New entrepreneurial fields of activity can be digitally structured and processed by combining AI program elements according to principles of modularity, which can be called up from the Cloud. Innovative forms of platform economy are reshaping the relationship between producer and consumer (Jaekel 2017; Daum 2017).

The social opportunities that are opening up around the globe as digitalization progresses are also diverse. More and more people can participate in the universally available knowledge that is accessible to them in the transcultural web environment. The widespread use of digital technology simplifies and accelerates the exchange of information both within society and internationally. The Internet opens up better opportunities for people to organize themselves and articulate their political will in social networks more quickly than is possible in conventional democratic election processes. The "Arab Spring" in 2011 was an impressive example of this—also in the sense that newly created digital public spaces can quickly disappear again.

For countries of the Global South, with their highly fragmented and exclusive social structures, the opportunities offered by digitalization are of particular importance. For many marginalized population groups, the Internet gives them the opportunity, for the first time, to organize themselves in order to represent their interests (Bolz 2017, p. 19). This has also a gender-political dimension, since women can overcome sociological barriers through technological means. Through digitalization, more people in developing countries than ever before are gaining access to information that enables them to exercise their democratic participation rights on a more substantive basis. The processes of hopelessly overburdened state administrations are accelerated by e-government, and "transparent bureaucracies" make the use of state funds more transparent. In addition, digitalization in countries of the Global South opens up the possibility of catching up very quickly or even getting into the fast lane through adapted technological leaps (Scherf 2017, p. 1). With the availability of digital communications technology, it is no longer necessary in developing countries to build up a nationwide wired infrastructure for analog communications at high cost, as was the case in the industrialized countries. A good example is the use of smartphones in some African countries, where network coverage has been established that is superior to that of some regions in Europe.

In more and more countries in Africa, Asia, and Latin America, the exclusionary mechanisms of the traditional financial system are being overcome with Fintech's, digital financial service providers that offer their services to those previously excluded from access to conventional commercial banks. In countries of the Global South, digitalized payment transactions at low transaction costs, even across national borders, are making everyday life much easier for many people; this is particularly true for groups of people in sub-Saharan Africa and Central America who urgently need to receive remittances from emigrated or refugee family members quickly. A study by the German Institute for Economic Research (DIW) shows how the use of "mobile money" has increased in recent years (Lehmann-Uschner and Menkhoff 2020). However, the digitalization of payment transactions also harbors new risks of dependency (Häring 2018).

In Ghana, an "Easy banking for everyone" system enables digital payments even without the Internet by using money cards with fingerprints (Albrecht-Heider 2017). In Paraguay, the company Tigo Money, in cooperation with other service providers, offers cashless financial transactions via cell phone almost everywhere. Poor people and those living in remote regions can use digital smart contracts based on blockchain technology to gain access to services such as microcredit or microinsurance, which in the analog world are usually not available to them due to the high transaction costs (Kemper 2017).

However, basic prerequisites must be in place to take advantage of the social opportunities offered by digitalization. One of these prerequisites is the provision of digital infrastructure. Without a basic nationwide digital infrastructure, the opportunities for digital use will remain limited, especially for people in remote rural regions. Whether poor people can use the Internet is also a question of costs for accessing the Internet, which in turn depends on the relevant national regulations. As long as the digital infrastructure is predominantly provided by private companies, the "digital divide" between rich and poor, between city and country, cannot be overcome.

Another important basic prerequisite for taking advantage of the social opportunities offered by digitalization is the "digital literacy" of people, to ensure that they can responsibly perceive the individual opportunities offered by digital change. "Digital literacy" means more than mastery of software and computer technology; it requires the acquisition of context-related technical, functional, and behavioral skills (Warschauer et al. 2010). "Digital literacy" must be integrated into elementary education and continue as lifelong learning. Without continuing education, learning the digital skills of today is not sufficient to be able to follow the rapid changes in the digital world; the newest digital tools are available in rapid succession or are expanded to include new functionalities. However, the deficient education systems of many developing countries are even less capable of such comprehensive "digital literacy" than those of many industrialized countries. The digital knowledge gaps remain unequally distributed around the globe, so that the "digital divide" threatens to become entrenched. Economic and political asymmetries that are created or reinforced by digitalization stand in the way of the global responsibility for autonomous developments in the countries of the Global South, which the international community of states agreed to in 2015 by adopting the Sustainable Development Goals (SDGs); Goal 10 of the SDGs expressly calls for reducing inequality within and between states (BMZ 2017b, p. 9).

There is no doubt that digital disruption opens up individual and social opportunities, but it also harbors considerable risks. Up until now, every revolution in human history has had its negative consequences; it would therefore be irresponsible or naive to believe that this should not be the case with the digital revolution. The innovation potential of digitalization is enormous, but the destructive potential of digitalization must not be underestimated. Established mechanisms of action become functionless, existing structures dissolve, and social norms deteriorate.

2 Digitalization Risks

The attempt to outline a risk scenario of the digital present and future must consider many facets. The consequences of digital change in the world of work are immediately apparent, as a conflict of interest between job security and digital innovations obviously exists. The effects of digital innovations are felt directly and usually very quickly in the workplace. In many areas, the digital working world will be a completely different one than the working world we are used to. Some occupational groups will be particularly affected by digitalization, while others may benefit from it, such as programmers, software developers or IT specialists. Jobs with simple qualifications and job characteristics will be lost in the digitalized working world, as these jobs can be quickly replaced by digital technologies. Medium-level qualification groups also lose importance to the extent that their activities can be formalized and standardized using digital technology. At the same time, constant availability and self-exploitation of the workforce are becoming cultural norms of the digital working world. To cushion the radical changes in the world of work and counter growing inequality, a new social contract is needed, supplemented by massive investments in lifelong learning (World Bank 2019, p. 125f). In industrialized and developing countries, digital change is triggering far-reaching changes in production structures. This is because the digitalization-induced increase in the capital intensity of production processes is leading to the relocation of manufacturing processes back to industrialized countries, as the locational advantage of lower labor costs in developing countries loses significance. According to World Bank estimates, up to two-thirds of the current jobs in the manufacturing industries of developing countries could be lost because of increasing digitalization and automation. This would particularly affect unskilled or low-skilled workers, who risk falling into a poverty trap (World Bank 2016, p. 23).

Less concrete than the consequences of digital change in the world of work are the longer-term social consequences of the digital revolution. On a meta-level, the tendency towards the "colonization" of many people's everyday lives by digital technologies can be observed (Greenfield 2017). Online, a comprehensive but invisible standardization of individual experiences is taking place that shapes social interaction and self-presentation (Villhauer 2018, p. 163). *Persuasive technologies* capture the attention of Internet users to control their general and buying behavior through personalized advertising. The balance of power between manufacturers, retailers, and end consumers is shifting because of these relationships changing form. Direct contact between manufacturers, dealers, and consumers is being replaced by the mediation services of international tech groups on their digital platforms, supported by virtual language assistants such as Alexa and Siri to control and handle the customer's purchasing needs. Services complementary to the supplier's product are imposed on the customer in a cost-saving way, such as printing a concert or train ticket. The "classic" market-based exchange mechanism—goods for money—is losing importance in the attention and data focus of the digital economy; the data and "clicks" of Internet users, which serve as the most important production and marketing tools for the digital providers, act as a medium of exchange. On the provider side of digital markets, there is a strong incentive to monopolize user data by offering Internet services apparently free of charge and to eliminate competition (Wambach and Müller 2018, p. 24 ff).[2] The regulated capitalism of continental Europe is being replaced by a digital capitalism that generates the highest returns and follows the laws of the strongest. Undisturbed by nation-state regulations, globally active digital service providers such as Amazon, Facebook, and Google determine the rules of the game in more and more economic sectors. Such processes run even faster in countries with weak governance structures. In the Global South, the relationship between the private sector and the state is therefore shifting to some extent, as the law of action can be determined by large tech companies. In addition, there is the question of data protection. How are citizens' rights protected? The high European standards are seldom reached worldwide, or a completely different model of transparency through technology is sought. There is also the intentional monopolization of digital services by the state, as in the People's Republic of China, where the Internet company Google is not (yet) allowed to operate its search engine; the monopolist here is the search engine of the Chinese company Baidu, which also manages the strictly censored Internet encyclopedia Baidu Baike.

More and more regimes are taking advantage of the opportunities for increased surveillance and censorship of digital media. Authoritarian governments, such as those in Iran or the Democratic Republic of the Congo, withhold unwanted information from their own people by temporarily shutting down the Internet or blocking websites. In the People's Republic of China, the digital infrastructure for surveillance and censorship is already the most developed and perfected; everything is prepared for the nationwide use of a digital points system of "social trustworthiness," which is already being tested in several provinces. This system (called "Social Credit System") is intended to reward "good people" and punish "bad people," with the

[2] In his dystopian novel "The Circle," Dave Eggers (2014) describes how the Internet group "Circle," which has taken over the business areas of its competitors Apple, Facebook, Google, and Twitter, is monopolizing the digital economy, providing all customers with a single internal identity and eliminating anonymity.

Chinese Communist Party determining what is "good" and what is "bad." As the National Development and Reform Commission of China reported, tens of millions of air tickets had been denied to "untrustworthy people." The example of China shows most clearly how digitalization provides the instruments of repression to the high-tech dictatorships of the twenty-first century.

Digitally enabled democratic participation is abused and corrupted by influential interest groups, influenced by Russian hacker groups such as APT28 and APT29, or disrupted by cyber troops of the Chinese People's Army. With *social bots*, automated computer programs, it is possible to manipulate voting decisions by spreading *fake news* in social and unsocial networks in a targeted manner, using fake accounts that simulate human identities. Liberal democracy is being replaced by digital democracy, in which the number of *likers* and *haters* on Facebook or the number of followers on Twitter and Instagram relativize majority decisions won in elections. Anonymous hate commentary and *influencers* determine the formation of opinion in the "net community," which only superficially forms a large community, but in fact consists of users who, in their isolation, join the pressure of the supposed public majority opinion. In addition to the loss of privacy in social and unsocial media, the language whose words are replaced by emojis is also impoverished. What we are dealing with here is a profound structural change of the public sphere, in which classical disciplinary and educational functions in the public sphere are being overridden. Not the best informed, but the loudest determines the discourse.

3 AI Algorithms Do Not Know Human Behavior

AI algorithms can be programmed to be capable of learning; with the method of *Reinforcement Learning* it is even possible to build a "reward impulse" into the algorithm, which takes effect as soon as the program has successfully completed a defined task. On the surface, this may look like a simulation of human behavior, but it is merely the result of programming.[3] Human characteristics cannot be assigned to AI. According to the current state of knowledge, the fear that further developed Artificial General Intelligence (AGI) could replace the individual as an autonomous subject in the future also seems largely unfounded. Although AGI may be the declared goal of AI research, AGI could only come close to or even surpass human thinking if it were possible to design AGI like a "digital neocortex," similar to the multisensory and motor part of the human cerebral cortex. However, it is

[3] A well-known example is "indolent cannibals," a computer simulation programmed by Larry Yaeger in 1994, in which "agents"—a multitude of small individual programs—were given abilities such as eating, moving, and reproducing. After only a few iterations, the digital "agents" cannibalized their offspring instead of mating and reproducing; see Kelly (1995) and Johnston (2008).

currently considered unlikely that this goal can be achieved in the near future, if ever.[4]

Even without the use of AGI, the negative effects of the ongoing digitalization of everyday life are already obvious. It is the increasing amount of time that people of all age groups spend on digital platforms—with network-relevant self-occupation, without any comprehensible benefit for them. It is the erosion of the concept of truth by the many supposed, often contradictory "truths" and unverifiable "facts" offered by the Internet. It is the malice, which is unleashed upon dissenters on large web portals. It is the brutally simplifying reduction of political discourse to 280 characters of a *tweet* on Twitter. With the permanent digital culture of affirmation in the public sphere of social media, privacy is lost—and with it the separation between the public and the private sphere as an essential achievement of Europe's democratic culture (Weidenfeld 2018, p. 13).

However, an undifferentiated *teclash*, a fundamental rejection of digital technologies, cannot be justified in view of the advantages that can be exploited in many areas through digitalization. Digitalization also offers people political and social opportunities for freedom. But an ethic is needed for the technologies of the future, a value system for the digital century (Spiekermann 2015; Hemel 2020). This is because without digital ethics the ongoing digitalization of living and working environments threatens to lead to an irreversible error and we could forget to ask who is freed by digitalization and what for? (Villhauer 2018, p. 169). AI techniques can facilitate everyday life in many ways, but it is a question of ethics to use AI techniques to construct "killing robots," *lethal autonomous weapons systems* (LAWS), and to export them to developing countries. For ethical reasons, the Federation of German Industries (BDI) demands a ban on Lethal Autonomous Weapons Systems and advocates an international treaty on the use of AI in military conflicts (BDI 2018). If digital technologies are developed at a rate faster than society's ability to understand and regulate their implications, then there is a risk of misuse, with potentially disastrous consequences. In this respect, the urgent need for responsible digital innovations marks the limits of digitalization. In any case, economies and societies will not be able to do without human activity in the future either. Committed employees will continue to be needed for creative jobs and for activities that deal directly with people. The AI of IT systems is capable of learning but can only ever be creative to a limited extent. Algorithms of *Generative Adversial Networks* (GANs) may produce images that resemble a model—for example a painting by Rembrandt—but GANs are not used to make creative art, but only to program a sequence of self-learning steps of the algorithms.[5]

---

[4] Raymond Kurzweil (2005) provided an exception to this assessment. Kurzweil sees singularity, an intelligence explosion, reached as soon as the digital neocortex of the AGI exceeds the thinking ability of humans.

[5] That pictures, which owe their existence to the calculated creativity of GANs, can reach the established art market was demonstrated by the auction of the AI portrait "Edmond de Belamy" on October 25, 2018 at Christie's in New York. It sold for 432,500 U.S. dollars.

AI has become an indispensable tool for analyzing huge amounts of data at high speed, drawing conclusions, and simulating decision-making processes. But AI does not have human consciousness and AI lacks human emotions. In the binary logic of algorithm-controlled AI, empathy is not a factor. Therefore, AI cannot be left to make decisions which have to be made by people who are globally responsible and capable of empathy. It is important to address the possibilities, limitations, and ethical aspects of digital transformation—if only to counteract false expectations that digital change can make societies and globalization more just. The gloomy vision of the technological age at the dawn of the third millennium, which Yuval Noah Harari sketched out in his novel “Homo Deus” (Harari 2017), is countered by individual responsibility for the humane use of freedom as the guiding principle of globalized digitalization.

References

Albrecht-Heider C (2017) Auch ohne Internet digital zahlen. Chancen. Das Magazin für Entscheider aus Politik und Wirtschaft, Spring/Summer, pp 37–40

BDI [Bundesverband der Deutschen Industrie] (2018) Künstliche Intelligenz in Sicherheit und Verteidigung. Handlungsempfehlungen der deutschen Industrie, BDI, Berlin

BMZ [Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung] (2017a) Die digitale Revolution für nachhaltige Entwicklung nutzen. Die Digitale Agenda des BMZ. http://www.bmz.de/de/mediathek/publikationen/reihen/infobroschueren_flyer/infobroschueren/Materialie312_Digitale_Agenda.pdf [Accessed 10 Jan 2019]

BMZ [Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung] (2017b) Der Zukunftsvertrag für die Welt. Die Agenda 2030 für nachhaltige Entwicklung. BMZ, Bonn

BMZ [Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung] (2019) Digitalisierung für Entwicklung Den digitalen Wandel gemeinsam gestalten. https://www.bmz.de/de/mediathek/publikationen/reihen/infobroschueren_flyer/infobroschueren/sMaterialie405_digitalisierung.pdf [Accessed 23 Sept 2020]

Bolz N (2017) Die Pöbel-Demokratie. Cicero 3:15–22

Datareportal (2020) Digital 2020 July Global Statshot Report https://datareportal.com/reports/more-than-half-the-world-now-uses-social-media [Accessed 6 Jan 2021]

Daum T (2017) Das Kapital sind wir. Edition Nautilus, Hamburg

Eggers D (2014) Der Circle. Kiepenheuer&Witsch, Köln

Greenfield A (2017) Radical technologies. The design of everyday life. Verso, London

Harari YN (2017) Homo Deus. Eine Geschichte von morgen, 15th edn. C.H. Beck, München

Häring N (2018) Schönes neues Geld. PayPal, WeChat, Amazon Go - uns droht eine totalitäre Weltwährung. Campus, Frankfurt a. M

Hemel U (2020) Kritik der digitalen Vernunft. Herder, Freiburg i. Brg

Hilser K (2018) Wer profitiert von den digitalen Dividenden? Der World Development Report 2016. In: Sangmeister H, Wagner H (eds) Entwicklungszusammenarbeit 4.0 – Digitalisierung und globale Verantwortung. Nomos Verlag, Baden-Baden, pp 83–93

Jaekel M (2017) Die Macht der digitalen Plattformen: Wegweiser im Zeitalter einer expandierenden Digitalsphäre und künstlicher Intelligenz. Springer/Vieweg, Berlin

Johnston J (2008) The allure of machinic life: cybernetics, artificial life, and the new AI. MIT Press, Cambridge

Kelly K (1995) Out of control: the new biology of machines, social systems, and the economic world. Basic Books, Boston/Mass

Kemper N (2017) Blockchain-Technologie: Funktionsweise und Potenziale für Entwicklungsländer. Entwicklungspolitik Kompakt (= KfW Development Research), 17

Kurzweil R (2005) The singularity is near: when humans transcend biology. Viking, New York

Lehmann-Uschner K, Menkhoff L (2020) Mobile Money treibt finanzielle Entwicklung Afrikas voran. DIW Wochenbericht, 22

Scherf T (2017) Megatrend Digitalisierung: Ein zentraler "Game-Changer" für Entwicklungsländer. Entwicklungspolitik Kompakt (= KfW Development Research), 7

Spiekermann S (2015) Ethical IT innovation: a value-based system design approach. Auerbach Publications, Boca Raton

Villhauer B (2018) Ethische Fragen der Digitalisierung. In: Sangmeister H, Wagner H (eds) Entwicklungszusammenarbeit 4.0 – Digitalisierung und globale Verantwortung. Nomos Verlag, Baden-Baden, pp 161–169

Wambach A, Müller H-C (2018) Digitaler Wohlstand für alle: Ein Update der Sozialen Marktwirtschaft ist möglich. Campus Verlag, Frankfurt/New York

Warschauer M et al (2010) New technology and digital worlds: analyzing evidence of equity in access, use, and outcomes. Rev Res Educ 34(1):179–225

Weidenfeld N (2018) Das große Beichten. Wie die digitalen Medien die Werte der Puritaner in die Welt bringen. Süddeutsche Zeitung No. 293, December 20, 2018

World Bank (2016) World development report 2016: digital dividends. World Bank Group, Washington DC

World Bank (2019) World development report 2019: The changing nature of work. World Bank Group, Washington DC

Digitizing the Humanities

Malte Rehbein

Contents

1 Humanities (As Academic Discipline)

In brief, Humanities encompass a set of academic disciplines whose object of study is human culture. The so-called "Digital Humanities" try to answer the question of how the ubiquitous phenomenon of digitization affects these Humanities. Digital Humanities aim at providing digitally-supported academic study (scholarship) of human culture. The digital collection, preservation and making accessible of artifacts of human cultures, i.e., cultural heritage digitization (see there), is closely related to Digital Humanities.

