bitkom

Comments on the White Paper on AI

Bitkom comments on the White Paper on Artificial Intelligence – A European approach to excellence and trust

12. Juni 2020

I. General Remarks

The White Paper on Artificial Intelligence intends to develop and establish a European approach to excellence and trust in the field of Artifical Intelligence. Bitkom comments on the Commission's considerations in line with the above mentioned objectives of the White paper.

At the outset, we would like to stress that we welcome the objectives of the European Commission in general: Support the EU in becoming leader in AI (chapter 4: ecosystem of excellence) complemented by introducing new safeguards for citizens (chapter 5: ecosystem of trust).

In principle, we do not see the need for a specific AI-regulation throughout Europe. Before such regulation is introduced, it should be examined in detail from a legal point of view where there are blank spots on the EU regulation map and where significant restrictions of the digital single market are imposed by regulations in member states. This applies in particular to potential regulation that is explicitly introduced as a consequence of the increased use and dissemination of artificial intelligence in the economy and society. In our view, it has not yet been proven that the considerations made in the paper give rise to a general need for additional regulation of AI.

In this context, the scope of the white paper's suggestions should be further clarified in terms of whether the proposals refer to AI, machine learning or automated decision making and its impact on humans, which we will elaborate on in later comments.

There must be evident need for regulation, particularly in the context of the goal to establish an attractive ecosystem for innovation and excellence. Regulation is inevitably associated with effort and costs because of the relatively fixed (and thus degressive) cost structure, which is particularly challenging for start-ups and small companies. The need for regulation must therefore be comprehensibly documented and justified, which must be taken into account in all future considerations of regulation, particularly given the EU's goals to create the framework conditions for an ecosystem of excellence.

Bitkom
Bundesverband
Informationswirtschaft,
Telekommunikation
und Neue Medien e.V.
(Federal Association
for Information Technology,
Telecommunications and
New Media)

Lukas Klingholz
Big Data & Artifical Intelligence
T +49 30 27576 101
l.klingholz@bitkom.org

Benjamin Ledwon
Head of Brussels Office
Office +32 2 60953-21
b.ledwon@bitkom.org

Albrechtstraße 10
10117 Berlin
Germany

President
Achim Berg

CEO
Dr. Bernhard Rohleder

www.bitkom.org

Apart from fundamental principles, the application of AI should not be regulated in a dedicated AI directive, but on a sector- or topic-specific basis if necessity emerges. Regulations across various industries already consider and regulate data-driven applications. If new regulation needs to arise in the application of artificial intelligence, it makes most sense to link it with existing rules, standards[1] and regulatory framework within sectors.

With regard to the „ Ecosystem of Excellence", the main objectives are on the one hand to develop a European ecosystem in which research results can be applied as best as possible in the public and private sector as well as in the society as a whole. On the other hand, it is important that public and private investments complement each other optimally.

With regard to the „Ecosytem of Trust", it is important that the private sector is closely involved in the definition of the various requirements. Standards and certificates developed by business. Private sectoral self-government is just as relevant for a proper regulatory framework of the application of AI as the expertise of the regulatory authorities from the different sectors.

In further commentary we focus on the chapters four and five of the White Paper because these chapters deal with the conditions necessary to develop „Ecosystems of Excellence and Trust" in the European Union.

[1] For example ISO/IEC JTC1 SC42 already has a work board programme in place for this purpose.

II. An Ecosystem of Excellence

1. General Remarks

Bitkom welcomes the goal of establishing an ecosystem of excellence along the entire value chain (research & innovation and incentives for AI adoption in the economy).

2. Specific Remarks

A.Working with member States

Action 1: The Commission, taking into account the results of the public consultation on the White Paper, will propose to the member states a revision of the Coordinated Plan to be adopted by end 2020.

Bitkom Assessment

It must be more clearly worked out how, in the view of the Commission, various actors (both public and private) contribute to actually investing 20 billion annually in AI. In this context, a systematic inventory of current investments on the one hand and a realistic growth path of investments on the other hand should be aimed at.

Sustainability and energy effiency is not a new challenge for businesses. Just as with other technologies, companies have a strong self-interest in exploiting the potential of technologies for more sustainability. AI offers much potential for achieving sustainability goals. Bitkom therefore welcomes the Commission's statements and considerations to fully exploit the potential of AI in this context.

B. Focusing on the efforts of the research and innovation community

Action 2: the Commission will facilitate the creation of excellence and testing centres that can combine European, national and private investments, possibly including a new legal instrument. The Commission has proposed an ambitious and dedicated amount to support worldclass testing centres in Europe under the Digital Europe Programme and complemented where appropriate by research and innovation actions of Horizon Europe as part of the Multiannual Financial Framework for 2021 to 2027.

Bitkom Assessment

We fully support the goal to establish a lighthouse centre of research in Europe. We would suggest to establish a structure in which the lighthouse centre has a coordinating role in the european research and innovation community. The lighthouse center must be clearly linked to existing structures of excellence, such as CLAIRE and ELLIS. It also should strive to connect with standardization to foster market development.

Furthermore, we welcome the approach of concentrating on the sectors where Europe has the potential to become a global champion. In addition to excellence in research, a consistent approach on the transfer of knowledge & AI adoption is central regarding this.

It needs to be further clarified which financial instruments and incentives will be used to achieve these objectives.

C. Skills

Action 3: Establish and support through the advanced skills pillar of the Digital Europe Programme networks of leading universities and higher education institutes to attract the best professors and scientists and offer world-leading masters programmes in AI.

Bitkom Assessment

We support the objectives of the section on the creation of academic excellence and upskilling in the workforce. The measures to achieve these objectives need to be further specified.

D. Focus on SMEs

Action 4: the Commission will work with memberstates to ensure that at least one digital innovation hub per member state has a high degree of specialisation on AI. Digital Innovation Hubs can be supported under the Digital Europe Programme. The Commission and the European Investment Fund will launch a pilot scheme of €100 million in Q1 2020 to provide equity financing for innovative developments in AI. Subject to final agreement on

the MFF, the Commission's intention is to scale it up significantly from 2021 through InvestEU.

Bitkom Assessment

We support this goal. Existing hub structures and hub-like structures must be taken into account and further developed in a proper way in order to achieve the goals.

E. Partnership with the Private Sector

Action 5: In the context of Horizon Europe, the Commission will set up a new public private partnership in AI, data and robotics to combine efforts, ensure coordination of research and innovation in AI, collaborate with other public-private partnerships in Horizon Europe and work together with the testing facilities and the Digital Innovation Hubs mentioned above.

Bitkom Assessment

We fully support this approach.

F. Promoting the adoption of AI by the Public Sector

Action 6: The Commission will initiate open and transparent sector dialogues giving priority to healthcare, rural administration and public service operators in order to present an action plan to facilitate development, experimentation and adoption. The sector dialogues will be used to prepare a specific 'Adopt AI programme' that will support public procurement of AI systems, and help to transform public procurement processes.

Bitkom Assessment

In the scope of funding, we miss the reference to leveraging by public procurement, taking into account that half of EU's GDP is public money. Also, the paper misses a reference to an EU-funded national and transnational lighthouse-project that would facilitate private development and implementation of AI.
In general, we fully support the goal to promote and accelerate the deployment of products and services based on AI by the public sector. In addition to this abstract objective, an analysis is needed of how broad use of AI in the public sector can be promoted in concrete regulatory and organisational terms.

G. Securing Access to data and computing infrastructures

Bitkom Assessment

We fully support the goal to invest in strategic data and computing infrastructures which forms the basis for the digital transformation.

H. International Aspects

Bitkom Assessment

We support these considerations

III. An Ecosystem of Trust: Regulatory Framework for AI

1. General Remarks

Overall we don't see major gaps in EU legislation. The law applies without regard to a certain technology. Consequently it seems not only inadequate, but rather detrimental to create specific law only for AI.

Besides, there is no agreed upon mechanism for classifying AI applications as such.

Chapter 5 is strongly oriented towards the concept of AI. The key questions for different economics operators arising from chapter 5 are the following two.

- Is the relevant data-driven application an „AI application"?
- If the relevant data-driven application is an „AI application": Is this specific AI application a high-risk application?

If a future regulatory framework - as follow up of the White Paper on AI - plans to regulate AI, the concepts of AI and algorithmic systems must be defined in a way which makes them easy to handle for the economic operators involved (developer, deployer, producer etc.) to determine if a specific data-driven application meets the criteria AI/algorithmic system.

It is very important that the regulatory framework under discussion for high-risk AI applications does not create serious burdens that prevent companies and society from developing and using high-risk AI applications in and in the EU.

2. Specific Remarks

Opportunities and risks of AI

Bitkom Assessment

It is to be welcomed to systematically consider the opportunities and risks of new technologies. Therefore, the scenario of not using a new technology should be compared to the scenario of using the specific technology. The use of artificial intelligence offers many advantages in many industries and areas. These advantages and potentials must be weighed against the risks of their use. We believe that the overall potential of artificial intelligence is very high.

Assumption that the lack of trust is main factor holding back AI uptake

Bitkom Assessment

We do not fully agree with this thesis as the empirical basis for this claim is missing. There are multiple other possible reasons, which are responsible for the low uptake of AI such as lack of legal certainty due to GDPR or missing standards etc. Standards can help to increase trustworthiness.

Role of HLEG seven key requirements

Bitkom Assessment

While we welcome the seven requirements in principle, it remains to be noted that further steps need to be taken for their practical application on a broad scale, precisely because a legal implementation of ethical criteria is not directly possible 1:1.

A. Problem Definiton

We strongly recommend the highest possible degree of technology neutrality in a regulatory framework for implementation.

Risks for fundamental rights, including personal data and privacy protection and non-discrimination

Bitkom Assessment

To what extent is the existing legal framework not sufficient to limit these risks in the operational use of AI applications and to ensure compliance with fundamental rights? From our point of view this is not clear enough in this section.

Risks for safety and the effective functioning of the liability regime

Bitkom Assessment

We reject an additional liability regime for the application of AI-based technologies. Liability regimes should be set up in a technology-neutral way.

B. Possible Adjustments to existing EU legislative framework relating to AI

Effective application and enforcement of existing EU and national legislation:

Bitkom Assessment

We agree that transparency is key and thus, provision of clear information needs to be guaranteed. It is however completely unclear where the GDPR's requirements in that regard are considered insufficient – and the White Paper does not provide any indication here. Also, often it remains unclear whether the paper refers to personal or non-personal data. With regard to personal data, GDPR appears as sufficient means to close potential gaps.

Limitations of scope of existing EU legislation:

Bitkom Assessment

The highest possible degree of technological neutrality should be maintained in the regulatory framework. We therefore reject the idea of creating product safety legislation especially for AI based technology.

Changing functionality of AI systems:

Bitkom Assessment

We are of the opinion that the changing functionalities of AI in high-risk applications are already largely covered by regulations. There is no evidence of where specific regulatory gaps exist. These may have to be adapted sector by sector in the respective regulatory frameworks.

Existing regulatory frameworks (e.g. in the health sector or in the transport/automotive sector) already cover the topic of changing functionality of AI systems. In this context, please also note our explanations further down in this section (d) on robustness and accuracy "Requirements ensuring that outcomes are reproducible"). We strongly recommend taking a sectoral look at the advantages and disadvantages of locked algorithms from a user's point of view.

We point out that the wording "software" or "software update" is not clear or misleading in this context. A model can evolve (by adjusting parameters in the course of continuous learning) without the software itself changing. Is this a software update in the literal sense of the word?

Uncertainty as regards the allocation of responsibilities between different economic operators in the supply chain:

Bitkom Assessment

We do not see this uncertainty. Where such ambiguities exist regarding responsibilities they should be addressed in technology neutral regulatory frameworks such as EU product liability legislation and in vertical regulatory frameworks rather than in an AI-specific and non-technology-neutral new framework & special law.

Changes to the concept of safety

Bitkom Assessment

See considerations above. We reject AI or mandatory technology specific regulation. Potential changes should be addressed in technology neutral regulatory frameworks such as safety legislation and in vertical regulatory frameworks rather than in an AI-specific and non-technology-neutral new framework & special law.

However, we welcome the proposal to dock these considerations and plans closely to the considerations and plans of ENISA. To reach a high security-level for AI, technical standards particularly concerning robustness should be developed and approved, possibly based on the certification mechanism foreseen in the Cyber Security Act.

Based on the Security-by-Design Principle, all relevant stakeholders along the value chain should be addressed and the trader or deployer of the AI-system must not be left alone.

Report on the safety and liability implications of AI, the IoT and robotics

Bitkom Assessment

No further comments. Comments regarding „Section B" sufficiently represent our position.