"Humanities" as a category of various scholarly disciplines was, among others, defined by Wilhelm Dilthey (1833–1911)—at that time in particular distinction to the natural sciences. Dilthey based his definition on a philosophical doctrine of the meaning and understanding of expressions of human life (hermeneutics). In this traditional understanding, the Humanities were characterized both by the objects and subjects of their investigation as well as by its specific methodology. Historical-hermeneutic procedures were regarded as methodologically guiding for the Humanities but were quickly supplemented by various empirical and non-empirical approaches. The spectrum of interests of the Humanities includes objects and questions from the fields of philosophy, history, theology, literature, language and

M. Rehbein (✉)
Chair of Digital Humanities, Universität Passau, Passau, Germany
e-mail: malte.rehbein@uni-passau.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_62

art, among others, as well as questions concerning the conditions of human life and human coexistence. In addition to the creation of scholarly knowledge sui generis, the Humanities also have a meaningful and orienting effect, embedding ethics and aesthetics, and provide a contribution to gaining insight and understanding in addition to explainable knowledge.

The Humanities are closely linked with other fields of cultural studies, the social sciences as well as human sciences, from which they cannot be clearly distinguished. In a comprehensive understanding, Humanities as "Arts and Humanities" ties in with the medieval tradition of the so-called liberal arts (*artes liberales*) and can still be found today in academic titles (Magister Artium, Bachelor of Arts, etc.). In terms of research and teaching facilities, the Humanities are usually based in colleges, universities as well as several, often publicly funded research institutions and societies.

2 Classification

There is no generally accepted definition of "Digital Humanities." Rather, the internationally accepted term "Digital Humanities" (DH) encompasses a collection of research practices in the Humanities which base their methodology at least partly on the application of computer-aided procedures, working techniques and tools. In principle, therefore, the term covers all aspects of the Humanities under the conditions of digitality. DH is understood as a research paradigm, at the core of which stands a gradual, digitally and computationally supported, formalization and explication of the objects, methods, and processes of research. This oscillates between the datacizing of research or the measurement of culture for a primarily quantitative calculus on the one hand and abstract modeling with regard to qualitative approaches up to the formal-logical calculus on the other. In doing so, the traditional hermeneutically influenced approaches to the objects of study are complemented by a "computational thinking" shaped by digitality. This, depending on the point of view, questions the tradition and supports, enhances or even replaces it. Current research practice shows two poles in this respect: to regard the procedures, working techniques and tools as a means to an end (the pragmatic pole) or to allow them to become a fundamentally different way of thinking (the paradigmatic pole).

Beyond merely applying them, DH includes the conception, development and critical evaluation of these computer-aided procedures, working techniques and tools as well as the study of the theories and models behind them. In this understanding, it is an academic (auxiliary) discipline at the interface between the Humanities and Computer Science, in a similar way as can be found in other branches of science (e.g., Business Informatics, Bioinformatics, Health Informatics, Geoinformatics). The field of DH often includes all forms of academic communication in the Humanities and dissemination of research results via new media.

In the general context of academia, DH can thus be understood as a pragmatic transfer of the Humanities towards digitization. In doing so, it must ensure that digitization does not create black boxes, but meets scientific criteria of transparency, traceability and reproducibility. At the same time, a paradigmatic transformation of the Humanities in general can be observed: The new methods transferred into them are changing the Humanities' disciplines up to the point of establishing new ways of thinking. This is induced by the conception of research activity in the Humanities as a kind of information processing.

3 History of Digital Humanities

The first approaches to a supporting practical application of computer technology for research in the Humanities date back to the 1940s. The work of Roberto Busa on a concordance of medieval Christian philosopher Thomas Aquinas' complete works, the *Index Thomisticus*, which comprises around 11 million words, is considered as ground-breaking in this respect. The decisive factor for Busa was the realization that the processing of larger and complexly interrelated data sets is not possible without technological support. The generalization of these considerations henceforth formed the leitmotif for computer-aided procedures in the Humanities. Since the 1950s, international projects followed, which took up Busa's ideas, especially in textual scholarship.

With the broad establishment of data centers in the 1970s, computer-aided methods had become increasingly popular in the Humanities. However, it was not until the advent of the Personal Computer in the 1980s and the associated greatly simplified access to computing capacity that they could find widespread application. In the wake of the invention of the Internet, mobile technologies, and the ubiquitous availability of computer technologies, since the early 2000s the so-called Digital Humanities (often also, e-Humanities) represent a research paradigm within the Humanities that is currently experiencing strong growth. Recently, large scale digitization efforts as well as current advancements in Computer Science have had a reinforcing effect on the further development and propagation of DH.

4 Characterization

The Digital Humanities seek to enable the Humanities, through digitization and computer-supported research processes, to formulate questions and work on topics that have traditionally been difficult to realize for pragmatic reasons (e.g., the difficult accessibility of source material) or scientific considerations (e.g., the size or complexity of the material). For this reason, new procedures are being developed at the same time as digitization is taking place to find relevant digital material, to organize, retrieve, and, above all, to analyze it and thus make it systematically usable either selectively or in the entire research process.

Thus, some central features and prerequisites of automatic information processing can be seen as constitutive for the Digital Humanities, some of which differ significantly from the traditional approaches of the Humanities. These include formalization by means of models that are representative for the respective research question and explicitly describe the theoretical assumptions and research data on which the analysis is based. The same applies for a mathematical or algorithmic operationalization of the research methodology. Here, it is central to computer-aided information processing to diligently consider the nature of the Humanities' research objects and data. In contrast to the natural and social sciences, the distance to the objects of investigation often does not allow researchers to collect data directly at the place and time of their origin, but rather forces them to construct them using artifacts that have been handed down by chance or in a targeted manner. The latter are characterized by their historicity and cultural contexts as well as a constructiveness based on complex and subjectively biased, intentional processes of human action. Thus, digital research data in the Humanities are often not objective, vague, incomplete, or ambiguous. Moreover, they are heterogeneous in terms of depth and density of structure as well as their degree of complexity. For many media data (text, image, sound, object), semantic structures are still often only latent and must first be recognized and explicitly formally described for digital information processing. It is one of the central research tasks of the Digital Humanities to develop appropriate concepts and models for this purpose and to link these to the theories and traditions of the Humanities on the one hand and to the perspectives of modern information processing on the other.

5 Research Designs

The research designs followed via Digital Humanities typically include analyses of content, spatial and temporal references, structure and interrelationships as well as networks and other relations from text-based or other media data. Exemplary questions are, for example, the determination of authorship and genre of unknown texts, the change of cultural phenomena across epochs and spaces or the social and historical network analysis.

Depending on the research goal and available resources, the actual analysis is preceded by a qualitative data exploration. Manual and (partially) automatic methods of text and image recognition, annotation (e.g., Part of Speech Tagging, Named Entity Recognition) and record linkage are used. A constant prerequisite is the existence of digitized data with relevance and representational capability for the research question. Digital-born data are also playing an increasing role. Often of decisive importance for research in the DH is the unrestricted access to these data in the sense of Open Access.

Important technologies of DH are still the so-called X-technologies (XML and related), relational and non-relational databases, algorithmic approaches, GIS as well as computer-linguistic text mining methods. In addition, there are, among other things, image recognition through computer vision, information visualization, virtual and augmented reality and new methodological approaches through simulation and reconstruction. Of increasing importance for the Digital Humanities are various approaches to so-called Artificial Intelligence whose potential and limits are being explored in various scenarios of Humanities research.

In the current research practice of Digital Humanities, quantifying approaches are emphasized. In particular, the number and distribution of selective phenomena that can be determined in research data are analyzed descriptively or inferential-statistically. This initially takes up the efforts for the systematic and formalizing collection and documentation of cultural artifacts on a larger scale, which have been established since the nineteenth century at the latest, as well as the quantitative approaches that became popular in the mid-twentieth century. Under new catchwords such as "Distant Reading" (as a method), "Culturomics" (as a research approach), or "Big Data" (as a general leitmotif), a nomothetic search for (ir)-regularities, patterns and structures is underway. Within these paradigms, in contrast to the traditional Humanities, the reading or viewing of the individual artifact is deliberately done without, in favor of the macroscopic perspective.

Hereby, the fundamental paradigmatic ideas of data science have also found their way into the Humanities. In contrast to many other areas of application of "Big Data," however, the sheer volume of data is not necessarily considered the leading factor in the Digital Humanities, but rather its complexity and special qualitative characteristics. This means that data analyses can only lead to valid knowledge and understanding in the Humanities if they are critically interpreted and fed back to basic theoretical assumptions.

In addition to the quantifying approaches, the Digital Humanities also support the qualitative approaches that are closer to the traditions of the Humanities such as hermeneutics. In these approaches, it is not the number of entities, but rather their specific nature that is the focus of the analysis. The resulting in-depth scholarly indexing of historical, literary and other cultural sources and their provision as an information system or digital edition has been a core task of the Digital Humanities since its inception. In addition, the epistemological value of modeling has been repeatedly emphasized. Especially in this area, standards have been developed since the 1980s and have been applied internationally in many contexts. Examples include the Text Encoding Initiative (TEI) for annotating digital or digitized texts and the CIDOC Conceptual Reference Model (CRM) for documentation in the field of cultural heritage.

Acknowledgement Digital methods in the Humanities require both a sophisticated understanding of information technology and an ability to be critical towards the underlying models and algorithms, their possibilities, and their limitations, all to be seen in the context of principle functions of research in the Humanities. Due to its position on an intersection between different areas of academia, research in DH often involves a high proportion of collaborative and interdisciplinary work. While the concrete operationalization of models and methods as well as the configuration of the support by the machine can be seen as close to Computer Science, hypothesis formation, theoretical embedding, contextualization, and interpretation of the results of data analyses are core tasks of the Humanities. In research practice, interdisciplinary cooperation thus involves

dialogical, ideally dialectical, processes that iterate between human thought and algorithmics. In the Digital Humanities, a modern understanding of science is expressed in a research culture that bridges disciplinary cultures not only between the Humanities and other branches of science and scholarship, but also within the various disciplines and perspectives of the Humanities. Digitization has a stimulating effect here in that the inherent necessity of formalization at a high level of abstraction can also lead to a common language.

Opportunities that arise from these new integral research perspectives as well as its embedding in the omnipresent central theme of digitization also justify the current public investment in the Digital Humanities. This includes the establishment of digitization and research centers, the creation of research data infrastructures, project and program funding, and the establishment of professorships and chairs, which will institutionalize the DH at both university and non-university level and create academic curricula. The identity of DH as an academic discipline includes the founding of specific journals (including Digital Scholarship in the Humanities, Digital Humanities Quarterly), professional associations under the umbrella of the Alliance of Digital Humanities Organizations (ADHO), and the regular organization of international conferences since the 1980s.

In the context of digitization, the Humanities are being enriched in a complementary way by various new methodological approaches and tools. In addition, in the medium term, a change in the objects of research in the Humanities can also be foreseen. In the future, human activity will increasingly appear to be linked to technology, which, in addition to social media, includes in particular data collections via the Internet of Things. Their providers will thus become new actors in society, shaping the conditions of human life and coexistence. Both these technologically induced social changes and the transformations of science are criticized. The criticism of the Digital Humanities is made as a scientific self-reflection, both from within itself and from the traditional Humanities.

To some extent, the critical reflection also mirrors the general discussion about the scientific value of Big Data, in which the reproach of decoupling from theory being formulated in the most enduring way. In addition, an orientation of research interests according to the availability and possibilities of data and tools and not vice versa is critically observed, which is feared to have a negative impact on the character and quality of research. The potentials of the use of computer technology and data analysis in the Humanities are also confronted with their limitations, which manifest themselves above all in dealing with the typical nature of cultural artifacts. This is where the criticism of Digital Humanities comes in, following on from earlier discussions about quantifiable research which states that human life and coexistence can be experienced especially through those things that cannot be quantified. Thus, Digital Humanities revisits long smoldering debates about the sovereignty of interpretation between hermeneutics, positivism, and behaviorism. Critics see the data-centered approaches as a purely descriptive determination of the structure of data sets and their inherent correlations at the expense of interpretation as an understanding and fathoming of basic conditions, options, intentions and causalities of human action. While skeptics often accuse the Digital Humanities of colonizing the traditional functions and tasks of the Humanities, their protagonists point out the complementarity of the methodological approaches and locate the Digital Humanities as an integral part within and not outside the Humanities.

For research in the Digital Humanities, legal issues, especially regarding access to and processing of data (copyright, rights of use and exploitation, privacy concerns, etc.), are often relevant and require clarification. Ethical questions are also increasingly being addressed. As in the digitization of cultural heritage, the question of sustainability and long-term preservation as well as research data management has not yet been conclusively clarified.

Cultural Memory

Andrea Schilz and Malte Rehbein

Contents

A. Schilz
Munich, Germany

M. Rehbein (✉)
University of Passau, Chair for Digital Humanities, Passau, Germany
e-mail: malte.rehbein@uni-passau.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_63

1 Culture, Heritage and Cultural Assets

The ‘cultural sector’ plays a dual role in the sense of an economic sector—a creating and a preserving one. Both sectors represent important parameters for hard and soft factors of sustainable social and economic development. Keywords are education, cultural tourism, attractiveness of the location. Under the generic term of the creative industry, exposed fields in the cultural sector can be subsumed under Fine arts, sound art, theatre, dance, literature. Although cultural preservation in the broader sense cannot be sharply separated from the field of cultural creation—music and theatre performances of historical works, for example, support the preservation of tradition—the culture preserving authorities can clearly be named as the memory institutions archive, library and museum. Just like the entire field of economy, this economically relevant field is also going through transformations shaped by digitisation. Therefore, cultural property and cultural heritage—objects of the institutions of memory—should not be discussed without reference to the digitisation of cultural heritage. The young academic subject of Digital Humanities is also related to this issue. As a bridging discipline between humanities and information sciences, it focusses on possibilities to support the digitisation of cultural assets for the humanities, cultural studies, history and art.

In a broad definition, culture encompasses everything that people aim to manipulate in a way, so that a transformation from a ‘natural existing’ state into a ‘cultivated’ one takes place. This was—hence the word stem from the Latin *cultura*—among other factors particularly momentous in the transition from nomadic to sedentary forms of life through the cultivation of arable soil and domestication of animals in the Neolithic period. As cultural property we can understand phenomena which are connected with this transformation process or the result of which (in the context of agricultural economies for example: a plough, a strain of crop, a form of settlement). In a more abstract approach, culture is also understood as the forms and processes of human life and living together. A distinction is therefore made between material and intangible cultural property.

The United Nations Educational, Scientific and Cultural Organization (UNESCO) defines a selected part of material cultural assets as ‘monuments: works of architecture, large sculptures and monumental paintings, objects or remains of archaeological nature, inscriptions, caves and connections of such manifestations (. . .); ensembles: groups of individual or interconnected buildings (. . .); sites: Works of man or collective works of nature and man and areas, including archaeological sites (. . .)’ (Art. 1 of the ‘Convention concerning the Protection of the World Cultural and Natural Heritage’, UNESCO 1972). This passage refers to physically tangible phenomena that potentially can be declared a ‘World Cultural Heritage Site’—for this they must be awarded that they are, inter alia, of ‘historical, artistic or scientific reasons of outstanding universal value’. Yet beyond this rare and coveted recognition, the spectrum of material cultural assets is much wider.

The immaterial cultural asset, on the other hand, denotes something not directly physically tangible. To declare the phenomena of this type of cultural property, which are considered to be particularly worthy of preservation, and thus to support their tradition worldwide, UNESCO adopted a 'The Convention for the Safeguarding of the Intangible Cultural Heritage' in 2003. It is aimed at practised, identity-creating traditions which 'communities consider (...) as part of their cultural heritage' (ibid.). As is it the case with physical cultural monuments, such phenomena (Intangible Cultural Heritage, ICH) are subject to strict selection procedures in order to be included in the respective UNESCO lists. These include so-called 'ICH elements' from the fields of customs, rituals, festivals, oral traditions, dance, song and music, cooking and crafts.

Both material and immaterial cultural assets bear witness to culture, which leads to the concept of cultural heritage. It has an objectifiable value: What is inherited is what is considered worth preserving. The principles on which 'cultural inheritance' is based are not dissimilar to those of inheriting in the economic and private spheres. On the one hand, objects are inherited that have an accepted material value (such as objects made of precious metals) or are linked to such (cadastre, title deeds). On the other hand, certain objects are considered worthy of inheritance in which forms of representation manifest themselves—in the private or bourgeois milieu, for example, the family album. On a macro level, religious and secular power is represented by means of sacred or profane buildings, in visible from afar (whose deliberate destruction, conversely, is intended to demonstrate that this very representation is rejected as 'inheritance').

In addition to clearly perceptible phenomena, there is another large field of cultural heritage: objects whose representative significance has been successfully negotiated by society and which therefore occupy an exposed, symbolically charged position—but without being visible in everyday life. This field includes objects that document technical progress and are therefore shown or kept in technology museums. This category of cultural heritage makes the educational aspect, which UNESCO includes in its name, particularly clear. The consensus on what is to be taught and learnt and which goods are to be handed down under the criterion of conscious preservation is subject to social premises. However, most of the preserved cultural goods are the result of accidental transmission. The status of such phenomena, which float freely in this pool of cultural assets not declared as 'tradition'—what exposes them highly to transience—can also change from time to time. For if we look at societies in space and time, value systems change, and with them cultural concepts and the consensus on what future generations should definitely come upon.

To what extent a decision to secure artefacts that are considered culturally is expedient, is in turn determined by external conditions. On the one hand, it can be determined by the fact that the social negotiation starts (too) late in relation to already created facts. One example is the history of objects of the industrial age that were destroyed before they could be preserved for posterity as valued relics of 'industrial culture'. In contrast to milestones of industrialisation (Watt's low-pressure steam engine) or prominent exhibitions (the Eiffel Tower), these were testimonies to unremarked production chains or proletarian everyday life. The fact that museification (museums of industrial heritage) and monument conservation (e.g. for model settlements) were applied to such goods remains rather the exception. In urban areas, factory and storage facilities are sometimes preserved through conversion, at least in their approximate external perceptibility. Here too, a revaluation, a changed concept of aesthetics, is one of the causes.

On the other hand, the preservation of a phenomenon can also be influenced by evaluations or actions depending on individuals (or even smaller groups) who behave in opposite ways to ruling paradigms. The entry of preserved objects into the collective memory, in turn, is conditioned by later social recognition. One example is the diaries of Anne Frank; they were first published under risk, but their value was only appreciated later on, when then socially accepted and anchored. But also, quite different characteristics of human culture are subject to categorically similar developments, such as the increase in value of design or pop-culture objects from the 1970s shows.

Lastly, however, the concept of what is declared ‘cultivated’ and therefore worth receiving measurable appreciation has shifted in the global context. Today, cultural goods are also accepted as such if they have their source in unsettled and/or non-agricultural or simply non-European societies—this was by no means always the case. Cult objects shown in the context of early ethnology, for example, were often only perceived as ‘exotic’ evidence of ‘savageness’ from a colonialist perspective. And something else has changed in the social consensus about what needs to be passed on. Testimonies are also being preserved as culturally valuable in a reflected manner, for the reason they are evidence of historical conditions marked by terror and organised mass murder. Exposed examples are memorials to genocides, which are on the UNESCO World Heritage List. For example, the Nazi concentration camp Auschwitz-Birkenau and the memorials to the genocide in Rwanda (Nyamata, Murambi, Bisesero, Gisozi).

There is a link between negotiating what is culturally valuable and is therefore valued as a cultural asset and what are considered suitable ways and means to bequeath cultural assets. Such storage locations for objects, which are called Cultural property, declared and designated as worthy of preservation, are classified as institutions of cultural memory. Each of these institutions has developed suitable organisational architectures and processes to adapt them to their key-object to be reliably passed on. But not only these manifestations and processes have been and are being adapted to the conditions of digitality, the objects themselves are also subjected to a transformation in this respect—for example through their (Retro-) digitisation and the associated new options for dealing with them, or through their novelty per se. Even sources generated in the digital realm, so-called Born Digital, are now ‘part of the tradition’. This suggests that task profiles in the field of cultural heritage are affected by digitisation in different ways, depending on their orientation, and that these profiles are variously designed.

2 Institutions of Cultural Memory

Institutions of cultural memory, or memory institutions, generally include libraries, archives and museums. As ‘museum-related types of institutions’ (Enumeration: Walz 2016, p. 129 ff.), which serve the concept of cultural memory can be named ‘Memorials for people and events; Historical places; Monuments with historical equipment; clerical treasuries; parks with reconstructions of past life; children’s museums; art galleries’ (ibid.). A certain special position have soil monuments which are a quantitatively and qualitatively significant component of cultural memory, although their presence is mainly merely latent: To perceive them requires prior knowledge. This kind of monuments is also institutionally connected, as the authorities are usually responsible for their identification and providing inventory and protection.

The task of all memory institutions is to promote knowledge and memory, mostly those of an organised group of people, e.g. a nation, and to preserve them over a long period of time. Here ‘long’ means a period, reaching out over the lifetime of individuals and generations, in some cases extending beyond the life of the organisation itself—potentially for eternity. It is precisely at this point that a special feature of the cultural memory claims its importance: In the long-term preservation of knowledge, the passing on of knowledge of meanings—possibly—can play a greater role than the passing on of the auratic original. Even its total loss potentially does not break the chain of transmission, if the corresponding knowledge content has been extrapolated beforehand and is preserved by consensus on a parallel path. A variant of this concept for the long-term preservation of knowledge, decoupled from the original, can be found again in the so-called Lots-of-Copies principle of the library. The objects that are collected and made accessible here are usually not unique, but rather multiple and, depending on the edition, approximately the same reproduced: Books are subject to multiple transmission—a principle that, in the history of ideas, especially scriptural religions use as a stabilising factor.

The aspect of preservation over long periods of time can also manifest in a completely different way: For museums and archives alike it is characteristic that their secured objects are unique—so-called unique specimens—or exist only in a few copies. A distinction is made between these two institutions of memory, however, in the way in which they combine knowledge and memory accessibility. In the museum (mostly representational) objects of the cultural heritage are kept ‘under glass’. In the case of the archive, the focus is on making the objects physically usable—often, but by no means exclusively—for research in focus.

3 Archives

3.1 Definition and Functions

The functional roots and development of an archive are closely linked to the respective institution for which the archive was created. It is therefore difficult to find a uniform, catchy formula for the term archive, which is used very widely and ranges from a private online photo archive to a state institution (cf. Ksoll-Marcon). The common feature is that something is collected and preserved that is unique (as a unique specimen) and is intended to survive time in an orderly fashion. A functional definition of an archive in the true sense of the word can be: 'An archive is there to secure, order, make documents accessible (. . .) and make them available for use' (Brenner-Wilczek et al. 2006, p. 13). Institutionalised archives clarify the questions of archive worthiness, capability and responsibility. In the case of official archives, they are usually subject to legal framework conditions, the parameters of which may be defined at state level, but also (supplementary and/or complementary) at federal state level.

3.2 History

Evidence of the systematic storage of official documents goes far back into history. Institutionalised archives are known from the Greek and Roman antiquity as well as from other cultures. In the occidental Middle Ages texts emerged (charters) that set or fixed laws, in which important facts for the respective recipients were certified (e.g. awarded professional privileges, the transfer of a fief, land ownership and customs rights). These documents possessed a high degree of binding force, comparable to today's documents on real estate ownership, for example. In contrast to today, however, there was no superordinate administration which kept mandatory registers. To avoid conflicts (which regularly dealt with economic, political or legal issues) or to preserve the legal status, every secular or religious rightsholder had to present the relevant charter physically, often via centuries. Thus, such documents accumulated, secured against theft and environmental influences, over long periods of time. With its growth, the introduction of an order and a system of classification increasingly became necessary, which enabled reliable retrieval in case of need. Thus, archives become instances, but above all also instruments of power. During the French Revolution, the use of archives became a civil right, with the guarantee of access to the manorial and state archives and thus ensuring equal information and equal rights.

3.3 Organisation

The orientation of archives is based on the organisation for which they provide archival tasks. For example, territorial archives include state and city archives; entities which maintain non-public archives are churches, universities or broadcasting companies; there is also non-public archive material from estates and from the media, associations, parties, trade unions, enterprises, families and aristocratic houses (especially such archives also own an identity-forming role). The organisational structure of public archives is based on their legally prescribed tasks of preserving archival records, protecting them, and making them usable). The concrete work that results from this for archives and archivists is the assessement of potential archival materials (appraisal), their indexing, ordering and recording, the provision of written or oral information, the scientific evaluation as well as the conservation and restoration of the archive material. In their internal order, archives often follow the so-called provenance principle, according to which archival documents are summarised and arranged according to their origin. The focus of the archival recording is on the specific archival material (the individual unit of archival material) or archival material related to its origin. It is also important to keep the history of the archival documents with them.

3.4 Collection Items

Archives can be typified on the one hand according to their carrier and on the other hand according to what they store. Here the difference between official and collection-oriented archives becomes particularly clear. The latter set themselves certain priorities, for example thematically (e.g. an archive for literature) or media-oriented (e.g. an archive for photography). Like museums, these archives are reliant on voluntary transfers or purchases to build up their collections in a targeted manner. In contrast, the archive is of an official nature—here the collection consists of material from defined areas of responsibility (e.g. the city archive for the city administration, the federal archive for the federal authorities) and so the material kept is thereby inherently categorised as worth archiving. The authorities are obliged to hand over their records to the relevant archive after a certain time. In compliance with legal requirements archives are used in various ways, for example by research (e.g. history and cultural studies), jurisprudence (e.g. heir determination) as well as certain professional groups (e.g. for journalistic research) and private individuals (e.g. for genealogical research).

4 Libraries

4.1 Definition and Function

The word library originally meant ‘bookcase’ in Greek. The possibility of removing these books points to the fundamental difference to the collection motivation of museums and archives: the principle of public lending. The ‘Public Library Tasks’ formulated by UNESCO are aimed at this principle and the educational mandate resulting from it. They are extensive and range from reading animation and reading promotion from preschool age to promoting knowledge of cultural heritage, understanding of art, scientific achievements and innovations, and maintaining the intercultural dialogue and support for cultural diversity up to securing the civil right of access to all types of state information (Public Library Manifesto 1994, p. 6). Education tailored to the target group is therefore the focus.

In contrast to the provenance principle of the archive, the library often follows the so-called pertinence principle in its collection order. From this perspective, libraries organise their collections according to thematic contexts. The library principle of order has an encyclopaedic aura, carried by the cognitive, intellectual world apprehension and penetration. This impression, together with the structural and architectural manifestation of a knowledge store, results in a high, communicatively effective symbolic charge. It refers to the privilege of educating oneself through the acquisition of knowledge and, following on from this, to the proverbial connection between knowledge and power, first formulated by Francis Bacon, ‘Nam et ipsa scientia potestas est.’—‘Because science itself is power’ (Meditationes sacrae 1597). To whom this offer is addressed has changed a lot over time.

4.2 History

Libraries as places for collecting written material go back to early history. The library of Alexandria is considered an outstanding example of antiquity. In the Western Middle Ages, monastic collections, later increasingly feudal collections, form the basis of libraries, which are expanded and first are accessible to small, privileged circles, but then—in the spirit of the Enlightenment—increasingly become public-bourgeois. This development will later continue with the addressing of a wider audience, which goes hand in hand with an increase in reading skills in the population. In the nineteenth century, with the help of compulsory schooling, the ‘decisive transition from restricted literacy on mass literacy’ can be regarded as completed after a development of around 300 years (Houston 2012). Towards the end of the nineteenth century, the public library based on the concept of the Anglo-American public library took shape.

Reading skills and educational movements are social factors that influenced the historical development of public libraries, but there were several others. In the scientific discourse, also mentioned are economy, urbanisation and industrialisation for the USA, the UK and France (Widdersheim 2018), democracy and emancipation are also cited as well as power and control (cf. ibid., pp. 269–272). The latter is also reflected in the historically determined characteristic features of the institution, 'three basic functions of libraries in social systems (. . .): the library as a repository for cultural or scientific memory, as cultic-lordly or hegemonic Place as well as a workshop and instrument for the promotion of human cognition' (Strauch and Rehm 2007, p. 51). The property mentioned here as a place of representation determines another characteristic of the library, its symbolic effectiveness.

The connection between knowledge and power was particularly used in the staging of national libraries, an institution that arises from the construct of the nation. The idea of the nation as a cultural, linguistic and 'genetic' unit (as opposed to the state) was adapted—from different traditions and motivations—by various European states and the USA in the nineteenth century. The library as an expression of cultural, economic and political capital was able to represent the national identity schemes and at the same time to consolidate the 'nation' itself internally through offers to the bourgeoisie. Much has changed here, too: Today there are 48 European national libraries (see The European Library), some of whose holdings are directly accessible and/or are shared via the aggregated virtual library Europeana—a nationally delimited, networked platform for digital copies from the cultural heritage sector (see European Commission, Strategy, Europeana).

4.3 Organisation and Collection Items

What a library collects are largely determined by its orientation. The spectrum of the collection items is as diverse as the library landscape is: There are public libraries and libraries that are privately or church owned. Accordingly, an initial categorisation of libraries can follow a territorial order, which includes national libraries or libraries of national importance—including state libraries—as well as state and regional libraries. National libraries have a special role because they are to collect comprehensively in relation to the state or language area. Libraries can orientate themselves on a wide range, but also on content-related focus, such as special and specialist libraries. A scientifically heterogeneous profile—depending on the faculty offer—can be found in university libraries. Children's and youth libraries, which are constituted by special departments in public libraries or in city and community libraries, play an important role in educational pedagogy, together with school libraries.

5 Museums

5.1 Definitions and Functions

A museum is a 'non-profit, permanent institution in the service of society and its development, open to the public, which acquires, conserves, researches, communicates and exhibits the tangible and intangible heritage of humanity and its environment for the purposes of education, study and enjoyment.' (International Council of Museums, Article 3, Statutes, 2007). Central to this definition is compliance with standards—which at the same time characterise the core tasks of museums—regarding 'permanent institutional and financial basis, mission statement and museum concept, museum management, qualified staff, collecting, preserving, researching and documenting, exhibiting and communicating' (Deutscher Museumsbund 2006, p. 7).

The profile of the museum is ambiguous, because for its self-image the complex and direct interaction of communication and reception is crucial. In contrast to archives and libraries, the explicit display of the objects to be preserved belongs to the core mission of the museum. But yet the act of selecting, the selection itself and the arrangement of the objects bare a meaning or a statement that sometimes goes far beyond those that the object in question had before it became an exhibit: What is exposed has a value—even if it should be negligible in material terms (e.g. exhibits in technical or open-air museums, which come from industrial or agricultural contexts) or if it is not immediately obvious (e.g. for some art installations).

The object, no matter of which genre, never stands for itself in a museum—it always also represents a consensus that the exhibited is something essential. What is declared as such depends, among other aspects, on political, religious and socio-economic conditions. These have a massive influence on how the paradigms of collecting are shaped. For this reason, the positivist character of collecting and the selective nature of museumisation harbours the danger of suggestion, or also, semblance. To encounter it in a reflective manner takes up a large part of the modern museum discourse. What is museumised, when and how, is decisive for today's curatorial concepts: What are the intentions and ideologies behind trendsetting and trendfollowing decision criteria, which both concern individual objects as well as entire fields of collection or research, and how and with what goal is the museum object communicated? At the same time, the history of the museum can be retrospectively described by means of these questions.

5.2 History

The history of the museum begins with the break-up of the theocentric world view in the Renaissance, starting from Northern Italy, and with the option of profane collecting, beyond 'cult or livelihood' (Sfedu 2006, p. 12). Highly elitist circles are socially and materially able to deal with patterns of world interpretation beyond the Christian canon, and to culminate their interest in art, technology, nature and various contemporary religious and profane fashions in chambers of wonder or art. The tendency to musealise things along contemporary discourses remains vital for the spectrum of the museum landscape.

In 1565 the Antwerp-born doctor and librarian Samuel van Quiccheberg sets up a first museological methodology. His approach was to structure the chamber of art (Kunstkammer) of Duke Albrecht V in the Munich Residence in six departments: Naturalia, Mirabilia, Artefacta, Scientifica, Antiquites, Exotica. Received as the 'earliest museologist in Germany' (Zäh 2003, p. 45), he also couples the educational idea of collecting—in contrast to the sensation (cf. Sfedu 2006, p. 15)—and wants the exhibits to be accompanied by a library. In Germany, for example, the Wunderkammer as an educational medium for the lower classes assumed shape in 1698 with the 'Kunst- und Naturalienkammer für den Realienunterricht' (Chamber of Art and Natural Resources for the teaching of real-life subjects) at the Francke's foundations in Halle (cf. ibid.), founded by the Pietist theologian and pedagogue August Hermann Francke (1663–1727). This form of a collection can be interpreted in terms of the history of ideas as a bridge between the baroque Theatrum Mundi and the early museum of the Enlightenment.

The first public museums of the eighteenth century were a new form of educational institutions based on privileged collections and aimed at a bourgeois-privileged public (cf. Sfedu 2006, pp. 17–31). In 1754 one of the first German museums, the Herzog-Anton-Ulrich-Museum in Braunschweig, was established. The first museum architecture of its own was the Museum Fridericianum in Kassel, which opened in 1779 (cf. Loers). In 1759 the British Museum was founded in London; 1793 the collections were installed in the Louvre in Paris and made accessible to the public. One year later the first technically oriented museum, the Musée des Arts et Metiers, was founded there.

In the nineteenth century, further museums, some of which were designed by architects who are still renowned today, were established through collections and foundations. In a complex interplay of framework conditions, weighted differently from country to country, the museum gains profile as a public institution in the nineteenth century—as a hybrid between civic interests and state representation and between feudal and entrepreneurial patronage (cf. Fliedl 2016, pp. 48–49). Characteristic of the collection policy of this era, whose paradigms have an impact well into the twentieth century, are both breadth (many fields and epochs) and concentration on the particular (aesthetics or curiosity as criteria). This is due on the one hand to the cabinet tradition described above, and on the other hand to the typical longing for a constitution of national identity.

5.3 Organisation

One way for museums to organise themselves is through associations. The role of a worldwide umbrella organisation is played by the International Council of Museums ICOM (International Council of Museums), with 119 national committees and 30 international committees (ICOM Strategic Plan 2016–2022). Among other things, the association aims to develop and disseminate standards and ethical guidelines for museums (Code of Ethics for Museums) and to promote the 'protection of world heritage and cultural diversity' (ICOM Germany). National umbrella organisations act as lobbies for museums and, in accordance with their self-image, promote adequate quality assurance. At the federal level there are, for example, in Germany, Austria and Switzerland, numerous other museum networks or associations that provide professional support and are effective in terms of communication.