C. Scope of a future EU regulatory framework

Definitions: Data, Algorithms, AI

Bitkom Assessment

There is no agreed mechanism for classifying AI applications as such. If a future regulatory framework plans to regulate AI, the concepts of AI and algorithmic systems must be defined in a way which makes them easy to handle for the economic operators involved (developer, deployer, producer etc., compare p. 22) to determine, if a specific data-driven application meets the criteria AI/algorithmic system.

Two concrete possible definitons are mentioned in the White Paper (Communication on AI for Europe[2] and the defintion from the HLEG[3]). We do not think that one of these defintions is approriate. We would like to emphasize that we find it very difficult to define AI precisely, also relative to data-based innovations and algorithms. When a definition is specified, it is in our view essential to make it as technology-neutral and as possible.

Looking at the definitions in the White Paper, in particular the definition of HLEG, we would like to make the following points about this definition, in the sense of an adapted and completed definition.

Artficial intellience (AI) systems are software (and possibly also hardware[4]) systems designed by humans that, given a complex goal, are taught by their designers or learn[5] from experi-

[2] Compare page 16, footnote 46 in the whitepaper
[3] Compare page 16, footnote 47 in the Whitepaper

[4] Regarding hardware we want to emphasize that this is only true for hardware with embedded software. Not for hardware itself.
[5] We would like to encourage you to speak of "optimize" rather than "learn"

ence how to act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal.

High risk, cumulative approach

Bitkom Assessment

We agree that a combination of defining relevant sectors and relevant use-cases within the sector could be a reasonable way to identify high-risk AI-systems, for the purpose of keeping strict and burdensome requirements proportional. While a sector-approach alone could lead to overly extensive regulation, solely looking at specific use-cases across all sectors could lead to a very high level of complexity and bureaucracy.

Potential high-risk sectors are mentioned in the paper (healthcare, transport, energy and parts of public sector)

How is „significant risk" concretely defined? This must be specified as clearly and practically as possible in order to enable an operationally manageable application and classification.

In this context, is the reference scenario considered enough (not using AI) when declaring a sector high-risk?

It is important to prevent entire sectors from being placed under general suspicion, as each sector involves applications, products and services with different risk requirements. In every sector, risks must be assessed on a case basis. We would like to emphasize once again that we believe that the majority of the applications in the various proposed high-risk sectors are uncritical in the sense of the White Paper (no significant risks in application, see p.17 bullet 2). Even though there are already sector-specific rules, the following is worth considering: a list of "high-risk sectors" is very difficult to manage and could significantly impair the use of AI for non-critical functions/services in all the sectors listed.

We therefore strongly recommend that clear criteria be defined according to which a sector is "high-risk". The explanations and justifications for this in the White Paper are very vague and not sharp enough. If necessary, the sectoral approach must subsequently be discarded if no clear criteria can be established which classifies a high-risk sector or if this classification leads to hampering general AI use and developments of AI applications (especially low risk applications) in this sector to be included in this list.

In this context problems regarding the differentiation of sectors may arise. Indeed, as digitisation increases, sectors overlap and are no longer distinct.

The complementary criteria to identify high-risk AI-systems, needs further specification to avoid legal uncertainty and over-regulation. This includes e.g. the clear definition of "significant" impact on affected parties or the "exceptional instances" that classify an AI-system as high-risk. Furthermore, as mentioned above, it must be clear and easy to classify if a specific data-driven application is an AI application6.

Also, it remains generally unclear whether the more general definition of high-risk for AI-systems is consistent with the separate proposal of high-risk in the scope of the proposed liability provisions for AI-systems.

Recruitment processes & applications impacting workers rights should always be considered high-risk

Bitkom Assessment

To what extent is the comparison with the reference scenario (no AI deployment) taken into account? To what extent is the existing legal framework not sufficient?

Biometric identification & biometric authentification

Bitkom Assessment

Already now, several security and data protection requirements apply to all applications in the field of remote biometric identification. It must be clearly worked out, which additional requirements are necessary by classifying remote biometric identification as high-risk, given the regulatory status quo.

D. Types of requirements

Bitkom Assessment

[6] Following the spirit of the whitepaper in general and the considerations on page 18 concretely.

In general, a lot of the requirements discussed in this chapter are very costly. This tends to mean that many AI applications can no longer be developed profitably. Therefore, the most unbureaucratic and unobtrusive implementation of the requirements is central to the design of a possible regulatory framework for high-risk AI applications.

a) Training data

- **safety rules**

Bitkom Assessment

In general, our comments regarding „Section B" sufficiently represent our position. Furthermore, compliance with safety standards must be closely linked to the results and findings of standardization activities.

- **anti-discrimination**

Bitkom Assessment

Authorities need to define clear and easy-to-use requirements and test criteria to identify potentially unlawful discrimination. In this context, sector-specific standardisation activities must be taken into account and encouraged.

We see several and major conflicts with GDPR here. Due to GDPR and data protection regulation, personal data can not be collected in a lot of cases. However, in many cases these would be necessary to meet the anti-discrimination requirements outlined in this section.

The important messages should be: It is not apparent why a new, additional anti-discrimination regulation specifically for data and AI algorithms is necessary. Discrimination is already covered by law. The focus should be on a non biased outcome of the AI system, as potential discrimination only occurs when the trained algorithm is applied, even if the data on which the algorithm is trained play a significant role.

- **privacy**

Bitkom Assessment

It must be made clearer where additional requirements are needed which are not covered by General Data Protection Regulation and the Law Enforcement Directive.

b) Keeping of records and data

- **accurate records regarding the data set used to train and test the AI systems, including a description of the main characteristics and how the data set was selected**

Bitkom Assessment

While a clarification regarding the documentation and retention obligation for development documentation is welcomed, we see the following problems here:

For already applied AI technology, it is very complex and costly. If a third party is auditing the records and data, there a major conflicts with trade secrets and security arising. This leads to the general question of how IT-security issues managed in the new regulatory framework by authorities?
Numerous data sets used in training AI systems could not be recreated and AI systems may be ingesting continuous flows of historic or real-time data over time. It would often be ineffective for companies to be required to keep such records or datasets when AI is frequently developed in a dynamic and iterative manner.
Also the trend of edge computing is not considered here (which is relevant and becoming more and more relevant in several sectors & industries). One characteristic of edge computing is that data is processed at the edge and not a (central) cloud.

Also, the consequences of the rise of federated computing paradigms for the availability of data records and datasets must be taken into account in this context. Especially, when the overall policy framework tries to promote this trend in other areas (see for example the data strategy)

Furthermore, many of the software-development processes and standards that have evolved over time and are used to help build trust in software do not exist for data; there are no common data naming conventions, no formatting standards or concurrent versioning systems used for data which make regulation in this area premature and impractical. Therefore, to require AI developers to keep records and data would be unlikely to lead to anything meaningful that could be garnered for assessment.

We also see various potential problems and legal conflicts with copyright law. Furthermore We see several conflicts regarding this requirement with GDPR in general and also in particular with the GDPR-based right to forget.

All the these additional requirements lead to processing costs and have a strong impact on the marginal/break-even decision of applying AI-based applications. Furtheremore, these requirements lead to additional energy/environmental costs.

- **in certain justified cases, the data sets themselves;**

Bitkom Assessment

The term justified case must be defined more concretely. Furthermore, there are special challenges in different AI technologies such as federated machine learning where data sets themselves are never collected. The role of anonymised and pseudonymised data in these cases needs to be worked out. Finally, we see several conflicts with GDPR.

- **Documentation on the programming and training methodologies, processes and techniques used to build, test and validate the AI systems, including where relevant in respect of safety and avoiding bias that could lead to prohibited discrimination.**

Bitkom Assessment

High administrative costs must be taken into account. Clear, operationally manageable rules and standards are necessary for implementation. Industry must be closely involved.

c) Informations provision

- **AI system's capabilities and limitations. This information is important especially for deployers of the systems, but it may also be relevant to competent authorities and affected parties.**

Bitkom Assessment

From our point of view, these questions are already sufficiently addressed in a large amount of cases, especially in B2B relations.

- **Discussion on information and labelling requirements**

Bitkom Assessment

Which specific additional informations, in addition to informations induced by EU data protection legislation should be provided?

We would like to stress that the objective of "avoiding unnecessary burden" is to be welcomed.

d) Robustness and accuracy

- **Requirements ensuring that the AI systems are robust and accurate, or at least correctly reflect their level of accuracy, during all life cycle phases;**

Bitkom Assessment
Precise definition of accuracy is necessary, especially when compared to processes and situatiuons without AI applications usage.

- **Requirements ensuring that outcomes are reproducible;**

Bitkom Assessment
From our point of view, this is not always appropriate and in the interest of the user. In several applications, new versions of AI systems come at short intervals. In this case, all intermediate versions must be kept available following this requirement. Therefore, we would argue to look in detail where this requirement is neccessary given the trade-off of insights into reproducibility on the one hand and the additional administrative and processing costs on the other hand.

This requirement is also problematic as it is not always possible to achieve this. AI systems change over time and outcomes are not reliably reproducible, therefore compliance with requirements of this nature would be impossible for many AI applications. Reproducibility of outcomes may require exactly reproducing the entire dynamic environment and the entirety of the data flows used to train the model and this would simply not be possible in practice in several cases.

Requirements ensuring that AI systems can adequately deal with errors or inconsistencies during all life cycle phases.

Bitkom Assessment
How is this exactly defined and how is monitoring to be ensured in practice?

e) Human oversight

Bitkom Assessment

Generally, we support to have human oversight. However, the degree of this possible oversight might vary from one case to another and should be limited to high-risk applications. Again, legal certainty for businesses is key and therefore, clear criteria need to be established that allow companies to determine what rules have to apply in specific situations.

In principle, the different gradations below are to be evaluated positively, since it shows that there is no "one size fits all" solution for human oversight.

The interaction between the four different non-exhaustive manifestations and their respective fields of application[7] should be considered and clearly defined.

- **Output of the AI system does not become effective unless it has been previously reviewed and validated by a human (e.g. the rejection of an application for social security benefits may be taken by a human only);**

Bitkom Assessment

Clear criteria, for example based on standards, must be defined when the human is applying the „review & validation" process. Otherwise, partial automation through the use of AI is taken ad absurdum.

- **Output of the AI system becomes immediately effective, but human intervention is ensured afterwards (e.g. the rejection of an application for a credit card may be processed by an AI system, but human review must be possible afterwards);**

Bitkom Assessment

This objective has to be compared with the regulatory status quo in the relevant vertical regulatory frameworks. For the most part, these rights of human intervention are already legally secured and no further regulation is needed. This leads to the question for which of the six high-risk areas this requirement should apply.
There must be clear and easily manageable rules and limits regarding the explainability of AI systems. Standards and certificates developed from business must play a central role here. Criteria for reversibility and unwinding as legal consequences must also be clearly defined.

[7] for which high-risk sector is the application of the respective manifestation appropriate, but for which area is it not?

- **monitoring of the AI system while in operation and the ability to intervene in real time and deactivate (e.g. a stop button or procedure is available in a driverless car when a human determines that car operation is not safe);**

Bitkom Assessment

Standards and certificates developed from businesses must play a central role here regarding the question of safety. In addition, it must be made clear how human oversight in the sense of the economic operators in Section E „Adresses" is to be implemented here with regard to responsibility/the distribution of obligation.

- **in the design phase, by imposing operational constraints on the AI system (e.g. a driverless car shall stop operating in certain conditions of low visibility when sensors may become less reliable or shall maintain a certain distance in any given condition from the preceding vehicle).**

Standards and certificates developed from businesses must play a central role regarding the question of safety. In addition, it must be made clear how human oversight in the sense of the economic operators in section E „Adresses" is to be implemented here with regard to responsibility/the distribution of obligation.

f) Specific requirements for remote biometric identification

- **Biometric identification & biometric authentification**

Bitkom Assessment

Already now, several security and data protection requirements apply to all applications in the field of remote biometric identification. It must be clearly worked out which additional requirements are necessary by classifying remote biometric identification as high-risk, given the regulatory status quo.

E. Adresses

Bitkom Assessment

The general statement that each obligation should address those actors who are best placed to address any potential risk is worrying. Those undertakings that are actually most

responsible for causing a risk must not be left of the hook and should always be in the first place regarding obligations to mitigate risks ("**polluter pays principle**").

Roles of different economics operators

- **Developer**
- **Deployer**
- **Producer**
- **Distributor**
- **Importer**
- **Service provider**
- **Professional or private user**

Bitkom Assessment

Are the boundaries of roles of the different economic operators and responsibilities clear? In general the Commission's objective to assign clear responsibilities to each of the different economic operators is to be welcomed.

We also recommend adding to the roles under „E. Roles of different economics operators" the "data provider", since the private sector is increasingly involved in a division of labour in data collection/processing.