The publicly visible side of any museum is its exhibitions. Not least because of a densely packed museum landscape in which the favour of the paying visitor is at stake, the spectrum is divided into permanent, special and travelling exhibitions. Each format requires planning and design tailored to the target group. Many factors play a role in the layout, the execution of which—provided that the resources are available—is split professionally. For example, curators conceive concepts, museum education advises on didactics, designers take over the (media) equipment or scenography. Specialists are also involved in museum work relating to the object as a physical unit. This area includes the preservation of museum goods under the aspects of prevention, conservation, or preparation and restoration (Museum Association, Standards for Museums 2006, p. 16). Large institutions sometimes have their own workshops or laboratories for conservation and restoration.

No less relevant is scientific research on the object and its accompanying sources, including questions of origin and provenance. The latter includes provenance research, the ethical principles of which were motivated, among other things, by Nazi looted art and which are described in the 'Washington Principles'' of 1998 in connection with a conference on assets from the Holocaust era. The 'Lost Art Database', operated by the Stiftung Deutsches Zentrum Kulturgutverluste, for example, attempts to do justice to this task. The topic of restitution also extends to so-called looted art (art stolen during wartime events) and to artefacts originating from colonial contexts. In France, for example, a report commissioned by the government at the end of 2018 was submitted for a critical examination of non-European museum objects (Sarr and Savoy 2018). Another aspect of research on museum goods is changes in attribution; they can have a strong impact on object valuations from a scientific and economic point of view.

5.4 *Collections*

The museum's tendency to expand the cultural narrative goes hand in hand with a historicisation of everyday life (the first open-air museum, the Skansen in Sweden, opened in 1891, with 'folklore' as its subject matter), media (the first film museums were in Munich in 1963 and Vienna in 1964), inventions (the Digital Computer Museum was founded in Marlborough, Massachusetts/USA in 1979; today: Computer History Museum, Mountain View, California) and pop culture (exhibitions on Kraftwerk, Lenbachhaus/Kunstbau, Munich 2011; Laibach, Tate Modern, London 2012; David Bowie, Victoria and Albert Museum, London 2013).

The illustrative recourse with links to the present shows how the heterogeneity of museums is becoming increasingly differentiated. A typological internal differentiation along their collection objects is correspondingly complex. It can first lead to a pragmatic three-way division into nature, culture and art museums, whereby certain technology museums 'which are not classified under culture museums' can be added as a fourth group (Walz 2016, p. 78). Sub-orders are suitable for art museums (antiquity, post-antiquity fine arts, applied arts) and for the broad spectrum of cultural museums. This also includes the somewhat diffuse field of 'special museums', which in turn can be narrowed down 'in two respects (. . .): to subject and place at the same time' (Walz 2016, p. 79, after Dippold 2007). This would include museums which, for example, address local production from a cultural-historical and possibly empirical point of view; as well counting within the scope of these 'special museums' are museums which are dedicated to a more in-depth cultural-historical presentation of certain (everyday) objects.

The categories sketched here are based on museum science discourses, not on a binding classification. The problem of establishing a classification also stems from the fact that contemporary social and political conditions and their changes affect the institution of the museum. Walz notes that 'finer classifications pose additional problems because the focus is on types that have since disappeared (e.g. dynasty museums in the form of multiple personal memorials, war museums as in and after the First World War places of mourning and remembrance) and some terms change their meaning (School museums in the interwar period as local history visual aids, today as historical sites of institutionally bound teaching and learning)'. (Walz 2016, p. 79).

References

Brenner-Wilczek S, Cepl-Kaufmann G, Plassmann M (2006) Einführung in die moderne Archivarbeit. Wissenschaftliche Buchgesellschaft, Darmstadt

Deutscher Museumsbund (2006) Standards für Museen. https://www.museumsbund.de/wp-content/uploads/2017/03/standards-fuer-museen-2006-1.pdf. Accessed 10.10.2019

Deutsches Zentrum Kulturgutverluste: Grundsätze der Washingtoner Konferenz in Bezug auf Kunstwerke, die von den Nationalsozialisten beschlagnahmt wurden (Washington Principles).

https://www.kulturgutverluste.de/Webs/DE/Stiftung/Grundlagen/Washingtoner-Prinzipien/Index.html. Accessed 10.10.2019

Dippold G (2007) Spezialmuseum — Geschichte und Wesen. In: Spezialmuseen. 13. Tagung Bayerischer, Böhmischer und Sächsischer Museumsfachleute; Neukirchen b. H. Blut, 15.-17. September 2004. München: Landesstelle für die Nichtstaatlichen Museen in Bayern, 2007. Museum-Bulletin 13:18–29

Fliedl G (2016) Das Museum im 19. Jahrhundert. In: Walz M (Hrsg) Handbuch Museum. J.B. Metzler, Stuttgart, S 47–52

Houston R (2012) Alphabetisierung. In: Europäische Geschichte Online (EGO) (Hrsg) Leibniz-Institut für Europäische Geschichte (IEG). http://ieg-ego.eu/de/threads/hintergruende/alphabetisierung/robert-a-houston-alphabetisierung. Accessed 10.10.2019

International Council of Museums (ICOM) — Statutes. As amended and adopted by the Extraordinary General Assembly on 9th June 2017 (Paris, France). https://icom.museum/wp-content/uploads/2018/07/2017_ICOM_Statutes_EN.pdf. Accessed 10.10.2019

Internationale Vereinigung bibliothekarischer Verbände und Einrichtungen (1994) Öffentliche Bibliothek, Manifest der IFLA/UNESCO [IFLA/UNESCO Public Library Manifesto 1994]. http://www.ifla.org/DE/publications/node/8768. Accessed 10.10.2019

Ksoll-Marcon M Archive, Version 09.05.2018, 17:32 Uhr. In: Staatslexikon 8 online [Staatslexikon der Görres-Gesellschaft]. https://www.staatslexikon-online.de/Lexikon/Archive. Accessed 10.10.2019

Sarr F, Savoy B (2018) The Restitution of African Cultural Heritage. Toward a New Relational Ethics. November 2018. https://restitutionreport2018.com/sarr_savoy_en.pdf. Accessed 10.10.2019

Sfedu T (2006) Museumsgründung und bürgerliches Selbstverständnis. Die Familie Leiner und das Rosgartenmuseum in Konstanz. Dissertationsschrift, Universität Konstanz. http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-22602. Accessed 10.10.2019

Stang R, Irschlinger A (2005) Bibliotheken und Lebenslanges Lernen. Kooperationen, Netzwerke und neue Institutionalformen zur Unterstützung Lebenslangen Lernens. Expertise zum aktuellen Stand. Bonn: Deutsches Institut für Erwachsenenbildung. https://www.die-bonn.de/esprid/dokumente/doc-2005/stang05_02.pdf. Accessed 10.10.2019

Strauch D, Rehm M (2007) Lexikon Buch — Bibliothek — Neue Medien. de Gruyter Sauer, Berlin

Walz M (2016) Grundprobleme der Museumstypologie. In: Markus W (Hrsg) Handbuch Museum. J.B. Metzler, Stuttgart, S 78–80

Widdersheim MM (2018) A political theory of public library development. In: Albright KS, Bothma TJD (Hrsg) Libri. International Journal of Libraries and Information Studies 68(4):269–289. https://doi.org/10.1515/libri-2018-0024 . Accessed 10.10.2019

Zäh H (2003) Quicchelberg, Samuel. In: Neue Deutsche Biographie 21, 4445 [Onlinefassung]. http://www.deutsche-biographie.de/pnd119331535.html. Accessed 10.10.2019

Referenced Websites

Bibliotheksportal des Kompetenznetzwerks für Bibliotheken, knb; ebd.: Bibliothekslandschaft, Öffentliche Bibliotheken. https://bibliotheksportal.de/informationen/bibliothekslandschaft/öffentlich-bibliotheken/. Accessed 10.10.2019

Computer History Museum., https://www.computerhistory.org/. Accessed 10.10.2019

Europeana. http://europeana.eu. Accessed 10.10.2019

Franckesche Stiftungen zu Halle, Die Kunst- und Naturalienkammer der Franckeschen Stiftungen, Geschichte. https://www.francke-halle.de/einrichtungen-a-1894.html. Accessed 10.10.2019

Lost Art-Datenbank. http://www.lostart.de/Webs/DE/Datenbank/Index.html. Accessed 10.10.2019

Convention concerning the Protection of the World Cultural and Natural Heritage, UNESCO 1972. https://whc.unesco.org/en/conventiontext/. Accessed 10.10.2020

Convention for the Safeguarding of the Intangible Cultural Heritage, UNESCO 2003. https://ich.unesco.org/en/convention. Accessed 10.10.2020

UNESCO UIS, Glossary: Museum, Definition http://uis.unesco.org/en/glossary-term/museum. Accessed 10.10.2020

European Commission, Strategy, Europeana. https://ec.europa.eu/digital-single-market/en/europeana-european-digital-library-all. Accessed 10.10.2019

The European Library. http://www.theeuropeanlibrary.org/tel4/. Accessed 10.10.2019 (Project Freeze am 31.12.2016)

Loers V (1994) "Neun Meter: zu sachlich, zu niedrig". In: 120 Meisterwerke, Höhepunkte einer fürstlichen Gemäldesammlung aus Renaissance und Barock. documenta und Museum Fridericianum, Kassel, S 26–31. http://archiv.fridericianum-kassel.de/kunsthalle_bisher.html. Accessed 10.10.2019

ICOM Deutschland, Ethische Richtlinien für Museen. http://www.icom-deutschland.de/schwerpunkte-ethische-richtlinien-fuer-museen.php. Accessed 10.10.2019

ICOM Strategic Plan 2016-2022. https://icom.museum/wp-content/uploads/2018/06/ICOM_STRATEGIC_PLAN_2016-2022_ENG.pdf. Accessed 10.10.2019

Deutscher Museumsbund (DMB), Geschichte. https://www.museumsbund.de/geschichte-des-deutschen-museumsbunds/. Accessed 10.10.2019

Österreichisches Museumsgütesiegel. http://www.museumsguetesiegel.at/shop/shop.php?detail=1. Accessed 10.10.2019

Digitization of Cultural Heritage

Andrea Schilz and Malte Rehbein

Contents

Like almost all sectors, institutionalized cultural memory is subject to major changes due to digitization. Three areas are of central importance here: digitization of cultural goods as technical digitization in the narrower sense, meaning analog-to-digital conversion of cultural artifacts as a prerequisite to the comprehensive concept of digitization; the creation of digital access to these digitized objects; and the transformation of processes in the management and performance of institutional tasks within the sector.

A. Schilz
Munich, Germany

M. Rehbein (✉)
Chair of Digital Humanities, Universität Passau, Passau, Germany
e-mail: malte.rehbein@uni-passau.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_64

1 Digitization in the Narrower Sense

1.1 Objects and Processes

The sector of cultural heritage encompasses memorial institutions, especially galleries, libraries, archives, and museums (sometimes referred to under the acronym GLAM). Here, digitization projects and enterprises deal with a heterogeneous spectrum of objects (physical artifacts, images, texts, audiovisuals). The typological range of pictorial sources include works of art (paintings, graphics, lithographs, etc.), commercial graphics (advertising graphics, posters etc.), or photographic sources (photographic prints, silver and collodion plates, film negatives, slides). Textual sources include mainly print media (books, newspapers, etc.), handwritten testimonies of the past, and serial documents (matriculation records, directories, files, etc.). So far, they constitute the largest part of digitization of cultural heritage resulting from respective inventories of libraries and historical archives. Age, origin, and context of the objects can generally vary greatly; the following exemplary categorical list of physical objects refers to the period from the sixteenth century up to the modern age: areas of art (sculptural works) and commercial art (serial manufactured works, e.g., ornamental porcelain objects; also, artifacts from hybrid contexts between crafts and objects of use—this includes, e.g., textiles from the field of fashion and traditional costume and areas from a wide variety of usage contexts that require further typification (craft or industrial equipment; means of payment; military objects, etc.). Audio sources (sound recordings) and audiovisual sources (film sources including silent films) form a further group. The most recent group of sources is "born digital"—sources that were originally created in the digital medium and lack a non-digital original.

The characteristic of heterogeneity not only refers to the media spectrum and media groups, but also to objects that in themselves are multi-medial (textual, pictorial, object-like). Heterogeneous characteristics which can be detected as inherent to the source may become particularly obvious in born digitals due to the huge variety of integral media components of, e.g., Internet publications. But analogous groups of sources also show clearly medially heterogeneous features. There are objects with both textual and pictorial characteristics (broadsheets, illustrated books and calendars, cartographic material, and cadastral works, etc.) and objects that have both object-like and textual and pictorial character—for example, an illuminated gospel book with a massive splendid binding. It is also due to these phenomenological conditions that the choice of digital representation and thus of the technique of digitization can be flexible. Buildings and sculptures respectively archival records and books give directions for three-, respectively two-dimensionality, but there may be ambiguous cases. Ambiguity can also exist with objects that carry text and image properties. Here, decisions are to be made about the quality of data acquisition as well as the choice and design of further processing (for example, text recognition and image optimization). Also, the acquisition process of metadata, an integral part of every digitization, is influenced by such decisions.

Digitization in the narrower sense is only omitted for born digitals. All others are subjected to digitization processes corresponding to an analog-to-digital conversion. The term retro-digitization is sometimes used to refer to this conversion of existing non-digital sources into digital representations—with reference to audio and audio-visual media and to source groups of a textual nature. Textual sources have a special feature with regard to potential steps in the digitization process. They can undergo a further stage of processing through automatic text recognition, which results in computer-supported searchability and analyses. As far as the source group of physical objects is concerned, three-dimensional capture—if the resulting 3D model is of suitable quality—can extend and improve the usability of the digital representation. A common feature of all types of digitized material is that potentially post-processing methods can be applied to them to increase or assure quality.

1.2 Objectives

Within institutions of cultural heritage, digitization can take on different functions. Its first purpose is to create a ubiquitous access to digitized representations of the collections' objects, which—in line with the zeitgeist of digitization as a general phenomenon—is designed to be as barrier-free as possible and independent of time and space but can also serve commercial interests. Second, digitization can aim at documenting cultural property as accurately as possible if the original is lost through natural decay or disaster and can therefore contribute to the preservation of cultural property. Third, for the practice of exhibition and scholarly evaluation, digitization can substitute a particularly valuable or fragile original and thus contribute to its protection, thus helping to resolve the immanent tension between preservation and accessibility. Fourth, due to their immateriality and arbitrary reproducibility, digitized material can contribute to bringing together artifacts virtually in different, historically verifiable contexts, without violating the consolidated and institutionalized structures of tradition. Finally, digitized material increasingly serves as a basis for computer-aided analysis, especially in the field of the so-called Digital Humanities.

In the planning and realization of digitization projects, as in all projects, the objective is fundamental to the result. Is its main focus quality or quantity (or both, assuming appropriate resources)? What is the primary target group—experts, interested visitors, or an anonymous audience? How the demands of research, education and didactics are weighted in the design and presentation depends on the project profile, the type of institution supporting the project and its resources. The premises and parameters under which digitization is carried out are in turn decisive for the establishment of processes.

1.3 Processes and Standards

Digitization in the field of cultural heritage is carried out in several stages and its workflows include many components that are only partially related to the actual technical process of digitization. These components, which concern the fields of project-, data- and knowledge management, set standards in the field of digitization of cultural heritage and can be so fundamental that they have a significant impact on specific processes in cultural heritage institutions. Specialized infrastructure is used thereby to automate and control processes of editing, provision, and archiving.

1.3.1 Management

In digitization processes, however different they may be, aspects of quality must be consistently observed. They are subject to professional standards, which functionally contribute to ensuring long-term accessibility, comparability, and changeability of data across institutional boundaries. The standards have different thematic structures, focuses and specialization.

An example of a holistically oriented presentation of such standards in form of a compact set of guidelines is the “Practical Guidelines on Digitization” of the Deutsche Forschungsgemeinschaft DFG (the self-governing organization for science and research in Germany). In comparison, the guidelines of the Federal Agencies Digital Guidelines Initiative (FADGI) of the United States of America, supported by the Library of Congress, include various publications, each of which covers a broad spectrum of media, including AV media. The Guidelines for Planning the Digitization of Rare Book and Manuscript Collections by the International Federation of Library Associations and Institutions (IFLA) set a more institutionally oriented focus with emphasis on project management. The Metamorfoze Preservation Imaging Guidelines (National Library of the Netherlands, National Archives) focus in greater technical depth on image criteria and digitization parameters regarding 2D materials (manuscripts, archives, books, newspapers, magazines, photographs, graphics, paintings, technical drawings).

Specifications of the standards can help to shape management structures. As an example of guidelines that potentially influences such structures, the FADGI’s “Digitization Activities – Project Planning and Management Outline” provides guidance on project management. The interlocking of management aspects in terms of planning or technical and methodological aspects is reflected in the organization of large digitization centers. Here, digitization process designs can be based on the DFG-guidelines, while complex data management is implemented with regard to publication and data storage, which also includes the operation of network platforms and long-term archiving.

1.3.2 Image Acquisition

The actual process of digitization in the narrower sense is the acquisition of digital images (as digital surrogates) of non-digital objects. Imaging processes used for material artifacts measure physical quantities of the objects they image—in the case of optical processes this is the light reflected from the objects. In digitization of cultural artifacts, digital cameras and scanners of various types are often used. Both are also suitable for creating three-dimensional digital representations. With regard to audio media, however, acoustic methods are used.

Digital imaging is characterized by the fact that the acquisition (via the sensor), storage (on a storage medium) and processing of image information are internal, thus eliminating intermediate steps; in analog photography, acquisition and storage are internal (on film), while image processing is external (film development). Technically and methodically, the same parameters must be observed in digital photography as in analog: white balance, ISO number light sensitivity, analog: of the film, digital: of the image sensor), focusing (adaptation of the optics to the distance between object and camera), duration of the aperture (degree of depth of field) and, depending on this, setting the exposure time. Furthermore, the focal length of the lens determines the angle of view. In addition, values that are essential for the digital images produced must be observed: color depth, resolution, target format and storage space.

Primarily, so-called RAW formats are available immediately after digital image generation, which contain all the image information recorded via the sensor. However, they are not used as digitization masters. RAW formats are proprietary, differ depending on the manufacturer of the camera, so that their processability depends on particular software products. In view of these imponderables, the DFG recommends the common, uncompressed raster graphics format TIFF (DFG 2016, p. 6) for long-term archiving, which can be read and edited with many proprietary and open-source programs. From these large master files, formats with lower storage costs can be transformed, depending on the application scenario (the compression format JPEG, for example, is well suited for web presentations).