Second, there is the question about the geographic scope of the legislative intervention. In the view of the Commission, it is paramount that the requirements are applicable to all relevant economic operators providing AI-enabled products or services in the EU, regardless of whether they are established in the EU or not. Otherwise, the objectives of the legislative intervention, mentioned earlier, could not fully be achieved.

Bitkom Assessment

If new regulations are created, they must be based on the existing legal framework and must not create additional and protectionist barriers for data and AI-based products from outside the EU.
It is very important that competent authorities are able to verify quickly and with legal certainty if the relevant requirements are met by economic operators offering AI-enabled products or services from outside the EU. The mentioned mutual recognition agreements

with third countries are central at this point. Experiences and best practices from other sectors[8] must be taken into account.

Furthermore it is important that appropriate transition periods or cut-off date regulations are created for applications that are already in use.

F. Compliance and Enforcement

Bitkom Assessment

Central to these considerations is the competence of the relevant competent authorities who are to carry out these assessments. The assessments have the potential to mutate into massive bottlenecks in the market launch and therefore in the speed of innovation uptake of the EU economies and societies.

- **prior conformity assessment would be necessary to verify and ensure that certain of the above mentioned mandatory requirements applicable to high-risk applications (see section D above) are complied with.**

Bitkom Assessment

We recommend, besides mandatory conformity assessements for high-risk AI applications to additionaly consider the potential of self-regulation.

Furthermore, it must be clearly worked out which requirements and related standards really have to be fulfilled within the framework of a prior conformity assessment. This varies from field of application to field of application and must be considered sector-specifically. This in turn makes it clear again how important it is that the assement frameworks discussed are linked with existing vertical framework.

- **The prior conformity assessment could include procedures for testing, inspection or certification. It could include checks of the algorithms and of the datasets used in the development phase.**

Bitkom Assessment

[8] For example: Healthcare mutual recognition (Link). Additionally the European Commission declared the data protecion rules of 13 countries as adequate to the european framework (see European strategy for data, p.5/35; Link)

We would like to emphasize again how important a cost-benefit analysis and simple practical implementation is in this context. With regard to testing, inspection and certification procedures, the relevance of self-regulation and the standards set by companies must be reaffirmed again. We welcome that the EU-Comission wants to use input of stakeholders and the European standards organisations in this context.

We also welcome that the conformity assessements should be part of the conformity assessement mechanisms that already exist and should be closely linked to them.

When designing and implementing a system relying on prior conformity assessments, particular account should be taken of the following:

- **Not all requirements outlined above may be suitable to be verified through a prior conformity assessment. For instance, the requirement about information to be provided generally does not lend itself well for verification through such an assessment.**

Bitkom Assessment

It must be clearly worked out which requirements have to be fulfilled within the framework of a prior conformity assessment. This varies from field of application to field of application and must therefore be considered sector-specifically. This in turn makes it clear again how important it is that the assement framworks discussed are linked with existing vertical framework.

- **Particular account should be taken of the possibility that certain AI systems evolve and learn from experience, which may require repeated assessments over the life-time of the AI systems in question.**

Bitkom Assessment

Similiar answer as for D. „Requirements“ /d. „Robustness and accuracy“: Requirements ensuring that outcomes are reproducible: In several applications new versions of AI systems are updated at short intervals. Therefore, there must be clear standards and rules which define when an additional assessement is actually necessary. If a repeated assessment is applied it is very important to design it in a way which minimises the additional administrative costs.

- **The need to verify the data used for training and the relevant programming and training methodologies, processes and techniques used to build, test and validate AI systems.**

Bitkom Assessment

As these considerations are a direct consquenence of D. „Requirements"/a. „Training data" & b „Keeping of records and data" we refer to our assessements from this part.

- **In case the conformity assessment shows that an AI system does not meet the requirements for example relating to the data used to train it, the identified shortcomings will need to be remedied, for instance by re-training the system in the EU in such a way as to ensure that all applicable requirements are met.**

Bitkom Assessment

In general, there must be operationally easy to handle criteria, which can be used to decide whether AI applications and systems meet the relevant requirements. AI applications trained with non-European data must be treated in the same way as systems trained with European data in this context. Disproportionate protectionist restrictions on non-European data must be prevented.

Furthermore, the recent developments of the Covid-19 crisis showed how important high-quality data and AI applications are for society as a whole to get necessary insights in the development and fight against Covid-19. Debates about restrictions of the use of non-European datasets and AI applications must always keep in mind the overall trade-off between risks and potentials.

- **Ex-ante and ex-post controls**

Bitkom Assessment

The role of a potential life cycle scheme for product security must be taken into account. The role of ex-post controls in general needs to be specified more specifically and based on standards: when are they necessary? How do they relate to ex ante conformity assessments and to the discussed repeated assessments? Overall, from a life cycle compliance

and enforcement perspective, the entire administrative burden must be kept in mind and must be minimized given an aspired level of safety.

G. Voluntary labeling for no high-risk AI applications

- **For applications which are not high risk the possiblity of voluntary labelling is discussed**

Bitkom Assessment

We support the goal of voluntary labelling. However, we have different views and risks as to whether voluntary labeling can achieve these goals. Provisions linked to such labels for low-risk AI must, on the one hand, support trust while, on the other hand, must not be overly burdensome. Otherwise such labels will not be used on a voluntary basis. Applicable provisions falling under the label could relate to transparency, robustness and human oversight. Rules such around enforcement and legal remedies for users should not apply in the same manner as under high-risk applications, which couold result is more severe harm.

In our opinion, the proposed approach oversimplifies the concept of trustworthiness which will be more effectively built by brands and determined by the alignment of incentives and whether the performance of AI systems is meeting consumers' expectations.

We also see the potential for additional uncertainty and confusion through voluntary labelling and a variety of different certificates and labels.

- **Once the developer or the deployer opted to use the label the requirements would be binding**

Bitkom Assessment

We agree with this approach in principle. But if the developer/deployer decides to not use the voluntary label anymore it must be possible to opt-out again from the framework.

H. Governance

- **Given already existing structures such as in finance, pharmaceuticals, aviation, medical devices, consumer protection, data protection, the proposed governance**

structure should not duplicate existing functions. It should instead establish close links with other EU and national competent authorities in the various sectors to complement existing expertise and help existing authorities in monitoring and the oversight of the activities of economic operators involving AI systems and AI-enabled products and services.

Bitkom Assessment

Agreement in principle- If additional regulation is necessary in certain areas, it should always be linked to existing sectoral regulation with the corresponding existing structures. The sectoral structures contain the historical regulatory expertise. Regulatory approaches, if necessary, should build on these vertical regulatory frameworks.

- **The EU enjoys excellent testing and assessment centres and should develop its capacity also in the area of AI. Economic operators established in third countries wanting to enter the internal market could either make use of designated bodies established in the EU or, subject to mutual recognition agreements with third countries, have recourse to third-country bodies designated to carry out such assessment.**

Bitkom Assessment

These observations should not lead to additional and protectionist barriers for data and AI-based products from outside the EU.

Stellungnahme

Rechtsfragen der digitalisierten Wirtschaft: Haftung für Systeme Künstlicher Intelligenz

Juni 2020

Zusammenfassung

Derzeit wird auf vielen Ebenen über die Regulierung von Künstlicher Intelligenz nachgedacht. Der Fokus liegt dabei bisher auf dem Haftungsrecht. Ziel der Überlegungen ist es, dass beim Einsatz Künstlicher Intelligenz das bisher erreichte rechtliche Schutzniveau nicht unterschritten wird. Die vielen aktuell vorgeschlagenen und diskutierten Rechtsänderungen würden bei gleichzeitiger Umsetzung jedoch in eine nicht gerechtfertigte Überregulierung münden, die eine weitere Entwicklung von Künstlicher Intelligenz in der EU stark bremsen, wenn nicht sogar verhindern würde. Aus Sicht des Bitkom sind insbesondere folgende Punkte zu berücksichtigen:

1. Das geltende Recht enthält flächendeckende Haftungsvorschriften für Schäden, die beim Einsatz technischer Geräte verursacht werden. Damit sichert es Risiken, die durch Systeme mit Künstlicher Intelligenz begründet werden können, zur Genüge ab. Individuelle Rechte und Interessen möglicher Anwender werden durch Diskriminierungsverbote, Produktsicherheitsvorschriften und Haftungsnormen angemessen geschützt. **Weder im Haftungsrecht noch bei der Marktzugangskontrolle ist es sinnvoll, neue Vorschriften zu schaffen, die nur für Anwendungen mit Künstliche Intelligenz gelten**. Bei einer Anpassung der geltenden Haftungs- und Marktzugangsvorschriften ist darauf zu achten, dass besondere Aspekte von Systemen Künstlicher Intelligenz von den Haftungsvorschriften angemessen erfasst werden. Dabei sind auch mittelbare Auswirkungen sorgfältig zu evaluieren, die solche Änderungen auf Vertragsbeziehungen im B2C- und B2B-Geschäftsverkehr haben können.
2. Das geltende Haftungsrecht verfolgt richtigerweise einen **technologieneutralen Ansatz,** der im Hinblick auf die Bewertung von Produktsicherheit und Haftung für Künstliche Intelligenz beibehalten werden sollte. Dabei orientiert sich die Haftung an Anwendungsbereich und typischem Gefahrenpotenzial eines Gerätes, nicht an der eingesetzten Technologie.[1] Rechtsänderungen sollten daher nicht an den zu allgemeinen Begriff der Künstlichen Intelligenz anknüpfen.

[1] So gilt z.B. die Halterhaftung für Kraftfahrzeuge unabhängig davon, ob das Fahrzeug mit Diesel, Benzin oder elektrisch angetrieben wird.

Bitkom
Bundesverband
Informationswirtschaft,
Telekommunikation
und Neue Medien e.V.

Thomas Kriesel
Bereichsleiter Steuern, Unternehmensrecht und -finanzierung
T +49 30 27576-146
t.kriesel@bitkom.org

Albrechtstraße 10
10117 Berlin

Präsident
Achim Berg

Hauptgeschäftsführer
Dr. Bernhard Rohleder

3. Es besteht derzeit **kein Bedarf für eine spezifische Gefährdungshaftung** mit Pflichthaftpflichtversicherung für Betreiber von Hochrisiko-Systemen mit Künstlicher Intelligenz, da die Risiken solcher Systeme (z.B. autonome Kraftfahrzeuge, Drohnen, autonome Reinigungsmaschinen) bereits durch bestehende Regelungen (nationale Straßenverkehrs- und Luftverkehrsgesetze) angemessen abgesichert sind. Bei einer dynamischen Ausweitung der Definition von Hochrisiko-Systemen besteht die Gefahr, dass Betreiber und Hersteller kurzfristig vor deutlich strengere Anforderungen in Bezug auf die Haftung und Herstellung von Systemen Künstlicher Intelligenz gestellt werden. Konkrete gesetzliche Regelungen für neue Hochrisiko-Systeme sollten daher frühestens dann und nur jeweils für bestimmte Bereiche eingeführt werden, wenn solche Hochrisiko-Systeme außerhalb der bereits durch bestehende gesetzliche Regelungen abgesicherten Bereiche vor der Markteinführung stehen. Andernfalls drohen aufgrund der Vielfalt verschiedener Systeme Künstlicher Intelligenz in bestimmten Bereichen unsachgemäße Regelungen.
4. KI-Systeme kommen derzeit und auch mittelfristig absehbar **nur als sog. „schwache" Systeme Künstlicher Intelligenz bzw. als Assistenzsysteme zum Einsatz**, die entweder Empfehlungen für eine abschließende menschliche Entscheidung liefern oder deren Lern- und Entscheidungsvorgänge in von Menschen vorgegebenen Bahnen und innerhalb vorgegebener Grenzen erfolgen. Deshalb kann solchen Systemen **grundsätzlich keine besondere Gefahrneigung** nachgesagt werden. Im Gegenteil: indem sie menschliche Fehler vermeiden (z.B. verspätete Reaktion, Übersehen wesentlicher Daten oder Informationen, Subjektivität), **können KI-Systeme menschliche Fehleinschätzungen und damit Schadensrisiken reduzieren**. Wenn der Gesetzgeber diese Möglichkeiten fördern und realisieren möchte, dürfen Betreiber und Hersteller von Systemen Künstlicher Intelligenz nicht mit Haftungsrisiken belastet werden, die über die Haftungsrisiken bei anderen Technologien hinausgehen. Auch aus diesem Grund dürfen Regelungen, die sich an der zumeist negativ konnotierten Komplexität, Autonomie oder Konnektivität eines Systems orientieren, nicht pauschal an den Einsatz Künstlicher Intelligenz anknüpfen.
5. Auch bei Systemen Künstlicher Intelligenz kann eine **absolute Fehlerfreiheit berechtigterweise nicht erwartet** werden. Entsprechend darf eine Sanktion den Hersteller oder Betreiber eines Systems Künstlicher Intelligenz nicht schon dann treffen, wenn ein Status der Fehlerlosigkeit nicht erreicht wird.

Inhalt

Seite

1 Einleitung

Der Einsatz von Künstlicher Intelligenz verspricht neue Möglichkeiten und Verbesserungen für Wirtschaft und Gesellschaft. Wie bei jeder neuen Technologie sind aber Risiken und Gefahrenpotenziale nicht ausgeschlossen.

Die Herausforderung für den Gesetzgeber besteht nun darin, den Rechtsrahmen für den Einsatz von Künstlicher Intelligenz so auszugestalten, dass Chancen der Technologie genutzt und gleichzeitig Risiken begrenzt werden. Allerdings muss der Gesetzgeber dabei nicht bei „Null" anfangen; denn das geltende Recht kennt bereits eine Vielzahl von Vorgaben, insbesondere von Haftungsvorschriften, die auch auf Systeme mit Künstlicher Intelligenz anwendbar sind. Daher ist zu fragen, ob und wieweit diese Rechtsgrundlagen ausreichen oder angepasst oder ergänzt werden müssen.

In diesem Positionspapier untersucht Bitkom die Haftungsvorgaben des geltenden Rechts für Anwendungen mit Künstlicher Intelligenz und evaluiert möglichen Handlungsbedarf.

2 Definition der Künstlichen Intelligenz

2.1 Begriffliche Annäherung

Es gibt bereits eine Vielzahl von Definitionen für Künstliche Intelligenz.[2] Gemeinsam ist diesen Definitionen der Aspekt, dass Künstliche Intelligenz (im Folgenden auch als KI bezeichnet) intelligentes menschliches Verhalten in Computer-Anwendungen simuliert oder zumindest diesen Anspruch hat. Künstliche Intelligenz ist damit ein Oberbegriff für Systeme, die Computer befähigen, einzelne kognitive Aufgaben, z.B. Bild-, Text-, Sprach- oder Mustererkennung zu bewältigen. Die von der EU-Kommission eingesetzte hochrangige Expertengruppe für Künstliche Intelligenz umschreibt KI-Systeme und ihre Fähigkeiten wie folgt: „Systeme der Künstlichen Intelligenz (KI-Systeme) sind vom Menschen entwickelte Softwaresysteme (und gegebenenfalls auch Hardwaresysteme), die in Bezug auf ein komplexes Ziel auf physischer oder digitaler Ebene handeln, indem sie ihre Umgebung durch Datenerfassung wahrnehmen, die gesammelten strukturierten oder unstrukturierten Daten interpretieren, Schlussfolgerungen daraus ziehen oder die aus diesen Daten abgeleiteten Informationen verarbeiten, und über das bestmögliche Handeln zur Erreichung des vorgegebenen Ziels entscheiden."[3]

Da diese Begriffsbestimmung noch keine ausreichende Klarheit für eine juristische Bewertung bietet, sollen im nachfolgenden Blick auf verschiedene Ansätze und Techniken

[2] Vgl. z.B. Bitkom Leitfaden „Künstliche Intelligenz", S. 28, 29, die Antwort der Bundesregierung auf Frage 1 einer Anfrage von Bündnis 90 / Grüne in Drs. 19/1982 vom 27.04.2018 oder die Strategie Künstliche Intelligenz der Bundesregierung 2018, S. 4
[3] Hochrangige Expertengruppe für Künstliche Intelligenz: „Eine Definition der KI: wichtigste Fähigkeiten und Wissenschaftsgebiete", April 2019, S, 6

der Künstlichen Intelligenz die Besonderheiten und Neuerungen der Technologie identifiziert werden.

2.2 Erscheinungsformen Künstlicher Intelligenz

Systeme der Künstlichen Intelligenz umfassen regelbasierte Systeme (Expertensysteme) und lernende Systeme. Gemeinsam ist allen derzeit eingesetzten Erscheinungsformen von KI, dass sie sich in den Grenzen der jeweiligen Programmierung und im Rahmen eines eng umgrenzten Aufgabenbereichs bewegen. Sie geben sich die zu erreichenden Ergebnisse und die Auswertungsverfahren für die Analyse von Inputdaten nicht selbst vor. Nutzungszweck, Ziel- und Analysevorgaben werden von Menschen programmiert und sind unveränderbar. Aufgrund dieser Beschränkungen bezeichnet man die derzeitigen KI-Systeme üblicherweise als „schwache KI“. Systeme mit „starker KI“, die selbstständig jede gedankliche Aufgabe erledigen könnten, zu der auch der Mensch in der Lage ist, selbständig den von ihrem Entwickler gesetzten Handlungsrahmen verlassen und sich selbstständig jeder gedanklichen Zielvorgabe annehmen könnten, sind gegenwärtig rein visionär. Systeme der „starken KI“ werden voraussichtlich auch in mehreren Jahrzehnten (wenn überhaupt) noch nicht existieren. Diese Stellungnahme setzt sich im Folgenden nur mit KI-Systemen auseinander, die gegenwärtig bereits Anwendung finden.

Expertensysteme übersetzen bewährtes Problemlösungswissen und bekannte Zusammenhänge in computergerechte Modelle und Regeln, die zur Lösung vorgegebener Aufgaben und zur Analyse von Datenbeständen eingesetzt werden. Dabei werden dem System bekannte Fakten und Zusammenhänge (Wissen) sowie Regeln für die Interpretation der Daten und für Schlussfolgerungen (Wenn-dann-Beziehungen) vorgegeben (symbolischer Ansatz). Das System wendet diese Anweisungen auf neue Datenbestände und Sachverhalte an. Anhand der zur Anwendung kommenden Wenn-dann-Beziehungen kann im Nachgang eindeutig nachvollzogen und erklärt werden, wie das System zu einem Ergebnis gekommen ist und ob das Ergebnis den vorgegebenen Regeln und Modellen entspricht. Der Output des Systems kann nach bekannten Schritten und Regeln deterministisch aus einem Input abgeleitet werden.

KI-Systeme, die auf Konzepten des **Maschinellen Lernens** (ML) beruhen, befähigen Computer, aus Daten Schlussfolgerungen und Informationen zu generieren, ohne dass ihnen dazu Regeln für das Schließen oder für die Informationsgewinnung vorgegeben sind. Vielmehr kann das System aus Trainingsdatensätzen (Beispieldaten) relevante Muster, Zusammenhänge und Informationen ableiten, verallgemeinern und für die Analyse neuer Datenbestände und Fragestellungen nutzbar machen. Dazu bauen Algorithmen[4] mit Hilfe der Muster aus den Trainingsdaten für eine bestimmte

[4] Unter Algorithmen sind aufeinander bezogene Anweisungen zu verstehen, die in definierten Einzelschritten ausgeführt werden, um ein Problem zu lösen oder eine Aufgabe zu bewältigen. Ein Algorithmus benötigt Eingangsdaten, die in vorgegebenen Prozessschritten zu Ausgangsdaten verarbeitet werden.

Fragestellung ein statistisches Modell (Abbildung eines Sachverhalts in einer mathematischen Funktion) auf. Andere Algorithmen übertragen die im Modell repräsentierten Lernergebnisse (das gelernte „Wissen“) auf neue Datenbestände und Fragestellungen, sodass das System bei Lernerfolg auch unbekannte Eingabedaten beurteilen kann (Lerntransfer). Das Modell wird im Laufe des Lernprozesses durch Auswertung weiterer Trainingsdaten über verschiedene Parameter nachjustiert und angepasst, um das Lernergebnis und damit die Ausgabegenauigkeit zu verbessern. Auf diese Weise können KI-Systeme ihre Ausgabemuster im Rahmen der vorgegebenen Parameter ändern und an verschiedene Situationen bzw. an einen Lernfortschritt anpassen.

Maschinelles Lernen wird mit verschiedenen technischen Ansätzen realisiert, gängige Ansätze sind überwachtes Lernen (supervised learning), unüberwachtes Lernen (unsupervised learning), bestärkendes Lernen (reinforcement learning) und Deep Learning mit Hilfe neuronaler Netze.

Beim **überwachten Lernen** wird jedem Trainingsdatensatz ein korrektes Ausgabeergebnis zugeordnet. Die aus den Trainingsdaten zu erlernenden Muster werden dem System auf diese Weise vorgegeben und sollen von ihm nach dem Lernprozess auch in unbekannten Daten identifiziert werden können.

Beim **nicht überwachten Lernen** erzeugt das ML-System nach der Eingabe von Trainingsdaten selbständig ein statistisches Modell, das diese Trainingsdaten beschreibt, erkannte Kategorien und Zusammenhänge enthält und somit Vorhersagen ermöglicht.

Beim **bestärkenden Lernen** beruht der Lernfortschritt darauf, dass das ML-System zunächst ein eigenes Entscheidungsmodell entwickelt. Im Anschluss daran erhält es zu getroffenen Entscheidungen die Rückmeldung (d.h. ein Feedback), ob es gut (Gewährung einer „Belohnung“) oder schlecht (Gewährung einer „Bestrafung“) entschieden hat. Das System passt daraufhin sein Systemverhalten dergestalt an, dass es anstrebt, die Belohnungen zu maximieren. Ziel des Lernens ist es, im Laufe der Zeit mit möglichst vielen Rückmeldungen das System zu optimieren.

Manche maschinelle Lernverfahren arbeiten mit Algorithmen, die auf dem Konzept der **neuronalen Netze** beruhen. Das Konzept besteht, in Anlehnung an die Funktionsweise des menschlichen Gehirns, aus einem Netzwerk kleiner Verarbeitungseinheiten, den Künstlichen Neuronen, die durch zahlreiche gewichtete Verbindungen, den künstlichen Nervenverbindungen, miteinander verknüpft sind. Die künstlichen Neuronenstrukturen sind üblicherweise in Schichten (Layers) angeordnet. Die Schicht zur Aufnahme der Eingangsdaten, z.B. einzelne Farbpunkte eines Bildes, bezeichnet man als Input Layer. Ein Output Layer berechnet daraus die Ergebnisse und kann z.B. angeben, ob auf dem Bild eine Katze oder ein Hund dargestellt ist. Während der Lernphase des neuronalen Netzes

werden die (mathematischen) Gewichtungen der Verbindungen und die Aktivierungsfunktionen der Künstlichen Neuronen angepasst. Auf diese Weise wird die Abweichung zwischen dem vom System berechneten Ausgabeergebnis und dem erwarteten Ausgabeergebnis schrittweise minimiert.

Zu den derzeit erfolgreichsten maschinellen Lernverfahren, die neuronale Netze nutzen, gehört das **Deep Learning**. Dieser Ansatz macht sich die Tatsache zunutze, dass neuronale Netze, die zwischen einer Schicht von Eingabeneuronen und einer Schicht von Ausgabeneuronen noch weitere damit verbundene Neuronenschichten aufweisen, über eine größere Ausdrucksmächtigkeit verfügen und daher komplexere Sachverhalte durch maschinelles Lernen bewältigen können.

Manche Systeme des maschinellen Lernens sind darauf ausgerichtet, dass der Trainingsprozess vor der produktiven Nutzung beendet wird. Das gelernte Verhalten wird dann auch bei neuartigen Inputdaten unverändert beibehalten. Es gibt aber auch ML-Systeme, die im produktiven Anwendungsbetrieb ihre Lerntätigkeit fortsetzen und daher ihren Output über das bisher Gelernte hinaus variieren und anpassen können. Das Verhalten letztgenannter ML-Systeme kann insoweit ganz wesentlich von den Daten abhängen, die dem System nach der Entwicklungsphase vom Anwender zugeführt werden.

Alle maschinellen Lernverfahren beruhen wesentlich auf der Auswertung von Beispieldaten und auf der Ableitung von Mustern aus diesen Beispieldaten. Dabei sind die Muster teilweise von Menschen selbst nicht zu erkennen und die Ausgabeergebnisse abhängig von den Beispieldaten und dem Lernfortschritt unterschiedlich. Verfahren des maschinellen Lernens im Allgemeinen und neuronale Netze bzw. Deep-Learning im Besonderen nutzen teilweise sehr komplexe und auf statistischen Wahrscheinlichkeiten beruhende Techniken und Verfahren. Daher sind Lernfortschritt, Anpassungen während des Lernverhaltens und das Zustandekommen einzelner Ergebnisse nicht in jedem Einzelfall exakt nachvollziehbar, erklärbar oder vorhersagbar.[5] Es gibt aber Methoden und Verfahren, Transparenz und Erklärbarkeit bis zu einem gewissen Grad herzustellen.[6] An deren weiterer Verbesserung wird intensiv geforscht.