The principle of scanners is based on line-by-line scanning. The original image is illuminated, and the reflected light is directed to an optoelectronic line sensor via a rod lens—which is designed to focus the reflected light and eliminate stray light. The analog light signals are converted pixel by pixel into digital signals by the analog-digital converter, while at the same time either the original or the sensor optics are moved step by step perpendicular to the sensor to the next line.

A common type of scanner is the flatbed scanner. Models up to DIN A0 size allow high-quality digitization of large-format objects (such as maps). Book scanners are equipped with a book cradle to protect the object. Models with line scanning and those with two digital cameras are used, which are adjusted to the light conditions and the size of the book pages at the beginning of the scanning process and scan a single or double page of the book per process. Digitization service centers supplement their range of services, e.g., with thermographic scanners for the digital capture of watermarks. Institutions particularly dealing with restoration issues sometimes use multi-spectral and infrared cameras to make phenomena visible below the surface of objects (e.g., signatures in paintings, palimpsests in archival documents) by using non-visible light wavelengths in the imaging process.

The technological choice of method strongly on the properties of the objects. Historically, or due to the manufacturing process, they may be fragile, fragmentary, weathered, and/or worn. Special procedures have been developed and are used for handling them. A technique that makes more clearly visible relief structures on flat objects (e.g., coins) is the RTI-method (Reflectance Transformation Imaging). It is based on series shooting with a fixed camera position but with variable lighting incidence. As a result, the digital image shows an optically highlighted reproduction of the surface texture (hence the term 2.5-dimensional digitization technology), which can bring advantages for research and presentation.

1.3.3 Digitization in 3D

The acquisition of digital representations in three dimensions aims at generating a solid model of the object with or without detection of the surface condition (texture). It is playing an increasing role as a method of cultural asset digitization in museums and other collections but is also used in scholarly practice in archaeology and historical disciplines, among others.

The range of imaging techniques for the digitization of cultural property in 3D is broad and ranges from photogrammetry to strip light scanning and technical more demanding laser scanning or LiDAR (Light Detection and Ranging) to computer tomography. The technological choice is first determined by the nature of the cultural property. A rough checklist contains basic information (material, size, weight, special features, etc.) such as color or gold frames, etc.), questions about the accessibility of the object (whether it is mobile or transportable, immobile, and possibly only accessible with the aid of equipment), and its fragility (whether special precautions are to be taken, e.g., stabilization during capture). Second, the purpose of digitization is decisive for the choice of the technology and the quality of its implementation, considering existing resources and results to be achieved (e.g., for target groups in the scientific or public sector). It is not only the imaging technologies that vary, but also their quality and costs. Post-processing of data needed depending on the technology and the expertise that may be required for this must also be considered—a very important factor, especially for highly specialized procedures. Computer tomography or X-rays, for example, are considered for special knowledge interests, namely when the interior of an object is to be researched or recorded non-destructively (e.g., a firearm mechanism). Planning and implementation can therefore vary greatly depending on the objective or the cognitive interest in the field of 3D.

1.3.4 Digitization of Text Carriers

In the large field of digitization of texts, a definition of objectives at the beginning is crucial. With regard to the project design, the weighing of quantity and quality factors must be considered together with a cost calculation. An important question here is the compatibility of the purpose to digitize large collections to the greatest extent possible and ensure high-quality access for scientific and other purposes. The definition of work processes, whose planning and realization is also far-reaching in economic terms, results from this determination of direction.

Projects that focus on high-quality reproduction first require a stringent selection of objects. The result should contain as much information as possible about the object itself—about its material condition, state of preservation and any paratexts, e.g., handwritten notes. Potential objects for this are often manuscripts whose digital representation corresponds technically and visually to "images of texts." One technical method of producing such digital copies is digital photography, under conditions of an environment that is adapted to scientific quality standards.

Large-scale and mass digitization projects have different priorities and are primarily aimed at serial sources or print media. The collections to be digitized are largely homogeneous material, such as books. The categories "large scale" and "mass" are sometimes used vaguely. In broad terms, large scale digitization aims to capture material within an organization quickly, but in as closed units as possible (collections, bundles, magazine series, and the like). The aim here can also be to provide at least rough information on the content, e.g., by means of structural content capture (Coyle 2006).

Mass digitization, on the other hand, is concerned with material that is not subject to any specific internal regulations or is not recorded according to such regulations. It concentrates on quantitative target achievement while adhering to quality assurance measures such as checks for correct reproduction and completeness. This category includes conversions on an industrial scale—in the case of libraries, for example, the non-selective digitization of complete collections (ibid.). Both digitization approaches, which are geared to large quantities, use scanning methods (e.g., scanning robots) to accelerate the process. Combining high quality with high productivity is the demand of both renowned houses and large digitization campaigns.

1.3.5 Information Recognition

Digitizing text-based objects usually encompasses a second step in addition to pure image digitization, i.e., the creation of a digital surrogate of the original as a raster graphic and its professional indexing via metadata (see below): the conversion of the text as image into machine-readable character codes for further processing such as automated information retrieval and opening up new accesses both for the general public (e.g., topic-based searches in and between literary works) and for scholarship (distant reading—mass text evaluation by means of algorithms under focused knowledge interests). To achieve this, two techniques are in use: manual text acquisition (keying) automated text recognition. In manual processes, to correct reading errors and oversights, the double-keying method is often used, in which two people transcribe the same text independently of each other. For quality control, any differences are automatically filtered out by means of Optical Character Recognition (OCR) (DFG 2016, p. 37)—the process for automated text recognition. Concepts for OCR already existed in the 1930s (Mori et al. 1992, p. 1030). With the introduction of the first commercial computer UNIVAC I in the 1950s, corresponding programs began to take shape (ibid.). In the meantime, OCR has become a widely established procedure (cf. Mühlberger 2011).

Quality of information recognition is measured by accuracy. It determines whether and how useful a captured text is for different purposes. The DFG-guidelines recommend a statistical procedure for collecting samples to check the degree of accuracy (DFG 2016, p. 35). It decides whether and how useful a captured text is for various purposes. For example, scholarly editions require the highest degree of accuracy while for mass digitization more moderate standards often are applied. In general, accuracy below 95% is not recommended (ibid.). For historical originals, recognition accuracy is often still a challenge due to poor document quality or fonts that are difficult to process. As a rule of thumb, the accuracy decreases with the age of the original. Another ongoing research topic is the recognition of historical handwritten originals, whose texts still can hardly be automatically recognized with reliable quality. Current approaches using machine learning as Handwritten Text Recognition (HTR) promise progress.

1.3.6 Accessibility via Metadata

Once the digitized material has been produced, further steps are required for the purpose of documentation and long-term preservation. Professional indexing of the materials is a central task of every collection activity in memory institutions, because only this can form the functional basis for searching, finding, receiving, and evaluating the materials: “Without the description and structuring of objects through metadata, digital archives, libraries and museums are of limited use for science” (HAB, Digital Humanities: Data Modelling and Metadata). For this reason, a source—whether it be books, archival materials, pictorial or artistic works or other objects—must be reliably recorded and systematically indexed. This process includes, on the one hand, the process of inventorying (which involves many specific steps, from the initial recording of the source to the acquisition of codes to the place of deposit to digitization) and the creation of finding aids. In each case, essential source characteristics are recorded and stored in a structured manner. Elementary information refers not only to the object itself (type, category), data on its creation (year, place, manufacturer), its origin (previous owners, purchase, donation), its condition (scope, dimensions), condition (damage), alterations (restoration) and much more, but also to its content; information on this can be deepened in various descriptive ways.

This applies to traditional, non-digital catalogs in card boxes as well as to machine-readable databases. As “classical” catalog data, they contain formally structured information about sources in a standardized form as “data about data.” Only this metadata allows efficient organizing, using, and maintaining of the data. To fulfill these functions, metadata is used for the targeted retrieval of relevant information, the organization or management of digital resources, the interoperability of information between different systems, the identification and unambiguous addressing of information as well as issues of archiving and (long-term) preservation.

Metadata can be categorized as descriptive data (used to locate information), technical data (used to assess the quality of objects), structural data (used to place objects in larger contexts) and administrative data (used to document administrative aspects such as access rights).

The quality of metadata can be improved by using so-called controlled vocabularies. They provide rules or restrictions for metadata entries via a subject-specific index of limited term sets. These vocabularies, which promote interchangeability and interoperability, can also vary in complexity—they range from simple lists of terms to hierarchical taxonomies and complex ontologies.

Digital variants of metadata help to achieve a new quality of metadata processing and use, often in combination with concepts for keywording or annotation, which serve search engine optimization (SEO) and user-oriented recommender systems. Established information systems such as the Online Public Access Catalogue (OPAC) are being developed and redesigned accordingly. The underlying databases, which contain the digital catalogs of each publication, are virtually linked in such a way that bibliographic metasearches of a very large scope are possible. The WorldCat system, operated by the non-profit organization Online Computer Library Center (OCLC; Ohio/USA), reports that it displays the holdings of over 10,000 libraries worldwide. Metadata standards are also a necessary basis for digital libraries and archives.

Standards for metadata are chosen according to scope and objective. Their complexity ranges from Dublin Core—a standard for the deliberately minimalist creation of metadata for a wide range of purposes—to highly specifiable models such as the CIDOC Conceptual Reference Model (an ontology from the cultural heritage sector) and the recommendations for encoding electronic texts in the Digital Humanities of the Text Encoding Initiative (TEI). EAD (Encoded Archival Description) and ISAD(G) (International Standard Archival Description (General)) are widely used in the archives sector; libraries often work with MARC (Machine Readable Cataloging), METS (Metadata Encoding & Transmission Standard) and MODS (Metadata Object Description Standard); for museums, LIDO (Lightweight Information Describing Objects) among others is a standard format.

1.4 Examples of Mass Digitization (Text)

Mass digitization quickly touches on issues of intellectual property, copyright, and rights of use—socio-political and ethical implications follow. Both everyday practice in the digital realm of the Internet and discourses on the digitization of cultural goods are influenced and shaped by this topic. A comparison of the projects Google Books, HathiTrust, and Deutsches Textarchiv shows how different objectives affect the results of mass digitization which are made available to the public in electronic, searchable form.

The US company Google Inc. has been operating Google Books commercially since 2002 as a collection of millions of digitized books. The service is subtitled as "the world's most comprehensive index for full-text books" in 2021. It can be assumed that it will be the quantitatively most comprehensive text-based information system, but no information on the exact volume is available. Google Books has been criticized for its cooperation with publishers and libraries, among other things. The latter is also controversial for licensing reasons, especially because of the problem of potentially circumventing other rights holders (including authors). In the discourse on Google Books, it is clear that this commodification can be a polarizing factor for some positions.

An association of predominantly US-American universities and libraries (including the Library of Congress) has been operating the HathiTrust digital library since 2008. It aggregates digitized material with metadata from Google Books, from the Internet Archive organization and partner institutions. According to its own information (2019, HathiTrust-Overview-Handout), HathiTrust has digitized ten million books, over three million of which are legally declared as public domain. HathiTrust says that it differs from Google Books in that many works are accessible here that are not accessible in Google Books, partly because of "differing rights determination processes." It also cites the quality of the search engine in terms of full text searchability and various search functions.

The German Text Archive (DTA) was funded by the German Research Foundation (DFG) from 2007 to 2016 and is located at the Berlin-Brandenburg Academy of Sciences and Humanities (BBAW). The aim is to create a reference corpus of German-language texts from the mid-seventeenth to the end of the nineteenth century; a compilation of different text types with metadata, balanced according to scholarly criteria, which as a collection promises a representative and reliable database for linguistic) research (for example, on vocabulary and word history). In 2021, 6,482 recorded works with 318 million tokens are indicated. The DTA explicitly distinguishes itself "from other extensive text collections on the Internet, such as Google Books, Wikisource or the Gutenberg-DE project. The DTA differs from these text collections in the careful selection of texts and editions, the very high recording accuracy, the structural and linguistic indexing of the text data and the reliability of the metadata." (DTA Project Overview). The use of the texts is regulated by Creative Commons licenses.

1.5 Examples in the Field of Audiovisual Media

Another group of sources that are subject to digitization projects are analog sound recordings. Since sounds are transient, there is the only possibility in their recording to document them as authentically as possible. Sound documents have been known in analog form since the end of the nineteenth century and form an indispensable source base for disciplines such as history, history of language, ethnology, ethnography, and musicology. Common media were wax cylinders, records (mechanical principle) and tapes (magnetic). Due to the technical repeatability of passages, such sources—also in analog converted form, on magnetic tapes—enable the creation of coded transcriptions that can also record paratexts such as intonation, speaking pauses and dialect using notations. However, digitization (also) offers a clear advantage in terms of content analysis. It is based on the access principle of digital media, without delay (e.g., by fast forwarding and rewinding tapes) at any point and being able to repeat sections as often as required without causing wear and tear on the carrier medium or its copy. Acoustic processes are often used to digitize analog sound documents (audio digitization by means of analog-digital conversion), but optical recording by scanning the sound carrier surface can also be used for records (non-contact digital imaging) (Haber 2011).

The group of analog audiovisual media (AV media) includes film recordings of a private nature (e.g., Super 8 and video formats), from television broadcasters (16 mm) and the silent and sound film sector (35 mm). Preserving the latter type of source requires enormous resources, which is primarily due to the carrier material of the medium. Before acetate became established around the 1960s, celluloid (nitrofilm) was used. Its cellulose nitrate base burns autocatalytically and is extremely flammable. Conservative storage requires "a room temperature of at most 7°C with 20% to 50% relative humidity; for color films it should even be -10°C" (Archives in the Rhineland, archive advice). In addition, the carrier material itself decomposes to the point of pulverization (cf. ibid.).

The example of the German film heritage shows the situation: permanent losses through chemical decomposition and occasional losses through fires (1945; largest part of the Reichsfilmarchiv; 1988: film magazines of the Federal Archives) lead to an estimate of up to 100,000 lost films; only around 15% of German silent films have survived (Koerber 2014).

Digitization under commercial parameters (DVD, streaming formats) is—although it means that there are many copies of a film—not a solution, as the quality of the original is considerably below the standard (ibid.). For a systematic long-term preservation of films, original negatives, well-preserved copies in full resolution as well as master files must be archived and cared for in terms of conservation (ibid.). For archiving-oriented digitization, the guidelines "Preservation and Reuse of Film Material for Television" of the European Broadcasters Union (EBU) recommend a digital reproduction system (film scanner) with a resolution of at least 4000 pixels across the line or 12,750,000 pixels per frame for a 35 mm film (Fossati 2009, p. 76).

1.6 Born Digital

Archives and national libraries face the challenge of preserving objects that were not created in analog but already digitally. This includes e-publications and e-books from both the scientific and the non-scientific area. Documents that are generated via "administrative action through the use of e-government processes" (BMI, e-government) should also be mentioned as a source type. For example, the German "Law for the Promotion of Electronic Administration" issued in 2013 is binding for public administration that has to promote electronic services with regard to "electronic evidence (. . .), payment in administrative procedures (. . .) electronic filing and of replacing scanning" (ibid.). Also involved are "publication obligations through electronic official gazettes, (. . .) documentation and analysis of processes, (. . .) provision of machine-readable databases by the administration ('Open Data')" (ibid.).

The Born Digital area opens up a new field for future historical sources, for whose research a modified form of source criticism is necessary. This basic method of historical work in the historical and cultural studies asks about the authenticity of sources, their temporal stratification, whether they convey objectively correct information, and who wrote or changed them when and with what intent. In a special, because new way, Born Digital sources that originate from the field of web archiving must be subjected to such tests.

Web content is in a special area of tension. On the one hand, it will serve as the primary sources of the future and therefore urgently need a consistent backup that is adequate for research. Because statements of the present time—in everyday, scientific, artistic, political, and other respects—are not reflected in any other medium as broadly as in the Internet. On the other hand, websites are ephemeral per se (if a server goes offline, pages and permalinks are deleted), inconsistent (ongoing changes by operators and—since the introduction of Web 2.0 technologies around the turn of the millennium—by users), technically heterogeneous (content in text, image, film, audio, animation, form, game, etc., each in different formats and applications, which in turn are subject to transience through obsolete software), and they stand as Phenomena that are volatile in terms of content (invalid links on a page) with many others in an unstable structure (change in the content to which the links point).

The saving of content can, e.g., be addressed via the WARC format (Web Archive), periods of web pages can be created automatically, and the sheer mass of data must be countered with both the creation of archive storage capacity and selection. But solutions must also be found with regard to questions of rights and data protection as well as archiving-worthiness, -ability and -responsibility. At the beginning of the third millennium decade, the field of web archiving offers plenty of space for discourse, research approaches and developments.

1.7 Sustainability

The long-term preservation of digital objects is a challenge in view of the rapid technological change in hardware and software, data carriers and file formats. Technical and organizational problems concern the storage media, algorithms, and software (which transform and process stored data), the hardware (which is required to read the storage media and control the software), output devices and the semantic knowledge for interpreting the data and ensuring that they cannot be changed (question of integrity).

Factors that promote sustainable long-term archiving in the digital world include redundant storage of data in multiple, spatially separated copies, compliance with international standards, regular migration of data to current file formats, storage media or hardware environments, analog backup of digital data via microfilming and the preservation of computer systems along with ensuring their functionality. The latter could be a decisive factor for future generations of researchers. Data—as sources and cultural heritage of the future—will only be accessible using technology that is then outdated. To program emulators that can do this, "hands on" practice is required to retain sufficient knowledge of hardware, software, and operating systems. This knowledge can help to pass on digital cultural assets that began to manifest themselves increasingly around the middle of the twentieth century. One approach is the museum collection of computers of all categories (e.g., mainframe computers, workstations, personal computers, home computers, mobile devices) and their peripherals, but also of manuals and relevant literature as well as spare parts.