2.3 Anwendung Künstlicher Intelligenz

KI-Systeme können in begrenztem Umfang einzelne Tätigkeiten von Menschen übernehmen und diese schneller und mit konstanter Qualität ausführen. Dies gilt

[5] Dies ist allerdings bei menschlichen Entscheidungen auch nicht anders. Der Mensch kann lediglich bestimmte Motive angeben, die ihn bei der Entscheidungsfindung beeinflusst haben. Damit ist jedoch das menschliche Verhalten und die menschliche Entscheidungsfindung nicht komplett erklärbar
[6] Vgl. die Bitkom-Publikation „Blick in die Blackbox - Nachvollziehbarkeit von KI-Algorithmen in der Praxis“, 2019

besonders für die schnelle Analyse großer Datenmengen. In der Erkennung von Mustern und Korrelationen in großen, auch unstrukturierten Datenbeständen und in der darauf aufbauenden Informationsgewinnung (z.B. Identifizierung von Personen auf Fotos oder aus einer Kameraüberwachung oder Texterkennung in gescannten Dokumenten) haben KI-Systeme ihre besonderen Stärken. Je größer der für eine Entscheidung zu analysierende Datenbestand ist, desto eher sind KI-Systeme dem Menschen bei der Entscheidungsfindung überlegen. Dabei müssen die Daten nicht manuell eingegeben, sondern können aus verschiedenen Quellen automatisch über Schnittstellen zugeführt werden (z.B. über Sensoren, Kameras, Vorsysteme).

In allen Bereichen, in denen KI-Systeme bisher unternehmerisch genutzt werden, sollen sie Tätigkeiten und Entscheidungen menschlicher Anwender unterstützen und ihnen ihre Aufgaben erleichtern. Häufig stellen KI-Systeme nach Analyse der Eingabedaten dem Menschen nur einen Vorschlag für eine mögliche Handlung zur Verfügung, welchen der Mensch einer Nachbetrachtung unterziehen und ablehnen kann. Sie finden Anwendung z.B. bei der Qualitätskontrolle von Stoffen, dienen zur Bestimmung der Autorenschaft von Texten[7], helfen bei der Diagnose von Krankheiten, unterstützen bei der Wartung von Anlagen und beim Design komplexer Industrieaggregate.[8]

Dabei können KI-Systeme als Komponenten in komplexe IT-Anwendungen und Anlagen integriert sein und diesen Anwendungen ganz neue Qualitäten verleihen. So können KI-Systeme Produkten Dialogfähigkeit verleihen (z.B. Amazon Alexa, Apple Siri, etc.), oder den Autonomiegrad von Robotertechnologien erhöhen. Ein bekanntes und häufig diskutiertes Beispiel dieser Fallgruppe ist das autonome Fahren. Autonome Fahrzeuge nehmen mit Hilfe verschiedener Sensoren ihre Umgebung wahr, leiten daraus Informationen über ihre eigene Position und die der anderen Verkehrsteilnehmer ab, kontrollieren mittels Aktoren das Fahrverhalten selbstständig und können Strategien zur Ansteuerung eines vorgegebenen Fahrziels entwickeln. In Zusammenarbeit mit der Navigationssoftware können sie das Fahrziel ansteuern und Kollisionen auf dem Weg vermeiden.

2.4 Konsequenzen für die rechtliche Beurteilung

Festzuhalten ist zunächst, dass es sich bei Anwendungen der Künstlichen Intelligenz im Kern um Programme zur Datenverarbeitung, also um – ggf. in physische Produkte eingebettete und nur in Verbindung mit diesen funktionsfähige – Software handelt. Wie vorstehend gezeigt, werden KI-Systeme jedoch in ganz unterschiedlichen Erscheinungsformen und mit einer Vielzahl unterschiedlicher technologischer Ansätze

[7] Bitkom-Publikation „Blick in die Blackbox - Nachvollziehbarkeit von KI-Algorithmen in der Praxis“, 2019, Kapitel 2.2 und Kapitel 4.2
[8] Vgl. Bitkom-Leitfaden „Konkrete Anwendungsfälle von KI & Big-Data in der Industrie“, 2019, Kapitel 9, 10 und 12

realisiert. Diese Vielfalt steht einer pauschalen und differenzierungslosen rechtlichen Beurteilung entgegen.

So beruhen KI-Systeme in ihrer Unterart der Expertensysteme auf klar definierten und fest programmierten Regelwerken und Softwarebefehlen. Aufgrund ihrer Determiniertheit kommt es bei diesen Systemen gerade nicht zu den für KI-Systeme beklagten Einbußen an Transparenz und Nachvollziehbarkeit. Expertensysteme lassen keine Besonderheiten im Vergleich zu herkömmlichen Computeranwendungen erkennen, sodass insoweit auch keine neue rechtliche Bewertung oder spezifische Neuregelungen gerechtfertigt sind.

Die Aufteilung der rechtlichen Verantwortung zwischen Betreiber und Hersteller beim maschinellen Lernen hängt nicht zuletzt auch davon ab, ob das System seine Lerntätigkeit im Produktivbetrieb fortsetzt, oder ob die Lernfunktion mit Auslieferung des Systems abgeschaltet wird. Denn bei fortlaufender Lerntätigkeit können fehlerhafte Ausgabeergebnisse des Systems nicht allein in der Verantwortung des Herstellers liegen, sondern auf vom Betreiber eigenmächtig zugeführte Daten zurückzuführen sein.

Zwar kann Künstliche Intelligenz in autonomen Systemen eingesetzt werden, autonome Systeme und Künstliche Intelligenz sind aber nicht gleichzusetzen. Das wesentliche Merkmal autonomer Systeme ist eine Selbststeuerung ohne kontinuierlichen Eingriff menschlicher Akteure in den Systemablauf. Künstliche Intelligenz wird dagegen überwiegend in Assistenzsystemen eingesetzt, die eine menschliche Entscheidung unterstützen, diese aber gerade nicht ersetzen.

Die skizzierte Vielfalt der möglichen Anwendungsfelder, der unterschiedlichen Techniken und jeweils unterschiedlich zu gewichtender Einflussfaktoren führen zu der Erkenntnis, dass der Begriff der Künstlichen Intelligenz allein kein geeigneter Anknüpfungspunkt für eine rechtliche Bewertung oder weitere, noch zu entwickelnde Regelungen darstellt. Vielmehr muss eine juristische Betrachtung die jeweilige Erscheinungsform (Anwendung) Künstlicher Intelligenz und den Einsatz der Technologie im konkreten Fall berücksichtigen.

3 Haftungsrelevante Besonderheiten von Künstlicher Intelligenz

Systeme Künstlicher Intelligenz werden von Menschen geschaffen und sind wie andere Softwareprodukte nicht frei von Fehlern.[9] Es kann also zu Situationen kommen, in denen beim Einsatz von KI-Systemen Schäden verursacht werden. Dabei darf der Rechtsschutz für Schäden durch Künstliche Intelligenz nicht hinter dem etablierten Rechtsschutz für Gefahren anderer Technologien zurückbleiben.[10] Um die Angemessenheit der geltenden

[9] Vgl. z.B. zur Fehlerhaftigkeit von Software die Bitkom-Publikation „Zur Sicherheit softwarebasierter Produkte“
[10] Insoweit unterstützt Bitkom die Forderung im KI-Weißbuch der EU-Kommission COM(2020) 65 final, S. 11 f.

Haftungsregelungen und ihre Anwendung auf KI-Systeme beurteilen zu können, ist zunächst festzustellen, wo die Risiken und Fehlerquellen von KI liegen, die mit den Haftungsregelungen adressiert werden sollen.

3.1 Fehlerquellen und Risiken

Die Auseinandersetzung mit Risiken der Künstlichen Intelligenz macht einen großen Teil aktueller Veröffentlichungen zu Rechtsfragen der KI aus. In ihrem Weißbuch für Künstliche Intelligenz und den begleitenden Veröffentlichungen identifiziert die EU-Kommission folgende Risiken und Besonderheiten Künstlicher Intelligenz, die aus ihrer Sicht zu Schwierigkeiten bei der Durchsetzung von Schadensersatzansprüchen führen könnten und daher auch rechtliche Relevanz haben.

- Komplexität: KI-Komponenten können in andere Anwendungen integriert, KI-Produkte können mit anderen digitalen Systemen vernetzt werden. Aufgrund der Vielzahl der möglicherweise bei Nutzung und Herstellung eines KI-Produkts involvierten Akteure und wegen der möglichen Vielfalt interagierender Komponenten könne es zu Problemen bei der Zurechnung von KI-Fehlern und zu Schwierigkeiten bei Darlegung und Beweis von Haftungsansprüchen wegen Schäden durch KI kommen.[11]
- Autonomie: In der Fähigkeit, auf Wahrnehmungen aus der Umgebung ohne vorab festgelegte Anweisungen reagieren zu können (Autonomie) sieht die EU-Kommission ein Hauptmerkmal Künstlicher Intelligenz. Daraus ergibt sich nach Ansicht der Kommission das Risiko, dass das System die Grenzen der ursprünglich vorgesehenen bestimmungsgemäßen und vom Hersteller beabsichtigten Verwendung überschreitet. Zu diesem Risiko trage auch die Selbstlernfunktion von KI-Systemen bei. Dies mache eine Risikobewertung durch die Hersteller auch nach Inverkehrbringen des Systems sowie eine menschliche Aufsicht über das System erforderlich.[12]
- Konnektivität: Auf Konnektivität ausgerichtete digitale Technologien sind nach Ansicht der Kommission anfällig für Cyberangriffe. Dritte könnten sich unbefugt Zugang zum System verschaffen, um geschützte Informationen zu erhalten oder in den Ablauf des Systems einzugreifen. Geht die Konnektivität verloren, könnte das System insgesamt ausfallen. Die Anfälligkeit bestehe z.B., wenn ein Systemzugang für Updates geöffnet wird.[13]
- Datenabhängigkeit: KI-Systeme sind auf Daten angewiesen. Dabei kommt es bei der Entwicklung von Systemen des maschinellen Lernens (ML-Systeme) bereits auf die Trainingsdaten und deren Qualität im Hinblick auf den beabsichtigten Verwendungszweck an. Die vom System zu erlernenden Muster und Informationen

[11] Dies befürchtet die EU-Kommission in ihrem KI-Weißbuch, S. 14 f.
[12] Bericht der Kommission über die Auswirkungen künstlicher Intelligenz, des Internets der Dinge und der Robotik im Hinblick auf Sicherheit und Haftung, COM(2020) 64 final, S. 8 f.
[13] Argumentation aus dem Bericht der Kommission über die Auswirkungen künstlicher Intelligenz, des Internets der Dinge und der Robotik im Hinblick auf Sicherheit und Haftung, S. 6 f. und aus dem KI-Weißbuch der EU-Kommission, S. 17

müssen im Trainingsdatenbestand angemessen repräsentiert sein und in korrektem Zusammenhang stehen. Erfassung, Auswahl und Aufbereitung der Trainingsdaten sind also ein wichtiger Faktor für die Qualität des Systems. Lücken in Lerndatenbeständen führen zu unsicheren und fehlerhaften Einschätzungen durch das System und begünstigen ungewollte Korrelationen und Tendenzen in den Ausgabeergebnissen.

- Opazität (Blackbox-Effekt): Für die Zurechnung und den Nachweis von Fehlfunktionen in KI-Systemen und für die Durchsetzung daraus resultierender Haftungsansprüche ist Transparenz und Nachvollziehbarkeit von Kausalverläufen erforderlich. Es muss in einem Rechtsstreit überprüft werden können, ob ein entstandener Schaden auf fehlerhafte Algorithmen oder unzureichende Datenqualität zurückzuführen ist. Ergebnisse, Entscheidungswege und genutzte Entscheidungsparameter von KI-Systemen sind aber nicht immer vollständig nachvollziehbar und erklärbar. Die EU-Kommission diskutiert dieses Merkmal von KI-Produkten unter dem Begriff „Opazität".
- Veränderbarkeit: Software kann ihre Funktionen und Eigenschaften im Verlauf des Einsatzes in gewissem Umfang ändern. Die Änderungen können auf Updates, bei maschinellem Lernen zusätzlich auf der fortgesetzten Lernfunktion beruhen. Aus Sicht der EU-Kommission ergeben sich daraus Risiken, die im geltenden Haftungsrecht, das vor allem Sicherheitsrisiken im Zeitpunkt des Inverkehrbringens adressiert, nicht angemessen berücksichtigt sind.[14]

3.2 Konsequenzen für die rechtliche Beurteilung

In ihren Veröffentlichungen zu KI setzt sich die EU-Kommission vor allem mit Risiken für Verbraucher auseinander. Eine weitergehende Betrachtung für Anwendungen im Unternehmensumfeld erfolgt nicht. Jedoch wäre eine solche Differenzierung sinnvoll und notwendig, da die jeweiligen Einsatzgebiete für KI-Systeme, die Beherrschbarkeit der Einsatzrisiken für den Anwender und die möglichen Schadensszenarien sehr unterschiedlich sein können.