In addition, the systematic consolidation of competencies is trendsetting. This is where *nestor* (Network of Expertise in Long-Term Storage of Digital Resources in Germany) comes in as an example, a "cooperation network with partners from various areas that have to do with the topic of digital long-term archiving" (*nestor*). Institutions, experts, and project sponsors exchange ideas in a network. They share information, tasks, and developments of standards; synergy effects are promoted, and strategies are discussed.

2 Digital Accesses

Probably the most obvious effect of producing digital copies of collection objects is their systematic publication on platforms accessible at any time via the Internet. The fact that there is usually no primary restriction on the group of users and that use via direct reception or research usually costs nothing contributes to the resonance. Further uses, especially of data that is not in the public domain or not published under Creative Commons licenses, often require inquiries or the acquisition of rights against payment.

Many memory institutions follow this path of a parallel offer in the digital world and, as digital archives, libraries, and museums, fulfill the task of making objects accessible—as in the analog world—in different forms. In the case of frequently requested archival documents, the digitized version takes on a protective function for the original and can thus help to resolve the immanent tension between preservation and providing access on behalf of the archives. The library can offer much easier access for its lending activities and the museum can establish viewing options that would be considered unthinkable compared to the analog original—for example, a close-up view at zoom level. The organization in databases ensures more efficient searches, structured and networked data allow the investigation of far more complex contexts than would be possible without these means.

Digital archives, libraries and museums that are linked to existing institutions are parts of the respective institution—not an institution in themselves. This characteristic can rather be attributed to virtual archives, libraries, and museums. In terms of the range of exhibits, such platforms have the nature of networked systems—they do not display a self-contained data stock, but aggregate data from various databases. However, a categorical assignment to one of the three memory institutions mentioned is by no means always possible, as two examples of large projects show.

The "Europeana Collections" platform (funded by the European Union), which has been set up since 2009, claims to contain "57,608,278 works of art, artefacts, books, videos and audios from all over Europe" in October 2019; it "(works) with thousands of European archives, libraries and museums and (shares) cultural heritage for the enjoyment of art, education and research" (ibid). The full version of the "Deutsche Digitale Bibliothek" (DDB; financed by the Federal Government and the German federal states) has been online since 2014, which "as a central national portal (...) is intended to network the digital offerings of all German cultural and scientific institutions with one another" (DDB). Any institution providing data can curate, as can curators who are interested in presenting the objects in the DDB in the context of an exhibition theme" (DDB).

For museums as physically accessible institutions, digital representations can be used in many ways. The dominating technologies in museums are differentiated into location-based services (location-dependent offers, by means of which information or services are selectively provided to the end user) and services that can be used independently of location. The latter not only include services that are used after (social media, recommendation systems) but also before (website, marketing) the museum visit.

The digital museum can be perceived from both perspectives—following up a real visit (possibly with regard to exhibits perceived as particularly interesting) and preparing for it. To a certain extent antipodal attitudes towards the latter can be roughly sketched out by skepticism on the one hand, in the context of a rather critical position with regard to such translations into digital space: (paying) visitors would refrain from a real visit to the museum (which is cost-intensive in operation and accountable to governmental, non-governmental or private sponsors) because they would then already be informed about exhibits and content. On the other hand, in contrast, digital museums are considered to be a stimulating, promotional and therefore commercially beneficial measure to encourage physical visits.

Even in services linked to the physical museum (including guide media), 2D and 3D digitized material is used for various approaches to knowledge transfer (education, edutainment, gamification)—especially for multimedia applications in exhibitions. Mobile applications are sometimes combined with Augmented Reality (AR), i.e., an additional information layer is placed over the display of the mobile device (smartphones, tablets) via interaction with the real environment, on which texts, images or digital copies, animations (also, 3D models) as well as further information can be displayed. Fixed installations, on the other hand, which (also) work with digital captures in 3D, are interactive virtual environments (Virtual Reality, VR). This is the simultaneous representation and perception of reality and its physical properties in a real-time computer-generated, digital, interactive, and virtual environment. 3D models of cultural assets up to entire architectural spaces can be used here.

3 Transformation Processes

Just like the media spectrum and the way of access, the image of memory institutions has expanded and changed in the way they perceive themselves and others—two positions are outlined here as examples. One reflects the concrete change in an archive that looks back on a centuries-old tradition. The other is formulated under the term “Museum 4.0” in a project and a discourse position. Both are equally concerned with innovative approaches to the museum as an institution of memory.

An example of strategic digitization in the field of archives is the archive of the diocese of Passau, which since the early 1990s has been systematically digitally indexing and publishing, especially in parish registers (church records; baptismal, marriage, and death registers). As there is a strong demand for access to these sources—including for scientific purposes in medical history, historical demography, migration, and onomastics and (lay) genealogy—digitization is also changing the work of archivists and users. In short, digital archives have potentials to offer technical and scientific advantages for researchers and economic advantages for users, as time and money can be saved by avoiding physical presence (cf. Wurster 2012). Another aspect of the transformation is the cross-institutional virtual merging of archive holdings. For example, the digitized sources mentioned above are accessible via the Matricula Online platform, which contains comparable registers from Austria, Germany, and Poland. It is funded by the association “ICARUS - International Centre for Archival Research” (Vienna), which consists of 160 institutions in Europe, Canada, and the USA; the infrastructure is supported by the European Union, countries, and dioceses. Comparable projects exist in other archival fields with similar objectives.

In the field of museums, the three-year project “Museum 4.0 - Digital Strategies for the Museum of the Future,” which is funded to the tune of five million euros and is headed by the Prussian Cultural Heritage Foundation, can be mentioned as an example. It sees itself as a “visionary pilot project in which innovative applications of digital technologies for museum work are developed and tested in a shared virtual space. The focus is on the topics of mediation, communication, interaction and participation" (press release of the SPK, November 2016). Partners come from various museum sectors and from the fields of technical, cultural, and natural history. Here too, the aim is to "link collections (. . .) in virtual space in a completely new way and make them accessible in an innovative way" (ibid.).

Innovation, participation, and digital and virtual space are also crucial to the conceptual approach of the project "Museum 4.0," which could be shortened to a "semantic equation (. . .) which would read: M(useum) + M(INT) + M (arke) = Museum 4.0" (Henkel 2017). "MINT" stands for teaching and study subjects resp. professions in mathematics, information technology, natural or technical sciences in Germany; "Marke" translates as brand—in this context it might also be trademark. Here, the change in meaning of the increasingly diversified museum in its factual capacity as a "mass medium" (ICOM 2016, summary) is targeted, or rather, that this requires reactions in the sense of "museological engineering" (cf. ibid.). Parallel offers from science centers to flag ship stores—see the above formula—make use of "the vocabulary, grammar and scenography of museology" (ibid.)—not least this makes "conceptual further development of the museum" inevitable to be able to develop and assert "social relevance (Nina Simon) [and] integrative added value in the sense of a third place (Ray Oldenburg)" (ibid.) for the museum in competition.

At this point, a project can be referred to that can be read as a horizon of action that is offered under the conditions of digitization. The Humboldt Lab Dahlem (2012–2015) relied on experimental approaches in 30 sub-projects for exhibition planning at the Ethnological Museum and the Museum of Asian Art (National Museums in Berlin) in the Humboldt Forum with the involvement of the audience in seven "rehearsal stages." For example, the interactive, cross-house installation "Gedankenscherz" (thought joke) (reference to Gottfried Wilhelm Leibniz's text "Drôle de pensée" from 1675 on the electoral-royal chamber of art and natural produce), in which the use of digital means in combination with modified digitized cultural assets (contemporary drawings) was implemented (Büro Focus + Echo). Visitors were able to use gesture recognition to successively open up aspects of the "Kunstkammer," "the more subject areas were uncovered, the more the multi-perspective interplay was opened up of animation" (Pinkow 2013).

4 Criticism

Digitization—in the comprehensive as well as the specific sense of cultural heritage digitization—creates a dynamic field of tension for memory institutions, the (previous) effects of which are still being explored at the beginning of the 2020s. A first associative question that arises with the surrogate character of the digitized cultural heritage and its availability via the Internet is that of the irreplaceability of the physical space of experience. Here, a preliminary conclusion for the memory institutions is different in each case.

In the case of archives that are not primarily dependent on visitors anyway, digitization tends to have a relieving function for archivists and archival records alike. For libraries, the functional change seems at first glance to have similarly positive consequences for personnel and material resources. However, the library—unlike the archive—has an explicit educational mandate aimed at the general public. This is borne by the physical object of the book and the lending function. Since the emergence of public libraries, these decisive complementary criteria have justified the presence of those many physically available, accessible rooms that are also used specifically for accompanying activities. The replacement of the tangible book by its virtual counterpart therefore influences the profile of this memory institution. Possible prognoses are that the competence to provide specialist information, support media education and offer a social space will move to the fore in comparison to the traditional core function of lending books.

Museums find themselves in an ambivalent situation—they are popular but under pressure. Privatization is no longer the exception; economic requirements promote programs which aim to provide an experience value or leisure facilities. To place the role of digitization in this context, it is worth taking a brief look back. A milestone for the broad impact of electronic information processing was the Personal Computer. In the early 1990s, (expensive) hand-held scanners offered the first opportunity to process images with the help of computers, along with easy-to-use writing programs. Technical development was then rapid, with the World Wide Web penetrating society as a publication and communication medium at about the same time. At the turn of the millennium, digitization had become a formative factor in many areas. However, a certain dichotomous tendency was already apparent in the 2000s with regard to its adaptation. Affirmative reactions were soon perceptible in the field of large memory institutions and are now evident in projects that are impressive in terms of both quantity and quality. But these are not representative. In a sense, they reflect the privileged tip of the iceberg, while on the other hand, there are many small to medium-sized institutions that fall far short of what is potentially possible and still stick to their homepage or web presence. Their resources are sometimes massively volunteered or supported by associations (similar to some privately owned archives) and only irregularly can specialized tasks in the field of IT be performed. However, they too come under at least latent pressure to compete in the field of up-to-date digital presence—via museum marketing discourses, corresponding implementations elsewhere and the expectations of the public evoked by these. Here, publicly supported or sponsored accompanying measures are a means of choice to have a constructive influence and to help these memory institutions to use digitization of cultural assets in a self-directed way.

The global COVID-19 appeared to hit the cultural sector in general specifically hard. However, like in many other sectors, digitization serves as a technical foundation for an alternative to the physical visit of cultural heritage institutions and to access their collections during prolonged periods of lockdowns. Consequently, digitization gains even more momentum due to the crisis and it continues to increase its importance.

References

Coyle K (2006) Mass digitization of books. Preprint. J Acad Librariansh 32(6):2006

Fossati G (2009) From grain to pixel. The archival life of film in transition. Amsterdam University Press, Amsterdam

Haber C (2011) New optical technologies for preserving recorded sound: experiments with the Smithsonian's earliest recordings, 18801890 (Lawrence Berkeley National Lab). Smithsonian NMAK, 14 Dec 2011. History, Slide 4/46

Henkel M (2017) Museum 4.0 — zwischen Museologie, MINT und Markenbildung. Von der Weltausstellung zum Science Lab. Handel — Industrie — Museum, Tagungsband zur Jahrestagung von ICOM Deutschland 2016, ICOM Deutschland — Beiträge zur Museologie 6:41–52

ICOM Deutschland Jahrestagung: Von der Weltausstellung zum Science Lab. Handel — Industrie — Museum. 6. bis 8. Oktober 2016. Berlin, Deutsches Technikmuseum. Zusammenfassungen der Vorträge. http://www.icomdeutschland.de/client/media/601/abstracts_2016_online.pdf. Zugegriffen am 07.10.2019

Koerber M (2014) Warum Digitalisierung Filme nicht rettet. Interview mit n-tv (Andrea Schorsch, 06.02.2014). http://www.n-tv.de/technik/Warum-Digitalisierung-Filme-nicht-rettet-article12085591.html. Zugegriffen am 07.10.2019

Li N (1991) An implementation of OCR system based on skeleton matching. Computing Laboratory, University of Kent at Canterbury, UK, Canterbury

Mori S et al (1992) Historical review of OCR research and development. Proc IEEE 80(7): 1029–1058

Mühlberger G (2011) Digitalisierung Historischer Zeitungen. Aus dem Blickwinkel der Automatisierten Text- und Strukturerkennung (OCR). Z Bibliothekswes Bibliogr 58(1):10–18

Terras M (2008) Digital images for the information professional. Routledge, London

Wurster HW (2012) Das archivalische Erbe unterwegs zur Online- Wissenswelt. Vergangene Wirklichkeiten, Digitalität, Archivwissenschaft und die gegenwärtige Geschichtsforschung am Beispiel des Bistums Passau. Vortrag zur Präsentation des Tagungsbandes "Die virtuelle Urkundenlandschaft der Diözese Passau", Wien Kulturgutdigitalisierung 1471

Literature

Cohen D, Rosenzweig R (2005) Digital history. A guide to gathering, preserving, and presenting the past on the Web. University of Pennsylvania Press, Philadelphia

Eikvil L (1993) OCR. Optical character recognition. https://www.nr.no/~eikvil/OCR.pdf. Zugegriffen am 07.10.2019

Hassner T, Rehbein M, Stokes PA, Wolf L (2014) Computation and palaeography: potentials and limits. Dagstuhl Manifesto 2(1):14–35

Kenney AR, McGovern NY (2003) The five organizational stages of digital preservation. In: Digital libraries: a vision for the 21st century: a festschrift in Honor of Wendy Lougee on the occasion of her departure from the University of Michigan

Klimpel P, Rack F, Weitzmann JH (2017) Handreichung — Neue rechtliche Rahmenbedingungen für Digitalisierungsprojekte von Gedächtnisinstitutionen. 4. Aufl. https://irights.info/wp-content/uploads/2017/12/Handreichung_Recht_Digitalisierung-Gedaechtnisinstitutionen-4-Aufl-2017.pdf. Zugegriffen am 07.10.2019

Kohle H (2018) Museen digital. Eine Gedächtnisinstitution sucht den Anschluss an die Zukunft, 1. Aufl. Heidelberg University Publishing, Heidelberg

Library of Congress (2006) Technical standards for digital conversion of text and graphic materials. http://memory.loc.gov/ammem/about/techStandards.pdf.2479.Digitalisierung. Zugegriffen am 07.10.2019

Neuroth H, Oßwald A, Scheffel R, Strathmann S, Huth K (2010) nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3. Vwh, Glückstadt

NISO (2004) Understanding metadata. http://www.niso.org/publications/press/UnderstandingMetadata.pdf. Zugegriffen am 07.10.2019

Parzinger H (2015) Kulturelles Erbe und Digitalisierung. In: Klimpel P, Euler E (Hrsg) Der Vergangenheit eine Zukunft. Kulturelles Erbe in der digitalen Welt: eine Publikation der Deutschen Digitalen Bibliothek, 1. Aufl. iRights Media, Berlin

Rehbein M (2017) Digitalisierung. In: Jannidis F, Kohle H, Rehbein M (Hrsg) Digital Humanities. Eine Einführung. J.B. Metzler, Stuttgart, S 179–198

Saffady W (2000) Micrographics: technology for the 21st century. ARMA International, Prairie Village

Schmitt H (Hrsg) (2010) Archive im digitalen Zeitalter. Überlieferung, Erschließung, Präsentation; 79. Deutscher Archivtag in Regensburg; [Tagungsdokumentationen zum Deutschen Archivtag, Bd 14]. Verband Deutscher Archivarinnen und Archivare; Deutscher Archivtag, 1. Aufl. Selbstverl. des VdA, Neustadt a.d. Aisch

Tanner S et al (2009) Measuring mass text digitization quality and usefulness. Lessons learned from assessing the OCR accuracy of the British Library's 19th century online newspaper archive. D-Lib Magazine 15:78

Wurster HW (2015) Die Matrikeln der Diözese Passau als Quellen zur Migrationsgeschichte. Publiziert: "*En route to a shared identity. Sources on the history of Central Europe in the Digital Age*" (Blog); Artikel der Internationalen Tagung "Auf dem Weg zu einer gemeinsamen Identität. Quellen zur Geschichte Mitteleuropas im digitalen Zeitalter", 16. – 17.09.2014, Universität Wien (AT). https://dighist.hypotheses.org/276. Zugegriffen am 07.10.2019

Referenced Websites

Deutsche Digitale Bibliothek. https://www.deutsche-digitale-bibliothek.de/. Zugegriffen am 07.10.2019

Deutsches Textarchiv. http://www.deutschestextarchiv.de/. Zugegriffen am 07.10.2019

DWB. Volltextdigitalisierung. http://dwb.uni-trier.de/de/die-digitale-version/volltextdigitalisierung/. Zugegriffen am 07.10.2019

Europeana Collections. https://www.europeana.eu/portal/de. Zugegriffen am 07.10.2019

Google Books. https://books.google.de/. Zugegriffen am 07.10.2019

Hathi Trust. https://www.hathitrust.org/. Zugegriffen am 07.10.2019

ICARUS für alle. https://icar-us.eu/icarus4all/. Zugegriffen am 07.10.2019

Matricula Online. http://data.matricula-online.eu/de/. Zugegriffen am 07.10.2019

MDZ (Münchener Digitalisierungszentrum). https://www.digitale-sammlungen.de/. Zugegriffen am 07.10.2019

nestor. https://www.langzeitarchivierung.de/Webs/nestor/DE/nestor/Ueber_uns/ueber_uns_node.html. Zugegriffen am 07.10.2019

Stiftung Preußischer Kulturbesitz: Bund fördert Pilotprojekt "Museum 4.0". Pressemitteilung vom 15.11.2016. http://www.preussischer-kulturbesitz.de/pressemitteilung/article/2016/11/15/pressemeldung-bund-foerdert-pilotprojekt-museum-40.html. Zugegriffen am 07.10.2019

Guidelines and Standards

Archive im Rheinland, Archivberatung/Bestandserhaltung: Umgang mit Zellhom-Filmen. Filme aus Cellulosenitrat (Zellhorn-Filme) gefährden Archive. https://afz.Ivr.de/de/archivberatung/bestandserhaltung_1/umgang_mit_zellhorn_film/umgang_mit_zellhorn_film_1.html#. Zugegriffen am 07.10.2019

CIDOC. CRM. http://www.cidoc-crm.org/. Zugegriffen am 07.10.2019

DFG-Praxisregeln "Digitalisieren" (2016). (DFG-Vordruck 12.151 — 12/16). Deutsche Forschungsgemeinschaft. http://www.dfg.de/formulare/12_151/12_151_de.pdf. Zugegriffen am 07.10.2019

Dublin Core. http://dublincore.org/specifications. Zugegriffen am 07.10.2019

EAD. http://www.loc.gov/ead/. Zugegriffen am 07.10.2019

FADGI — Federal Agencies Digitization Guidelines Initiative. FA Still Image Digitization Working Group, FA Audio-Visual Working Group. http://www.digitizationguidelines.gov/guidelines/. Zugegriffen am 07.10.2019

IFLA Guidelines. IFLA (International Federation of Library Associations and Institutions) and ICA. (International Council on Archives) for UNESCO. http://www.ifla.org/files/assets/preservation-and-conservation/publications/digitization-projects-guidelines.pdf. Zugegriffen am 07.10.2019

ISAD(O). https://www.ica.org/sites/defaulfiles/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_EN.pdf. Zugegriffen am 07.10.2019

LIDO. http://network.icom.museun/cidoc/working-groups/lido/resources/. Zugegriffen am 07.10.2019

MARC. https://www.loc.gov/marc/. Zugegriffen am 07.10.2019

Metamorfoze Preservation Imaging Guidelines. National Library of the Netherlands, National Archives (Hrsg H van Dormolen). https://www.metamorfoze.nl/sites/default/files/publicatie_documenten/Metamorfoze_Preservation_Imaging_Guidelines_1.0.pdf. Zugegriffen am 07.10.2019 1472 A. Schilz und M. Rehbein

METS. http://www.loc.gov/standards/mets/. Zugegriffen am 07.10.2019

MODS. https://www.loc.gov/standards/mods/. Zugegriffen am 07.10.2019

Pinkow A (2013) Barocke Bildwelten interaktiv erleben. Projektdossier. In: Humboldt Lab Dahlem (Hrsg) Ein Projekt der Kulturstiftung des Bundes und der Stiftung Preußischer Kulturbesitz (2012-2015). Leitung: Heller M, König V, Ruitenbeek K, Wegner A. Stand November 2014. — http://www.humboldt-lab.de/projektarchiv/probebuehne-1/gedankenscherz/projektbeschreibung/index.htm1 @tx_hfprjdoc_prjpdf[action]=download&tx_hfprjdoc_prjpdf[controller]=Project&cHash=4d03 177d7a22e4304ad120221d078239. Zugegriffen am 07.10.2019

TEL. http://www.tei-c.org. Zugegriffen am 07.10.2019

Osten M (2006) Digitalisierung und kulturelles Gedächtnis. In: Aus Politik und Zeitgeschichte. Bundeszentrale für politische Bildung, Bonn. APuZ 5-6:3-8

Europe 4.0

Walter Frenz

Contents

This article further develops the contribution in EWS 2019, 121 et seq. which is related to AG Wahl's opinion on the German car toll.

W. Frenz (✉)
Unit for Mining, Environmental and European Law, RWTH Aachen University, Aachen, Germany
e-mail: frenz@bur.rwth-aachen.de

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5_65

1 Central Importance of the EU Law on Digitization

It has already become clear in numerous passages in this book what central importance the EU law has for digitization. This starts with the question of who is protecting the data and who does it belong to. The NIS Directive (Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union, OJ 2016 L 194, p. 1) obliges companies to provide efficient protection against attacks by third parties, but it is not sufficient yet, especially with regard to small and medium-sized companies (Frenz 2019 in more detail, also for the following). Nonetheless, uniform state measures can be stipulated across the Union, for example when it comes to minimum penalties for **cybercrime**. It is also possible to specify uniform criminal offenses in this area (Article 83 Section 1 TFEU).

Even the **drafting of contracts** under civil law can be controlled under Union law. An example of this is the current **guideline on sharing private sector data in the European data economy**,[1] which prescribes that agreements for the use of data should be concluded on an equal footing. In terms of primary law, competition law is particularly important, as it allows even competitors to cooperate to a limited extent to advance digitization through improved or even new products and services. In addition, it grants access claims against monopolists, for example for software usage and the cloud (on the whole, Frenz 2016b, p. 671). It also limits the possibilities of a market leader for comprehensive processing and use of personal data based on the conditions of use ("**abuse of conditions**"), as the Federal Cartel Office of Germany decided on July 2, 2019 at the expense of Facebook. Possibly, in the area of competition rules, a guideline to Competition Law 4.0 can ensure greater clarity and sensitivity among those involved.

2 Remaining Obstacles and Obstacles to Be Removed

2.1 Competition Law

2.1.1 Entitlement to Access from the Prohibition of Abuse

An important obstacle that could stand in the way of Industry 4.0 and its further development has already been addressed. It is a matter of locking up essential basic equipment for digitization. Union law, which now dominates **competition law** and at least structurally prescribes it in its national form, ensures an opening towards

[1] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions "Towards a common European data space," COM (2018) 232 final.

monopolists. Microsoft, for example, had to open its platforms to other companies, although a reasonable fee may be demanded for this (CFI, ruling T-201/04, ECLI: EU:T:2007:289—Microsoft I; Ruling T-167/08, ECLI:EU:T:2012:323—Microsoft II; Frenz Industry 4.0 and Competition Law).

2.1.2 Open Standard Setting by Associations

This opening must not be undermined by the setting of standards by private associations. Corporate associations, in which regulations for technical innovations are often drawn up, are also subject to the ban on cartels and must therefore act in such a way that the interests of individual actors are not neglected compared to those of other corporate entities. Such a disadvantage can lead to innovative players not being able to place their products on the market.

Accordingly, **standards** must be designed so **openly** that new developments are also marketable. Digitization is essentially based on this, since there are many very innovative small companies that have not been in the foreground so far, but have ideas and developments that may displease the previous “top dogs.” The latter are therefore not allowed to dominate standardization committees in such a way that the standards are only tailored to their products and thus innovative ideas have no chance to enter the market (closer Frenz, standard setting by associations).

The openness of standards is also an important part in connection with fundamental liberties. State behavior is prohibited through them. The private association work, for example on standardization, is therefore only recorded to the extent that it is used and controlled by the state. If there is scope for decision making and the state, for example, only approves the work of a private standardization committee, competition law also applies to state behavior (ECJ, C-184/13, ECLI:EU:C:2014:2147—API).

2.2 Fundamental Liberties

2.2.1 Free Movement of Goods

Another important field to record exclusive state or at least state-dominated behavior is fundamental liberties. The **free movement of goods** is affected when it comes to the cross-border movement of products that have emerged in the course of digitization and now require authorization to be sold in other Member States. The free movement of goods establishes the principle that every product that has been lawfully marketed in one Member State may also be sold in other Union states (ECJ, ruling C-120/78, ECLI:EU:C:1979:42 paragraph 14—Rewe/Federal Monopoly Administration for Spirits; closer Frenz 2015, paragraph 761 et. seq.). All de facto handicaps must also be avoided. This applies above all to the establishment of national product standards. They must be justified by written or unwritten reasons such as consumer or environmental protection.

Untrammelled free movement of goods is also important. Therefore, the Advocate General Wahl in his opinion of February 6th, 2019 (C-591/17, ECLI:EU:C:2019:99, paragraph 118), like the ECJ (ruling C-591/17, ECLI:EU:C:2019:504, paragraph 129), has not classified the German car toll as a mere sales modality, as it does not constitute a measure to regulate the manner in which goods may be marketed but rather concerns the way in which goods are transported. While sales modalities generally do not hinder the access of imported goods to a market in a Member State, except in the case of discrimination, a restriction in the type and manner of transport can even have direct effects on the cross-border movement of goods, as AG Wahl (paragraph 118) rightly stated (Frenz, EWS 2019, also for the following).

New sales channels must also be **allowed in principle**. This was shown in the *Parkinson* ruling with regard to the mail-order business of pharmaceuticals, which must not be prohibited in general. In addition, restrictions according to the ECJ must be justified in more detail and their necessity must be proven. The ECJ examined the suitability, necessity, and the appropriateness very carefully, and made its own considerations and conclude that competition has a more positive effect on the survival of companies adhering to tried and tested methods than preventing it (for more details see ECJ, ruling C-148/15, ECLI:EU:C:2016:776—Parkinson; about this Frenz 2017a, b, pp. 9 et seq.). Such an approach is in line with further developments, including those through digitization.

However, it is essential to maintain **high product quality**. This is particularly important as the internal market is essentially based on trust. Hence, a secondary legal protection is provided on this basis. To remove unsafe and illegal products, the European Parliament and the EU Member States tentatively agreed on new regulations on market surveillance and compliance for consumer products on February 7th, 2019. In this regulation, which will apply from 2021, the national authorities can work together more efficiently and thus improve product controls as well as strengthen those at the external borders (Europaticker of February 9th, 2019).

2.2.2 Free Movement of Services

The same applies to services. To be exact, they play an even greater role in Industry 4.0 insofar as they also include assistance in setting up and further developing digitization in numerous companies. This applies to software offers, for example. Both the active and passive **freedom of services**, namely both the offering and the untrammelled acceptance of such services, are protected. In this respect, too, neither legal nor actual hurdles may be set up that are not justified by a common good. This also includes state tax collection.

2.2.3 Removal of the Toll?

The planned German car toll for vehicles from other EU countries is also seen as an obstacle to the merging of Europe. However, Advocate General Wahl (opinion of February 6th, 2019—C-591/17, ECLI:EU:C:2019:99) considered it to be compatible with both the general prohibition of discrimination in Article 18 TFEU and with the fundamental liberties (goods and services) as well as with the transport policy. The ECJ (ruling C-591/17, ECLI:EU:C:2019:504) rightly accepted a violation.

Advocate General Wahl referred to the negligible effects of a car toll levied on both private and commercial journeys (opinion of February 6th, 2019—C-591/17, ECLI:EU:C:2019:99, paragraph 124). However, the ECJ fundamentally rejected a de minimis limit in line with antitrust law. Noticeable effects on cross-border trade are not required for an infringement (ECJ, ruling C-16/83, ECLI:EU:C:1984:101, paragraph 20—Prantl). Only hypothetical developments without sufficient causality due to uncertain, indirect, or mediate significance are to be excluded (ECJ, ruling C-140/94 et al., ECLI:EU:C:1995:330, paragraph 29—DIP). This should apply above all to measures that do not differentiate "according to the origin of the goods transported" and should not "regulate trade in goods with other member states" (cf. ECJ, ruling C-266/96, ECLI:EU:1998:306, paragraph 31), and thus to the German car toll (AG Wahl, opinion of February 6th, 2019—C-591/17, ECLI:EU:C:2019:99, paragraph 124).

Just as the standards of the Parkinson ruling protect new sales channels very well, possible restrictions on cross-border traffic must also be examined very carefully and prevented as well. Admittedly, Advocate General Wahl has presented detailed explanations but in any case only listed negligible effects on the fundamental liberties. This also applies to the freedom of services, for which AG Wahl refers to the indiscriminate collection of the German car toll and its low level compared to other EU countries. He therefore sees no indications that could point to an obstacle to market access (opinion of February 6th, 2019—C-591/17, ECLI:EU:C:2019:99, paragraph 132). Again, however, the offsetting of the car toll for German car users with the vehicle tax, in which other vehicles do not participate, is hidden. The ECJ (ruling C-591/17, ECLI:EU:C:2019:504, paragraph 126) rightly demands a uniform approach and affirms a violation. The **psychological effect** of levying **motorway tolls** with a simultaneous reduction in vehicle tax only for residents was completely unmentioned. This can also have an inhibiting effect on bringing goods or delivering or receiving services to such a state. However, it is difficult to determine in advance how strong such an effect will be.

This shows the difficult handling of the **de minimis threshold** within the framework of the fundamental liberties. This applies not only specifically to the car toll but also in general. It should be completely dispensed to increase the effectivity of the fundamental liberties. The ECJ also uses an irrelevance threshold at best sparingly and with a view to uncertain effects (see more in Frenz 2012, paragraph 464 and 901 et seq.).

2.2.4 Public Procurement Law

Public procurement law is also an expression of the **fundamental liberties**. Although it has now been regulated in detail in three directives[2] and many points have been further developed, it must still be seen and interpreted against the background of fundamental liberties (in detail Frenz 2018). It is therefore an instrument of public procurement in accordance with them. **Electronic public procurement** is a major relief (for more details, see Wanderwitz 2019, p. 26). Above all, it helps providers from other EU countries to take part in transnational tenders. When it comes to registers for listing suitable applicants of certain contract categories or vice versa, for example, excluded bidders because they committed economic or environmental crimes, the efficiency of a Union-wide concept is greater.

In the tenders, **requirements** can also be made that correspond to the progress made in digitization. Product standards can be defined in such a way that they originate from digitized production or that they ensure such, for example, when it comes to the further processing of products. In the case of **software solutions for day-to-day administration**, it can be required that these enable administration activities that are as automated and smooth as possible. In the case of supervision of the digitized administration, basic requirements can be placed on the qualification and experience of the bidders.

Public procurement law is therefore an **important vehicle** for efficiently implementing digitization in public administration. At the same time, it is itself very much shaped by digitization. The legal requirements for this result largely from the EU procurement directives, which have been implemented in national law.

2.2.5 Free Movement of Workers

Digitization is based on highly qualified **specialists**. An exchange of these **between the Member States** is therefore also ensured with the permeability of the borders. In principle, this guarantees the free movement of persons in the form of the free movement of workers and the freedom of establishment.

The **free movement of workers** is, however, decisively hindered by the fact that there is still **no electronic income tax card** for Germans who live abroad but work in Germany. They have made use of the possibility of crossing the border within the European framework, too. Thus, they cross the border to go to their workplace, so

[2] Directive 2014/23/EU of the European Parliament and of the Council from 26 February 2014 on the award of concession contracts, OJ 2014 L 94, p. 1; Directive 2014/24/EU of the European Parliament and of the Council from 26 February 2014 on public procurement and repealing Directive 2004/18/EC, OJ 2014 L 94, p. 65; Directive 2014/25/EU of the European Parliament and of the Council from 26 February 2014 on procurement by entities operating in the water, energy, transport and postal services sectors and repealing Directive 2004/17/EC, OJ 2014 L 94, p. 243.

that there are not only purely internal issues that would not be subject to the fundamental liberties (ECJ, 175/78, ECLI:EU:C:1979:88 paragraph 12—Saunders). Tax liability and collection is and will remain a national matter. However, the circumstances must be designed in such a way that the free movement of workers is not actually hindered, unless this is sufficiently justified by the need for effective tax collection.

2.2.6 Freedom of Establishment

The **freedom of establishment** enables both the establishment of new companies and the opening of branches. The latter can be used to transport ideas and products that have been developed in one Member State in the course of Industry 4.0 to other countries. The establishment of companies in other EU countries enables EU citizens to work self-employed where the best framework conditions prevail.

In this way, the individual Member States are encouraged to provide a framework that favors business start-ups and relocation. It is about promoting start-ups and dealing well with the administration as well as reducing bureaucracy. On the other hand, the **prohibition of subsidies contrary to Union law** specifically excludes **tax breaks** for companies from other EU countries. All economic operators must be taxed equally. This applies to both legislation and its implementation. Otherwise, the Commission sees a **benefit that is contrary to the system**. This has already resulted in sensitive tax back payments for various companies, which the Member States had to impose based on the decision of the Commission. This is what the Fiat, Starbucks, and Microsoft cases stand for (more in Frenz 2016a, p. 142).

Furthermore, the freedom of establishment must **not** be hindered by the fact that **higher requirements** are set for the settlement of self-employed people from other EU countries than for locals. Qualifications from other EU countries are therefore to be recognized generously (see for more details Sect. 2.3 on education). This can, for example, concern start-ups or IT centers. Formally, equally applicable regulations are made in many cases. However, there is often a problem with implementation that is practically uniform in every respect. This also applies to the tax area.

2.2.7 Free Movement of Capital

The free movement of capital is not to be neglected. It enables companies to invest their capital in other EU countries without unjustified obstacles. Thus, in conjunction with the freedom of establishment, an invention can be sold and, if necessary, manufactured in all EU countries very quickly. The free movement of capital not only ensures the transfer of capital to the company's own subsidiaries, but above all the takeover of companies in other EU countries. Blocking minorities of Member States in areas which they consider to be elementary for security of supply represent a particular problem. However, such a meaning must be justified sufficiently clearly by the respective Member State to avert mere protective claims. This often did not succeed (for example, ECJ ruling C-463/00, ECLI:EU:C:2003:272, paragraph 70—Commission/Spain (Golden Stocks IV) for the energy sector).

2.2.8 Stricter Standards in the Interests of Digitization

The requirements of Union law have become stricter over time because the fundamental liberties no longer merely contain a prohibition of discrimination but also a **prohibition of restrictions**. Even in the case of formal equality of treatment, a result must be excluded which in fact discourages citizens from EU countries from crossing the border to work independently or as an employee in another EU Member State.

On the other hand, the ECJ (ruling C-148/15, ECLI:EU:C:2016:776—Parkinson) has **strengthened the proportionality control**. Restrictions on fundamental liberties must be justified precisely. It is not enough to state objective justifications such as environmental or health protection. Rather, the state must demonstrate and prove that the measure it has taken actually serves this goal and is suitable, necessary and appropriate for achieving it. The measure must therefore actually have a benefit for the intended goal and must not be disproportionate to the restriction of fundamental liberties. For this reason, the pharmaceutical mail-order business could not be restricted or even banned entirely in favor of the classic pharmacies for reasons of health protection (ECJ, ruling C-148/15, ECLI:EU:C:2016:776—Parkinson; see Frenz 2017a, b).

The European Court of Justice **no longer restricts** itself to **checking the evidence** of obvious errors and misjudgements but examines the facts very carefully and includes alternative scenarios as well as its actual assessments and equalizes them with the Member State. This allows it to see very easily to what extent mere self-serving claims are being made to protect the local economy from unpleasant competition from other EU countries.

This shows that the fundamental liberties must continue to be practically lived as well as that existing and new obstacles must be overcome so that Union citizens are encouraged rather than deterred to cross-border activities. Only in this way that new technical developments can spread rapidly across Europe, so that all EU citizens have the chance to benefit from them. This is the basic pattern on which the European fundamental liberties as well as the competition rules are based.