Des Weiteren beziehen sich die Ausführungen der Kommission zu einem großen Teil auf digitale Technologien im Allgemeinen, nicht speziell auf Künstliche Intelligenz.[15] So ist Autonomie kein kennzeichnendes Merkmal von KI-Systemen. Zielfunktion, Nutzungszweck und Analysevorgaben werden von Menschen programmiert. Dabei sind für jeden Hersteller Sicherheit und Beherrschbarkeit des Systems wesentlich. Da KI-Systeme nicht notwendigerweise autonom agieren, ist auch das autonome Fahrzeug kein allzu treffendes Anwendungsbeispiel für Künstliche Intelligenz. Hinzu kommt, dass derzeit lediglich das teilautonome Fahren zulässig ist, d.h. der Fahrer muss jederzeit bereit sein, die Kontrolle über das Fahrzeug zu übernehmen (§ 1b StVG).

[14] Weißbuch der EU-Kommission, S. 16
[15] Der Bericht der Kommission über die Auswirkungen künstlicher Intelligenz, des Internets der Dinge und der Robotik im Hinblick auf Sicherheit und Haftung umfasst bereits dem Titel nach nicht nur Künstliche Intelligenz, sondern auch Internet der Dinge und Robotik.

Sicherheitsrisiken vernetzter Produkte führen zu Gefahren für die Allgemeinheit vor allem dann, wenn die Produkte an öffentlich zugängliche Datennetze angeschlossen sind. Dies ist ebenfalls kein kennzeichnendes Merkmal für KI-Anwendungen. Vielmehr werden KI-Anwendungen in Unternehmen überwiegend zur Analyse unternehmensinterner Daten genutzt. Diese Daten sollen regelmäßig z.B. aus Gründen des Datenschutzes oder des Geschäftsgeheimnisschutzes gerade nicht für Personen außerhalb des Unternehmens zugänglich sein.

Schließlich ist Komplexität kein kennzeichnendes Merkmal für sämtliche KI-Anwendungen. Zwar können KI-Anwendungen in komplexe Systeme mit interagierenden Komponenten integriert werden und Bestandteil einer Anwendung sein, auf die viele Akteure in unterschiedlicher Weise zugreifen. Auch können Funktionsweise, Entscheidungsmodell oder angewandte Lernverfahren für Systeme des maschinellen Lernens komplex werden. Eine KI-Anwendung kann aber auch in einem Datenanalysetool bestehen, das sowohl hinsichtlich der Analysefunktion als auch hinsichtlich des Anwendungsbereichs sehr beschränkt ist. Andererseits ist Komplexität eine Eigenschaft vieler technischer Produkte, die keine Künstliche Intelligenz beinhalten. Würde man also die Regulierung von Künstlicher Intelligenz auf Systemeigenschaften wie Konnektivität, Komplexität oder Autonomie ausrichten, käme es unweigerlich zu einer Fehlregulierung.

Unbestreitbar ist dagegen, dass KI-Systeme in besonderer Weise von Dateninput abhängen und einen von außen nicht ohne Weiteres erkennbaren Wirkungsmechanismus in den angewandten Entscheidungsmodellen und Entscheidungsverfahren aufweisen. Dies gilt aber nur für Systeme des maschinellen Lernens (ML-Systeme). Während “statische“ Expertensysteme Daten nach vorgegebenen und jederzeit nachkontrollierbaren Bearbeitungsregeln abarbeiten, ist für kontinuierlich lernende Systeme der Output nicht durch einen Input determiniert. Ein konkretes Verarbeitungsergebnis lässt sich für diese lernenden Systeme nicht in jedem Fall eindeutig vorhersagen, weil sich das Lernmodell theoretisch bei jedem neuen Dateninput anpasst und nicht alle Lösungsschritte konkret vorgegeben werden.

Dennoch halten sich die möglichen Ausgabeergebnisse lernender Systeme innerhalb eines abgegrenzten Ereignis- und Ergebnisraums (Domäne). Das System generiert Ergebnisse auf der Grundlage des von Lerndaten geformten Entscheidungsmodells, und mögliche Fehlerquellen sind bekannt. Fehler können dabei beruhen auf lückenhaften oder inadäquaten Lerndaten oder auf falschen oder ungenauen Parametern für die Mustererkennung und für das Entscheidungsmodell. Bei einer unzureichenden Komplexität des Modells (Underfitting) werden die Muster in den Trainingsdaten nicht erkannt, bei einer Überanpassung des Modells auf den Trainingsdatenbestand (Overfitting) lassen sich die erlernten Muster nicht auf unbekannte Daten übertragen. Selbst wenn also Ausgabeergebnisse und Verarbeitungsprozess eines ML-Systems nicht

immer im Einzelnen nachvollzogen werden können, lassen sich die Ursachen von unerwünschten Ausgabeergebnissen solcher Systeme von Sachverständigen lokalisieren und beurteilen.

Die Betrachtung zeigt, dass verschiedene Risiken, die typischerweise mit Künstlicher Intelligenz in Verbindung gebracht werden, vor allem im Zusammenhang mit maschinellem Lernen auftreten. Eine risikoadäquate Regulierung sollte sich daher auf Systeme des maschinellen Lernens (ML-Systeme) fokussieren.

4 Ethische Grundlagen

Die Ethik umfasst die Grundwerte und Prinzipien für menschliches Zusammenleben und menschliches Verhalten in einer Gesellschaft. Welche Konsequenzen diese Prinzipien für die Entwicklung und Anwendung von Systemen mit Künstlicher Intelligenz haben, ist bereits durch mehrere hochrangige Expertengruppen untersucht worden.[16] Auf der Basis dieser Untersuchungen haben sowohl OECD[17] als auch EU[18] Prinzipien und Leitlinien entworfen. Die auf EU-Ebene eingesetzte Expertengruppe leitet aus ethischen Grundsätzen und aus Grundrechten sieben ethische Anforderungen an Künstliche Intelligenz ab:

- Menschliche Verwaltung und Kontrolle
- technische Robustheit und Sicherheit
- Schutz der Privatsphäre und Datenqualitätsmanagement
- Transparenz
- Vielfalt, Diskriminierungsfreiheit und Fairness
- Gesellschaftliches Wohl und Umweltverträglichkeit
- Verantwortlichkeit und Rechenschaftspflicht.[19]

Ethische Bedenken bei der Nutzung von KI drücken sich aber auch in den Sorgen potenzieller Anwender aus. So wächst zwar in der deutschen Gesellschaft die Akzeptanz von KI.[20] Denn bei fehlerfreier Programmierung und nach ausreichendem Training mit geeigneten Daten schalten KI-Entscheidungen die Beeinflussung durch nicht sachgerechte menschliche Regungen, Empfindungen und Reaktionen aus, sodass KI-Entscheidungen auf der Grundlage von objektiven, sachbezogenen Kriterien getroffen werden.

[16] Zuletzt durch die Datenethikkommission, die ihr Gutachten am 23.10.2019 vorlegte.
[17] OECD Principles on AI
[18] Ethic Guidelines for Trustworthy AI
[19] Vgl. Ethik-Leitlinien für eine vertrauenswürdige KI der Hochrangigen Expertengruppe für Künstliche Intelligenz, deutsche Fassung, S. 17 ff. Die zusammenfassende Darstellung dieser Ziele findet sich auch in der Mitteilung der Kommission COM(2019)168, S. 4
[20] Vgl. Ergebnisse einer repräsentativen Umfrage für Bitkom aus dem Januar 2018 sowie einer repräsentativen Umfrage des Bitkom aus dem November 2018

Allerdings wird auch für bestimmte Einsatzbereiche wie im Beziehungsbereich, zu militärischen Zwecken oder in der Justiz Künstliche Intelligenz überwiegend abgelehnt.[21] Nach einer Umfrage der Bertelsmann-Stiftung ist 79 Prozent der Deutschen unwohl, wenn Computer über sie entscheiden, 73 Prozent der Befragten fordern ein Verbot von sogenannten vollautomatisierten Entscheidungen, die ohne menschliche Beteiligung getroffen werden. Außerdem besteht in der Bevölkerung gemäß der Umfrageergebnisse der Wunsch nach engmaschiger Kontrolle von Algorithmen.[22]

Bitkom-Bewertung

Ethische Grundsätze umfassen abstrakte und allgemeine Werte für ein allgemein als gut und richtig angesehenes menschliches Handeln. Da sie ein Idealziel vorgeben, gehen sie über konkrete Rechtspflichten weit hinaus. Daher wäre es eine Fehlinterpretation ethischer Leitlinien, wenn man diese vollumfänglich in Gesetzesrecht transformieren wollte.[23] Vielmehr muss für ethische Grundsätze im Einzelnen geprüft werden, ob sie sinnvollerweise als Rechtsnorm, als unternehmerische Selbstverpflichtung oder als moralischer Anspruch an den Einzelnen umzusetzen sind.

Zu bedenken ist weiterhin, dass ethische Grundsätze und Überzeugungen in einer Gesellschaft Wandlungen unterworfen sind. Daher dürfen die zu einem bestimmten Zeitpunkt allgemein anerkannten ethischen Grundsätze nicht in ihrer Gesamtheit absolut gesetzt werden. Das Recht muss auf veränderte ethische Überzeugungen reagieren.

Dennoch müssen ethische Grundsätze und persönliche Sorgen potenzieller Anwender von KI-Systemen ernst genommen werden. Unternehmen sind sich ihrer diesbezüglichen Verantwortung bewusst, bekennen sich zur Beachtung gesellschaftlich anerkannter Werte und richten ihre unternehmerische Tätigkeit an den ethischen Anforderungen aus.[24] Dabei greifen die Unternehmen die von der EU-Expertengruppe formulierten ethischen Anforderungen durchaus auf. So bekennen sich Bitkom-Unternehmen dazu, dass Systeme mit Künstlicher Intelligenz menschliches Handeln ergänzen und unterstützen, nicht aber Menschen überflüssig machen sollen. Die Unternehmen unterstützen Forderungen nach

[21] Nach einer Bitkom-Umfrage wollen sich Menschen bei der Partnerwahl nicht durch KI beeinflussen lassen
[22] Repräsentative Bevölkerungsumfrage der Bertelsmann Stiftung „Was Deutschland über Algorithmen weiß und denkt“ aus dem Mai 2018, S. 25 bzw. S. 29
[23] Die Hochrangige Expertengruppe für Künstliche Intelligenz sieht in ihren Ethik-Leitlinien für eine vertrauenswürdige KI, S. 2, Recht und Ethik als unterschiedliche Bereiche, die zusammenwirken sollen, aber nicht identisch sind. Auch die von der Bundesregierung eingesetzte Datenethikkommission ist in ihrem Abschlussgutachten auf S. 41 der Meinung, dass nicht jedes ethisch relevante Detail rechtlich reguliert werden sollte.
[24] So haben viele Unternehmen Leitlinien und Prinzipien für künstliche Intelligenz entwickelt, z.B. die Deutsche Telekom AG, IBM, Google. Auf der „International Joint Conference on Artificial Intelligence“ (IJCAI) Mitte Juli 2018 in Stockholm haben sich mehrere große Technologieunternehmen dazu bekannt, sich nicht an der Entwicklung von tödlichen KI-Waffen zu beteiligen, vgl. Lethal Autonomous Weapons Pledge

Transparenz beim KI-Einsatz und legen hohe Qualitätsstandards bei der Entwicklung von lernenden Systemen an.

Zu den Konsequenzen aus der ethischen Betrachtung gehört nicht zuletzt, dass die Verantwortlichkeit für den Einsatz von KI und dessen Folgen nicht an die Technik übergeben werden darf. Denn solange (natürliche oder juristische) Personen für ein KI-System verantwortlich und haftbar bleiben, besteht ein Eigeninteresse des möglicherweise Haftenden, die Technologie und ihre Ergebnisse zu kontrollieren und ihre schädlichen Auswirkungen gering zu halten. Darüber hinaus sind viele ethische Forderungen und Grundsätze bereits im geltenden Recht umgesetzt, z.B. der Schutz der Privatsphäre und der informationellen Selbstbestimmung, die Vermeidung von Personenschäden, das Verbot von Diskriminierung.

Künstliche Intelligenz ist aber nicht nur mit Risiken verbunden, die aus ethischen Gründen zu vermeiden sind, sondern hat unbestreitbare Vorteile und positive Effekte für Menschen (z.B. Verbesserung der Entscheidungsfindung, Erhöhung von Sicherheit, Verbesserungen bei Diagnose und Therapie von Krankheiten). Denn die Ergebnisse von Entscheidungsprozessen mit KI-Unterstützung sind regelmäßig mit weniger Fehlern behaftet als menschliches Verhalten im gleichen Anwendungsgebiet. Daher könnte es in vielen Fällen sogar ethisch geboten sein, Menschen den Zugang zu KI zu eröffnen und dadurch Schadensrisiken zu begrenzen.[25] Jedoch darf auch von KI letztlich nicht etwas verlangt werden, was unmöglich ist und auch von Menschen nicht leistbar wäre.

5 Haftungsgrundlagen des geltenden Rechts

Wirtschaft, Gesellschaft und Gesetzgeber haben das gemeinsame Ziel, dass der Einsatz von Künstlicher Intelligenz nicht zu Schäden führt. Spezifische Haftungsvorschriften kennt das geltende Recht dafür bisher aber nicht. Vielmehr gelten Haftungsregeln für Schäden durch technische Produkte unabhängig davon, welche Technologie im Einzelfall den Schaden verursacht hat (horizontaler Ansatz). Dabei bedeutet Haftung allgemein, dass eine Person für die Konsequenzen ihres Handelns einstehen und insbesondere durch ihr Handeln verursachte Schäden ausgleichen muss.

In der aktuellen Rechtsliteratur wird teilweise vorschnell nach neuen Haftungsregelungen für KI und autonome Systeme gerufen.[26] Als Begründung hierfür reicht den Autoren meist der Hinweis auf die Neuartigkeit der Systeme und ihre angeblichen Besonderheiten. Diese

[25] Vgl. die Argumentation von Wintermann in diesem Sinne unter https://www.netzpiloten.de/kuenstliche-intelligenz-menschen-chancen/ und https://www.piqd.de/zukunft-der-arbeit/ist-der-verzicht-auf-die-nutzung-von-kunstlicher-intelligenz-ethisch-vertretbar

[26] z.B. Borges: Rechtliche Rahmenbedingungen für autonome Systeme, in: NJW 2018 (Heft 14), S. 977 ff.

These soll hier hinterfragt und die Anwendbarkeit des geltenden Haftungsrechts auf KI untersucht werden.

5.1 Produkthaftung

a) Haftungsgrundlagen

Nach dem Produkthaftungsgesetz (ProdHaftG)[27] hat ein Unternehmen Schadensersatz zu leisten, wenn ein von ihm in Verkehr gebrachtes Produkt einen Fehler aufweist, und dieser Fehler dazu führt, dass ein Mensch getötet, sein Körper oder seine Gesundheit verletzt oder eine Sache, die für den privaten Gebrauch oder Verbrauch bestimmt ist, beschädigt wird (Integritätsschutz für besonders wichtige Rechtsgüter). In gleicher Weise haftet auch der Hersteller eines Teilprodukts, wenn dieses Teilprodukt den Schaden verursacht hat (§ 4 Abs. 1 ProdHaftG). Von der Produkthaftung nicht erfasst werden Schäden am fehlerhaften Produkt selbst und reine Vermögensschäden, die nicht unmittelbare Folge der Rechtsgutsverletzung sind (z.B. vergebliche Aufwendungen, entgangener Gewinn, Wegfall von Nutzungsmöglichkeiten). Ein Produkt ist fehlerhaft, wenn es nicht die Sicherheit bietet, die berechtigterweise zu erwarten ist (§ 3 Abs. 1 ProdHaftG). Eine Haftung nach dem ProdHaftG kann insbesondere ausgelöst werden durch:

- Konstruktionsfehler (Verstoß gegen technische Erkenntnisse bei der Herstellung),
- Fabrikationsfehler (Abweichungen von den Konstruktionsvorgaben für einzelne Produkte während der Herstellung),
- Instruktionsfehler (unzureichende Gebrauchsanweisung oder fehlende Warnung vor Gefahren).

Der geschädigte Anspruchsteller trägt die Beweislast für das Vorhandensein eines Produktfehlers, für seinen Schaden und den ursächlichen Zusammenhang zwischen Fehler und Schaden (§ 1 Abs. 4 ProdHaftG). Ein Verschulden des Herstellers, also ein vorwerfbarer Verstoß gegen eine Sorgfaltspflicht, ist keine Haftungsvoraussetzung für die Produkthaftung. Der Hersteller kann sich jedoch von einer Produkthaftung befreien, wenn er nachweist, dass in seiner Verantwortungssphäre kein haftungsrelevanter Fehler begangen wurde. Desgleichen ist die Haftung nach § 1 Abs. 2 Ziff. 5 ProdHaftG ausgeschlossen, wenn der Fehler nach dem Stand der Wissenschaft und Technik beim Inverkehrbringen nicht erkannt werden konnte (Einwand des Entwicklungsfehlers).

Die Ersatzpflicht des Herstellers für einen Produktfehler ist auf 85 Mio. Euro begrenzt (§ 10 ProdHaftG). Außerdem ist eine Selbstbeteiligung des Geschädigten bei Sachschäden von 500 Euro (§ 11 ProdHaftG) vorgesehen.

[27] Das deutsche Produkthaftungsgesetz setzt die Produkthaftungsrichtlinie der EU 85/374/EWG von 1985 um.

Mit Blick auf Künstliche Intelligenz und zunehmende Digitalisierung werfen diese Haftungsvoraussetzungen des geltenden Rechts verschiedene Anwendungsfragen auf.

b) Geltung der Produkthaftung für KI-Anwendungen

Nach dem Wortlaut des Produkthaftungsgesetzes ist als Produkt eine „bewegliche", d.h. körperliche Sache anzusehen. Ob Haftungsvorschriften für bewegliche Sachen auf KI-Systeme, die aus Computer-Algorithmen und damit aus Software bestehen, angewendet werden können, ist in der Rechtswissenschaft umstritten. Die EU-Kommission hat bereits verlauten lassen, dass nach ihrer Meinung das Produkthaftungsrecht auch auf Software Anwendung findet.[28] Eine solche Rechtsmeinung ist jedoch für die gerichtliche Auslegung der Vorschriften in den Mitgliedstaaten und durch den EuGH nicht verbindlich. Nach der Rechtsprechung unterliegt aber jedenfalls derjenige der Produkthaftung, der Geräte oder Gerätekomponenten mit eigen- oder fremdprogrammierter Software zur Steuerung oder Kontrolle des Geräts bzw. der Komponente in Verkehr bringt.[29]

c) Fehlerhaftigkeit eines KI-Systems

Nach § 3 ProdHaftG ist ein Produkt fehlerhaft, wenn es nicht die Sicherheit bietet, die berechtigterweise zu erwarten ist. Dabei richtet sich das einzuhaltende Sicherheitsniveau nach dem Stand von Wissenschaft und Technik zu dem Zeitpunkt, zu dem das Produkt in Verkehr gebracht wird. Der Stand der Wissenschaft und Technik wird vor allem durch technische Standards, z.B. DIN-, CEN- und ISO-Normen indiziert. Darüber hinaus können technische Anforderungen an besonders gefahrgeneigte Produkte auch in speziellen Gesetzen festgelegt sein, z.B. für Medizingeräte oder Pharmaprodukte. Haftungsrelevante Fehler in den Ausgabeergebnissen von KI-Systemen können sich nicht nur aus einer unzureichenden Programmierung der verwendeten Algorithmen ergeben, sondern auch aus unvollständigen oder für den angestrebten Lernzweck ungeeigneten Trainingsdaten.

d) Reichweite der Herstellerverantwortung

Die Produkthaftung greift nur ein, wenn das Produkt den schadensstiftenden Fehler bereits beim Inverkehrbringen (also bei Auslieferung bzw. Übergabe an einen Abnehmer) aufweist. Der Hersteller ist also nicht für die Weiterentwicklung von bereits im Markt befindlichen Produkten verantwortlich, und er hat auch nach Inverkehrbringen geänderte Sicherheitsstandards nur für neu in Verkehr gebrachte Produkte zu beachten. Desgleichen ist der Hersteller eines KI-Systems nach Produkthaftungsrecht nicht verantwortlich, wenn die Fehlfunktion durch Veränderungen beim Einsatz des Systems verursacht wird (z.B.

[28] Vgl. Amtsblatt EG vom 8.5.1989 Nr. C 114/42, Antwort auf schriftliche Anfrage Nr. 706/88
[29] So hat der BGH im Urteil vom 16.06.2009, Az. VI ZR 107/08 eine Produkthaftung des Fahrzeugherstellers für Gesundheitsschäden bejaht, die durch einen Fehler in der Steuerungssoftware eines Airbags verursacht wurden.

Umprogrammierung durch Nutzer, Einsatz außerhalb des vorgesehenen Anwendungsbereichs, Zuführung ungeeigneter Datenbestände). Ist der Fehler dagegen auf eine unzureichende Programmierung der Lernalgorithmen zurückzuführen, haftet der Hersteller auch dann, wenn sich der Fehler erst nach dem Inverkehrbringen zeigt.

e) Zurechnung und Beweislast

Für einen Produkthaftungsanspruch muss der Geschädigte nachweisen, dass der ihm entstandene Schaden durch ein fehlerhaftes Produkt verursacht wurde. Insbesondere bei komplexen Systemen, an deren Herstellung und Betrieb mehrere Akteure beteiligt sind, können Verursachungsbeiträge für einen konkreten Schaden schwierig zu ermitteln und nachzuweisen sein. Diese Schwierigkeit hängt aber nicht notwendigerweise mit KI-Produkten zusammen, sondern besteht bei vielen Schäden durch komplexe Systeme.

5.2 Produzentenhaftung

a) Haftungsgrundlagen

Die Produzentenhaftung leitet sich als Unterform der deliktischen Haftung aus § 823 BGB ab. Danach können Haftungsansprüche auch dann begründet sein, wenn aus einem Verhalten des Herstellers nach Inverkehrbringen seines Produktes die Verletzung eines von § 823 BGB geschützten Rechtsguts resultiert. Ein Hersteller ist also verpflichtet, auch nach dem Inverkehrbringen seines Produkts alles zu tun, was ihm nach den Umständen zumutbar ist, um Gefahren abzuwenden, die sein Produkt hervorrufen kann.[30] Allerdings muss der entstandene Schaden einem Hersteller zugerechnet werden können, d.h. kausal auf die Missachtung einer Sorgfaltspflicht (Verkehrssicherungspflicht) durch den Hersteller zurückzuführen sein.

Als Verkehrssicherungspflichten sind in diesem Zusammenhang anerkannt

- Konstruktionspflichten (der Hersteller muss eine nach Stand der Technik sichere Konstruktion wählen und absehbare Sicherheitsrisiken seines Produkts bei der Produktplanung vermeiden) und Fabrikationspflichten (der Hersteller muss seinen Betrieb so organisieren, dass Fehler durch Kontrollen entdeckt und frühzeitig beseitigt oder im Produktionsprozess vermieden werden),
- Instruktionspflichten (der Hersteller muss den Nutzer über Bedienung und mögliche Gefahrenquellen des Produkts informieren),
- Produktbeobachtungspflichten (der Hersteller muss auch nach Inverkehrbringen sein Produkt beobachten und Hinweisen auf Fehler und Gefährdungspotenziale aktiv nachgehen) und

[30] BGH, Urteil vom 16.12.2008, Az. VI ZR 170/07

- Gefahrabwendungspflichten (der Hersteller muss erkannte Gefahren bei Gebrauch seiner Produkte evaluieren, die Kunden vor Gefahren warnen und notfalls das Produkt aus dem Markt zurückrufen).