The idea of a single market with as few hurdles and restrictions as possible for economic activity can therefore also have a positive effect on Industry 4.0. Consequently, the traditional and classic matter of fundamental liberties can secure and promote great dynamism, especially for digitization. This applies in the economic as well as in the personal area.

2.3 General Prohibition of Discrimination

"Thou shalt not discriminate." This is the first sentence of AG Wahls conclusion from February 6th, 2019 (C-591/17, ECLI:EU:C:2019:99, paragraph 1). The fundamental liberties contain factual prohibitions of discrimination. In addition, there is the general principle of non-discrimination with its specific form of the prohibition of discrimination on grounds of nationality according to Article 18 TFEU and Article 1 Section 2 CFR. This is at the center of the final motions of AG Wahl such as the ECJ ruling on the German car toll. The car toll is not a direct discrimination, but an indirect one because domestic motorway users are compensated for the usage fee due by a reduction in the vehicle tax. Because of the different legal starting points, AG Wahl sees no discrimination when viewed in isolation, nor when viewed as a whole, because foreign vehicle owners are never more burdened than domestic ones (paragraphs 43, 54 et seq.). For the latter, however, there is a compensation and integration of the car toll in the vehicle tax and thus a de facto preferential treatment (ECJ, ruling C-591/17, ECLI:EU:C:2019:504, paragraph 44 et seq.).

Therefore, it must be considered that foreign vehicle owners also pay a vehicle tax, albeit in their own country, and this is not reduced. Therefore, it is a much more difficult burden for the car toll in Germany. This is especially true when the discrepancy in vehicle tax is very high. Cross-border commuters using motorways are particularly affected. This has sensitive repercussions on fundamental liberties—including the free movement of workers.

2.4 Transports and Networks

2.4.1 Transport Policy in Parallel with the Fundamental Liberties

Fundamental liberties must be adjusted towards transport. Therefore, as soon as the transport of goods and services can be hindered by state measures, sufficient justification is required. This also applies to the fact that not many **measures, which are minor** in themselves, **lead to considerable hindrances** to the **cross-border exchange of services**. For this reason alone it is to be rejected that AG Wahl rejects the significance of an impairment for the German car toll both for the free movement of goods as well as for the free movement of services (opinion of February 6th, 2019—C-591/17, ECLI:EU:C:2019:99, paragraphs 124, 132; more detailed above Sect. 2.2.3).

Because of the great importance of transport for the cross-border exchange of services, this broad concept cannot stop at transport policy. AG Wahl (opinion of February 6th, 2019—C-591/17, ECLI:EU:C:2019:99, paragraph 140 et seq.) interprets it therefore narrowly and essentially reduces it historically to the fact that the Member States are not allowed to choose unilateral, less favorable structures against the background of the numerous harmonization measures that have now been taken. In this respect, he denies any disadvantage for the foreign drivers.

According to the static comparison required by the European Court of Justice (ECJ, ruling C-195/90, ECLI:EU:C:1992:219, paragraph 20 et seq.—Deutsche LKW-Maut), the existing relationship between domestic and foreign road users must not be changed to the detriment of the latter. The relief in the case of the car toll may not be masked out only for residents via an exemption for the vehicle tax (Zabel 2015, pp. 186, 188 et seq.). The fact that the latter is only collected from residents does not change the fact that it has been re-regulated together with the car toll and thus there is an overall regulation that practically compensates the burdens for residents—but not for other Union citizens. This is also how the ECJ sees it (ruling C-591/17, ECLI:EU:C:2019:504 paragraph 162).

There is no exception, as Article 92 TFEU does not contain any and is also an expression of the general prohibition of discrimination (Frenz 2010, paragraph 3136). This also shows the necessary overall view with the other regulations to secure the internal market and the exchange of goods and people. This is especially true in the age of digitization, which can develop particularly well when economic areas grow together and there is more intensive exchange across borders. Only the free movement of services is superseded by transport policy under Article 58 TFEU.

In more depth, the harmonization measures now available show the **need for traffic to be as frictionless as possible**. It is therefore difficult to interpret transport policy in a restrictive manner and thus deepen rather than close the gaps that remain despite harmonization measures. Not only the transport of goods is essential for the realization of the internal market, but also the transport of people. Without them, the free movement of workers and the freedom of establishment, the freedom of persons, but also the free movement of services can hardly be achieved. They therefore play a significant role in the internal market (Article 26 Section 2 TFEU). Employees who work across borders, in particular, use motorways, buses and trains. For this reason, cross-border problems must also be eliminated in **local public transport**. One example is the LIMAX (Liège-Maastricht-Aachen-Express).

2.4.2 Pathways

However, pathways are also elementary. This applies in particular to the **transport of electricity**. The digitization can help in this regard by determining the need for electricity at other locations and providing appropriate quantities. In this way, the need for network expansion can be reduced. At the same time, the existing networks can be used more efficiently. The coal phase-out in Germany is to be accompanied by a modernization and better use of the electricity grid through optimization, expansion, and market-related measures. The final report of the Coal Commission of January 26th, 2019 (p. 80) states: “In addition to the necessary new grid sections and grid expansion, numerous smart solutions provide possible ways to use the existing grids more intelligently. Here, digitisation offers considerable potential.”

2.4.3 Infrastructure

This enables digitization to be implemented efficiently in the energy industry. It promotes the interconnection of **energy networks** within the framework of energy policy in accordance with Article 194 Section 1 sub-paragraph d TFEU and serves to establish and expand trans-European networks in the field of energy infrastructure in accordance with Article 170 Section 1 TFEU. This development and expansion can be decisively controlled by digitization. On the one hand, this allows the number of networks required to be reduced. On the other hand, digitalization can be used to calculate and determine better which networks are necessary and how they can be used most effectively. This approach is also suitable for the European networks in the areas of **transport and telecommunications infrastructure**.

With the help of digitization, an **area without internal borders** can be realized much more effectively in accordance with the objective of Article 170 Section 1 TFEU. It is therefore part of the *effet utile* of this regulation to use and promote it as comprehensively as possible. This can also have positive repercussions on the individual Member States, for example, through better cross-border usability to guarantee security of supply in the energy sector. In view of the simultaneous phase-out of nuclear power and coal this could be hugely useful.

2.5 Education and Culture

The basis of digitization as a demanding development is education. This applies to schools and universities. It is therefore important to strengthen **networking in the education sector**. This includes **student exchange programs**. Article 9 of the treaty between Germany and France on Franco-German cooperation and integration of January 22nd, 2019 provides for the expansion of exchange programs between the two countries, especially for young people—including measurable goals.

Knowing other countries is also crucial for digitization to receive suggestions for one's own environment and to combine existing progress with those achieved in other countries. Stays at a young age are particularly formative. However, to be able to actually use the achievements in studies and training in other EU countries, it is necessary to improve the recognition. The fundamental liberties stipulate the principle that the formal qualification is not important, but that the **knowledge and skills** actually acquired **must be adequately considered and recognized** (see already ECJ, ruling C-340/89, ECLI:EU:C:1991:193—Vlassopoulou).

The German-French treaty of January 22, 2019 also seeks to promote the **mutual recognition of school-leaving qualifications** in line with Article 10. This involves the creation of Franco-German excellence instruments for research, training and vocational training as well as integrated Franco-German dual study programs. Article 11 provides for the networking of education and research systems and the associated funding structures. This national cooperation is particularly important because, according to Article 165 TFEU, in the field of education the EU only promotes cooperation between the Member States that are fully responsible for the content and, if necessary, supports and supplements it.

The same applies to **culture** in line with Article 167 TFEU. Article 9 of the Franco-German Treaty of Aachen provides for special programs and a **digital platform** aimed particularly at young people. This links cultural exchange with digitization and makes it significantly more effective. Overall, according to this Article 9, a **common area of freedom and opportunities** as well as a common cultural and media area should be created. Both nations should grow together as closely as possible and become an area of freedom and culture with the greatest possible development opportunities. Digitization can play a decisive role here by creating quick and extensive connections and thus decisively promoting exchange.

2.6 Embedding Artificial Intelligence

Digitization cannot take place without artificial intelligence. It is an important part of it. Here it is necessary to create rules as to how **artificial intelligence** can be further developed and used. Employee protection rights apply here. **Workers' protection** has always been a key driver of European legislation. It is therefore important to design this new field of possible competition between employees and digitization or artificial intelligence in such a way that both sides are considered. Employees require special protection.

2.7 Industrial Policy

Conversely, economic development is largely dependent on Industry 4.0. **Industrial policy** mainly consists of advancing technology and thus strengthening the global competitiveness of EU products. It is also important to reduce unemployment. Investments are therefore to be encouraged. These are also largely shaped by Industry 4.0. Artificial intelligence and its further development are crucial components for further economic competitiveness. Therefore, these elements can be significantly promoted within the framework of industrial policy in line with Article 173 TFEU and for better exploitation of the industrial potential of policy in the fields of innovation, research, and technological development. This is particularly true for small and medium-sized enterprises, so that they too have a favorable environment. However, consumer protection must not fall by the wayside. This applies above all to necessary information on the dangers and risks of products. The legal options for product controls have recently been strengthened (see above Sect. 2.2.1).

2.8 Research Policy

Digitization is essentially based on its continuous development. It must therefore also be an integral part of the **European research area** provided for in Article 179 TFEU, in which there is freedom of movement for researchers as well as scientific knowledge. Hence, technologies are freely exchanged. The European **research programs**, which are often financially strong, offer the opportunity for digitization to be researched further and thus advanced within their framework. The funds are to be distributed accordingly. It is particularly valuable that cooperation with and between companies, research centers and universities is promoted (see Article 180 sub-paragraph a TFEU). Digitization thrives on new **research** findings that are applied and implemented in **practice**. Therefore, a **close connection** between the two is necessary.

2.9 Digitized Administration

Efficient administrative cooperation is an essential basis for an ever closer growing together of Europe. According to Article 197 paragraph 1 TFEU, this concerns the implementation of Union law. The Member States can be supported by the Union by facilitating the **exchange of information** as well as people and training programs. Administrative cooperation between individual Member States and between them as well as the Union is not ruled out and is essential for further networking among the EU states. This is served by bilateral agreements. In some cases, there is a long tradition here that leads to even closer cooperation.

A prime example is the cooperation between Germany and France, which was established in 1963 and which found a profound continuation in the Aachen Treaty on January 22nd, 2019. Its preamble speaks of an unprecedented network of bilateral relations between civil societies and government agencies at all levels.

The recording of all areas of life is elementary, especially for digitization. If the EU states work together within the framework of the European Union as well as with each other, a much greater expansion of Industry 4.0 can be achieved. This expansion then takes place in a coordinated manner and does not lead to new borders, as if each EU Member State goes its own way, which is completely detached from developments in other EU states. This would later make it more difficult to bring the activities together and move them forward across borders. In the past, the railway network developed so differently in different EU countries that the locomotives had to be changed at the border to continue their journey. Comparable phenomena should not even arise in the area of Industry 4.0. Therefore, procurement law is an essential vehicle in the field of public procurement (see Sect. 2.2.4 above).

Digitized administration leads to the collection of a great deal of data. Are they to be protected or used? The EU Commission is open to access to public sector data and its further use and sees it as an "important driving force for mass data analysis and artificial intelligence" (COM (2018) 234 final, p. 3). The Commission therefore presented a **proposal for a new version of the directive on the re-use of information from the public sector** (COM (2018) 234 final, p. 1 et seq.). Thus, real-time access to dynamic data using appropriate technical means or high-quality data is to be provided (COM (2018) 234 final, p. 1 et seq., closer compared to German law Guckelberger, Verwaltung im Zeitalter 4.0).

Based on digitized administration, the exchange between the borders is generally much easier. On this basis, a **uniform European data management** can be approached. This could extend to, for example, driving licenses by introducing a European driving license which completely replaces the national documents. However, according to the current structure, a European driving license must also be issued if the requirements in the respective Member State are met. Another option is to allow only European standards to be decisive or to restrict national requirements considerably.

2.10 Consequences of Autonomous Driving

Union-wide standards may be necessary in view of the fact that **autonomous driving** will become more relevant and therefore **specific requirements for the driver arise**. In this way, it could be determined to what extent the driver must remain vigilant even with autonomous driving. Such standards are necessary because autonomous driving takes place across borders. Thus, drivers need to know what level of attention they can and should expect from other traffic participants. Based on this, you have to design your own driving. This is independent of which EU country the participants come from. Sometimes this cannot even be seen, for example, when rental cars are on the way. It is then hardly possible for every driver to familiarize himself with the level of attention required in the respective EU country in which he is driving. The standards must therefore be as uniform as possible across the Union.

2.11 Law Enforcement and Criminal Law

The digitization of government agencies also affects the police and the public prosecutor's office. It enables the construction of much better **data processing systems**, which can be used to track down possible criminals much more quickly. This is done by comparing the data with one another and thus identifying certain characteristics or finding indications of previous criminal offenses, for example, tax violations in other Member States. The prosecution can thus be clearly effectuated. This is also necessary in view of the increasing hacker attacks (see above the articles by Müggenborg and Hartmann).

If EU-wide data comparisons and searches are made possible in this way, it does not matter whether the prosecution is carried out nationally or whether it is increasingly entrusted to **Europol**. The decisive factor is the **cross-border line of sight and criminal prosecution**. This can be accelerated by installing cross-border information systems for the police and police officers no longer having to stop at the border.

Since **cybercrime** takes place **across borders**, criminal law must react accordingly and, ideally, have uniform threats of punishment, both in terms of the offense and the legal consequences. In this respect, requirements under Union law can be made. Article 83 paragraph 1 TFEU provides for particularly serious crime with a cross-border dimension as possible areas and explicitly computer and organized crime. If the area of digitization is shaped by harmonization measures of the Union, minimum provisions for criminal offenses and penalties in this area can be laid down by directives (Article 83 paragraph 2 TFEU).

3 The EU Perspectives

Digitization is gradually permeating all areas of life. It can therefore also play an increasingly important role in the numerous fields of Union competencies. This also applies to **climate protection**. Digitization enables improved models and forecasts for the further development of the climate and thus also for the achievement of the 1.5-degree Celsius target, which implies that the earth is not warmed up by more than 1.5 degrees Celsius compared to the pre-industrial level. This was again put into focus at the UN climate conference in Katowice in November 2018. In the context of Industry 4.0, products and production processes can be developed that require as little energy as possible and are therefore largely CO_2-free.

At the same time, waste can be avoided by making the best possible use of the available resources and processing them in such a way that as little waste as possible is produced or, at most, there is waste for recycling that can be easily recycled—for example, through a modular system for automobiles.

These are just a few examples of the numerous fields of application of Industry 4.0 and thus ultimately of Europe 4.0. Education, research and industrial policy are important starting points. The fundamental liberties ensure standards that do not hinder cross-border exchange. Standardization can be achieved based on the **internal market harmonization**.

The European Union cannot only form an important basis for the further development of digitization; it is also decisively shaped by it. On the one hand, as already shown, this applies to the transformation of the administration. Here, too, digitized procedures are being used to an even greater extent. The European Court of Justice (ECJ) and the Court of Justice of the European Union (formerly known as Court of First Instance—CFI) are also significantly influenced by the developments in digitization. The modernization of the judiciary towards **Justice 4.0** will be adapted to its specific circumstances and will also find its way into the Court of Justice of the European Union. The protection of personal data must be carefully observed—this was substantiated by the ECJ (ECJ, ruling C-362/14, ECLI:EU:C:2015:650—Schrems).

Based on digitization, the European Commission has much better monitoring options, for example when it comes to Member States' violations of Union law. At the same time, it can make a significant contribution to further digitizing the administration, be it within its own framework, be it for the Member States in the implementation of Union law (see above Sect. 2.9).

This efficiency of the Union's institutions is particularly important in a **globalized, digitized world**. It is above all in demand for **large units** that have international weight to shape and advance developments, especially in the field of digitization. Close cooperation between the Union and its Member States is therefore essential. Equally elementary is a joint appearance in relation to China, the USA, and Russia. The Brexit is counterproductive. Worldwide, the Union is only perceived as a unit and must therefore be developed further, but not downsized. The influence of the EU is very important in many areas—especially in climate protection and digitization. These issues can be shaped by Union law, for example when it comes to environmental and consumer protection issues. Conversely, the EU also has an incentive to grow closer together. The sometimes skeptical Union population can thus clearly perceive that a strong European Union brings advantages and no disadvantages. Digitization can therefore inspire European unification—Europe 4.0.

References

Frenz W (2010) Handbuch Europarecht tome 6: Institutionen und Politiken. Springer, Heidelberg

Frenz W (2012) Handbuch Europarecht tome 1: Europäische Grundfreiheiten, 2nd edn. Springer, Heidelberg

Frenz W (2015) Handbuch Europarecht tome 2: Europäisches Kartellrecht, 2nd edn. Springer, Heidelberg

Frenz W (2016a) Steuerwettbewerb und Beihilfeverbot - die Fälle Fiat, Starbucks und McDonald's, DStZ 2016, pp 142 et seq

Frenz W (2016b) Industrie 4.0 und Wettbewerbsrecht, WRP 2016, pp 671 et seq

Frenz W (2017a) Das Parkinson Urteil: Versandhandelsverbot als Folge? – Anmerkung zu EuGH, Urt. v. 19.10.2016, C-148/15. GewArch 2017: 9 et seq

Frenz (2017b) Das Parkinson Urteil: Versandhandelsverbot als Folge?, GewArch. 2017, 9 et seq

Frenz W (2018) Vergaberecht EU und national, 2018. Springer, Heidelberg

Frenz W (2019) Europa 4.0 und die deutsche PKW-Maut – zwei Welten, EWS 2019: 121 et seq

Wanderwitz M (2019) Digitale Vergabe, Blockchain, Smart Contracts und die Digitalisierung des Vergabeverfahrens, VergabeR 2019, pp 26 et seq

Zabel M (2015) Die geplante Infrastrukturabgabe ("Pkw-Maut") im Lichte von Art. 92 AEUV, NVwZ 2015: 186 et seq

Index

W. Frenz (ed.), *Handbook Industry 4.0*,
https://doi.org/10.1007/978-3-662-64448-5

D