Die deliktische Haftung nach § 823 BGB schützt im Vergleich zur Produkthaftung einen weiteren Kreis von Rechtsgütern und umfasst auch Verletzungen von Persönlichkeitsrechten und eigentumsähnlichen Rechten. Reine Vermögensschäden werden aber auch nach § 823 Abs. 1 BGB nicht ersetzt. Die deliktische Produzentenhaftung kann neben der Produkthaftung geltend gemacht werden.

b) Haftung für fehlerhafte Software

Anders als nach dem Produkthaftungsgesetz ist die Haftung des Unternehmers für fehlerhafte Produkte nach § 823 Abs. 1 BGB tatbestandlich nicht auf Sachen beschränkt. Daher findet die Produzentenhaftung Anwendung auf Schäden, die durch fehlerhafte Software verursacht wurden.[31]

c) Zurechnung und Vorwerfbarkeit von Rechtsgutsverletzungen

Künstliche Intelligenz in Form des maschinellen Lernens beruht auf kontinuierlicher Anpassung des erlernten Verhaltens. Solange die Lernfunktion aktiviert ist, kann es daher schwierig sein, ein konkretes Ergebnis des Systems vorherzusagen oder im Nachhinein den exakten Entstehungsprozess eines konkreten Ausgabeergebnisses nachzuvollziehen. Damit ist jedoch nicht die Voraussehbarkeit von Schäden beseitigt. Denn die Ausgabeergebnisse, die bei einem KI-System überhaupt in Betracht kommen, sind durch das Design des Systems vorgegeben. Auch bleiben KI-Hersteller zur sorgfältigen Organisation des Programmierprozesses, zur Produktbeobachtung und zur Abwendung erkannter Gefahren verpflichtet und sind daher für Schäden, die durch eine Verletzung dieser Pflichten verursacht werden, weiterhin verantwortlich. Man könnte sogar argumentieren, dass denjenigen, der ein KI-System auf den Markt bringt, erhöhte Sorgfaltspflichten treffen, weil nicht absehbar ist, wie sich konkrete Schäden im Einzelnen realisieren. Allerdings sollten an die Ergebnisse eines durch KI gesteuerten Prozesses nicht höhere Anforderungen gestellt werden als an menschliches Handeln in einer vergleichbaren Situation.

d) Pflichtenkreis des Produzenten

Das Recht verpflichtet den Produzenten, das ihm Mögliche und Zumutbare zu tun, um Gefahren seiner Produkte abzuwenden. Bei den Anforderungen an erforderliche Maßnahmen zur Gefahrabwendung berücksichtigt die Rechtsprechung die Höhe des zu

[31] Vgl. z.B. LG Berlin, Urteil vom 08.11.2007, Az. 31 O 135/05

erwartenden Schadens, die Wahrscheinlichkeit von Schadenseintritten sowie Nutzen und Aufwand einer Sicherheitsmaßnahme. Eine absolute Gefahrlosigkeit von Produkten erwartet sie nicht. Jedoch erhöhen sich die Sorgfaltsanforderungen an den Hersteller, wenn schwerwiegende Schäden, insbesondere an Leben, Gesundheit und körperlicher Unversehrtheit mit nicht zu vernachlässigender Wahrscheinlichkeit drohen.[32]

Es kann daher genügen, dass der Hersteller vor dem Gebrauch des Produkts warnt oder bei besonderen Gefahren das Produkt vom Markt zurückruft. Wenn jedoch davon auszugehen ist, dass diese Maßnahmen nicht ausreichen, um das Sicherheitsrisiko des Produkts auf ein zulässiges Maß zu reduzieren, kann der Hersteller zu Reparatur oder Nachrüstung verpflichtet sein.[33] Danach können den Hersteller nach § 823 Abs. 1 BGB je nach Risiko im Einzelfall auch Pflichten zur Pflege oder zum Update eines KI-Systems treffen.

5.3 Haftung für Produktsicherheit

Um Risiken für Leben und Gesundheit von Menschen sowie für die Integrität von Sachwerten zu begrenzen, kennt das geltende Recht nicht nur Haftungsvorschriften zum nachträglichen Ausgleich eingetretener Schäden, sondern auch Präventionsvorschriften, die einen Schadenseintritt im Vorfeld verhindern sollen, indem sie grundlegende Sicherheitsanforderungen an Produkte gesetzlich festschreiben. Da diese Sicherheitsvorgaben beim Inverkehrbringen in der EU zu erfüllen sind und damit den Marktzugang regulieren, sind sie weitgehend durch EU-Recht harmonisiert. Das EU-Recht kennt sowohl allgemeine Produktsicherheitsvorschriften (z.B. Richtlinie 2001/95/EG) als auch sektorspezifische Vorschriften zur Produktsicherheit (z.B. Verordnung (EU) 2017/745 über Medizinprodukte oder die Maschinenrichtlinie). Ob die vorgeschriebenen Sicherheitsanforderungen erfüllt sind, wird in Marktzulassungsprüfungen über Konformitätsbewertungen festgestellt.

Zwar regeln die Produktsicherheitsvorschriften in erster Linie den Marktzugang für Produkte. Sie können jedoch auch für die Haftung der Hersteller relevant werden, soweit sie als Schutzgesetze anzusehen sind, also auch auf den Schutz bestimmter Rechtsgüter abzielen. Somit kann sich eine Haftung des Herstellers eines KI-Systems aus § 823 Abs. 2 BGB in Verbindung mit einem Schutzgesetz ergeben. Dabei trifft den Hersteller eine Ersatzpflicht auch für bloße Vermögensschäden, soweit das in Bezug genommene Schutzgesetz die Verhinderung von Vermögensschäden bezweckt.

[32] BGH, Urteil vom 17.03.2009, Az. VI ZR 176/08
[33] BGH, Urteil vom 16.12.2008, Az. VI ZR 170/07

Für das Produktsicherheitsrecht stellt sich wiederum die Frage, ob eigenständige Software und damit KI-Systeme unter den Produktbegriff subsumiert werden können. Eindeutig entschieden ist diese Frage im Medizinprodukterecht.[34]

Neben der zivilrechtlichen Haftung für Produktsicherheit aus § 823 Abs. 2 BGB, die nur bei Eintritt von Schäden in Betracht kommt, besteht eine ordnungsrechtliche Haftungsverantwortlichkeit über Straf- und Bußgeldvorschriften (z.B. §§ 39, 40 ProdSG), die durch Verstöße gegen das Schutzgesetz begründet wird, auch wenn diese Verstöße keine konkreten Schäden auslösen (Präventionswirkung des Schutzgesetzes).

5.4 Gefährdungshaftung

Der Gefährdungshaftung liegt der Gedanke zugrunde, dass derjenige, der eine Gefahrenquelle eröffnet, für die daraus resultierenden Schäden haften muss. Verschulden oder Missachtung von Sorgfaltspflichten sind nicht Voraussetzung für die Haftung. Es reicht aus, dass sich das Risiko der Gefahrenquelle in einem Schaden realisiert hat. Eine Generalklausel für Gefährdungshaftung kennt das deutsche Recht nicht. Vielmehr ist die Gefährdungshaftung auf Bereiche beschränkt, in denen schwerwiegende Personen- und Sachschäden vorkommen können. Dabei tritt die Gefährdungshaftung in besonders gefahrgeneigten Bereichen zur Vertrags-, Delikts- und Produkthaftung hinzu und stellt dem Geschädigten einen weiteren Adressaten für mögliche Ersatzansprüche zur Verfügung. So gibt es Gefährdungshaftungstatbestände im Straßenverkehr (§ 7 Abs. 1 StVG), im Luftverkehr (§ 33 LuftVG), im Bahnverkehr (§ 1 HaftPflG), für Elektrizitäts-, Gas-, Dampf- und Flüssigkeitsanlagen (§ 2 HaftPflG) und für Atomanlagen (§ 25 AtG). Bestehende Gefährdungshaftungstatbestände können nicht allgemein auf KI-Systeme übertragen werden. Allerdings sind Gefährdungshaftungstatbestände auch anwendbar, wenn beim Betrieb der gefahrgeneigten Systeme Künstliche Intelligenz zum Einsatz kommt. Wer einer Gefährdungshaftung unterliegt, hat für das daraus entstehende Risiko in der Regel eine Versicherung abzuschließen (vgl. z.B. § 1 PflVG, § 14 Allgemeines Eisenbahngesetz).

5.5 Vertragliche Haftung

a) Rechtsgrundlagen

Im Rahmen der vertraglichen Gewährleistung hat die Vertragspartei, die eine vertragliche Leistung zu erbringen hat (Anbieter), für die Qualität ihrer Leistung einzustehen. Entsprechend stehen dem Käufer nach Kaufvertragsrecht oder dem Besteller nach Werkvertragsrecht Gewährleistungsansprüche zu, wenn die Kaufsache bzw. das bestellte

[34] Nach § 3 Nr. 1 des Medizinproduktegesetzes und Art. 2 Nr. 1 der Verordnung (EU) 2017/745 über Medizinprodukte ist Software von Gesetzes wegen als Medizinprodukt anzusehen.

Werk einen Sach- oder Rechtsmangel aufweist (§ 437 BGB bzw. § 634 BGB). Nach Mietvertragsrecht muss der Vermieter dafür Sorge tragen, dass die Mietsache während der Vertragslaufzeit zum vertragsgemäßen Gebrauch geeignet ist (§ 535 BGB). Da diese gesetzlichen Regelungen ihrem Wortlaut nach nur für Sachen gelten, war lange Zeit umstritten, ob sie auch für Verträge über den immateriellen Gegenstand Software zu beachten sind. Inzwischen hat aber die Rechtsprechung geklärt, dass die vertragsrechtlichen Gewährleistungsvorschriften ohne Einschränkung auf Software Anwendung finden.[35]

Ein Sachmangel liegt vor, wenn der Vertragsgegenstand nicht für den vertraglichen Zweck oder die gewöhnliche Verwendung einsetzbar ist (§ 434 BGB). Die Eignung zur gewöhnlichen Verwendung von Software kann sich aus dem Stand der Technik ergeben.[36] Damit verweist das Recht für die Bestimmung von Gewährleistungsansprüchen für KI-Anwendungen und andere Software auf außerrechtliche, technische Vorgaben für Design, Herstellungsprozess und Qualitätskontrolle von Software. Dieser Verweis ist auch sinnvoll, damit das Recht mit den schnellen technischen Entwicklungen gerade im Softwarebereich Schritt halten kann. Daneben liegt ein Sachmangel auch vor, wenn der Vertragsgegenstand die vertraglich festgelegte Beschaffenheit nicht aufweist oder zum vertraglich vereinbarten Gebrauch ungeeignet ist. Die Vertragsparteien können also im Vertrag festlegen, welchen Qualitätsmaßstab sie an den Leistungsgegenstand anlegen wollen und in welchem Umfang den Anbieter entsprechend Gewährleistungspflichten treffen.

Die Gewährleistungsansprüche umfassen z.B. Rücktritt vom Vertrag oder Schadensersatz. Vertragsrechtliche Schadensersatzansprüche decken – anders als solche aus Produkt- oder Produzentenhaftung - auch den Ersatz reiner Vermögensschäden oder Datenverlust ab. Die Geltendmachung von Gewährleistungsrechten setzt nicht voraus, dass die Unterschreitung der Leistungsqualität auf einem Verschulden des Anbieters beruht. Die Gewährleistung greift jedoch nur im Verhältnis der Vertragsparteien zueinander.

Darüber hinaus haben Vertragspartner die Pflicht, alles zu unterlassen, was den anderen Vertragspartner schädigen könnte (§ 241 Abs. 2 BGB). Dieser vertragsrechtliche Integritätsschutz ist – anders als der deliktische Integritätsschutz – nicht beschränkt und umfasst auch rein vermögensrechtliche Positionen. Die Verletzung dieser Schutzpflicht führt zu einem Schadensersatzanspruch nach § 280 BGB.

[35] Vgl. z.B. BGH, Urteil vom 15.11.2006, Az. XII ZR 120/04, Rn. 15
[36] BGH, Urteil vom 24.09.1991, Az. X ZR 85/90

b) Zeitliche Begrenzung der Gewährleistung

Die vertragliche Verantwortlichkeit des Anbieters für Mangelfreiheit des Vertragsgegenstands bei Kauf- oder Werkverträgen reicht bis zum Zeitpunkt des Gefahrübergangs (Einräumung einer Nutzungsmöglichkeit am Vertragsgegenstand für den Kunden). Der Anbieter muss also nicht für Gebrauchsbeeinträchtigungen haften, die ihre Ursache in der Zeit nach Gefahrübergang haben. Da KI--Systeme in der Erscheinungsform des maschinellen Lernens ihr Ausgabeverhalten auch nach ihrer Bereitstellung für den Nutzer weiter modifizieren können, können Nutzungsbeeinträchtigungen und Fehlfunktionen ihre Ursache nach Gefahrübergang haben. Daraus ergibt sich die Frage, ob solche Fehlfunktionen noch vom Anbieter zu verantworten sind.

c) Beweislast

Im deutschen Recht gilt der allgemeine Grundsatz, dass derjenige Tatsachen vorzutragen und zu beweisen hat, der sich darauf beruft. Danach hat bei Geltendmachung von Gewährleistungsansprüchen der Anspruchsteller das Vorliegen eines Mangels bzw. die Verletzung einer vertraglichen Pflicht zu beweisen.

5.6 Berufshaftung

Für verschiedene Berufe, die auf einer besonderen Vertrauensbeziehung beruhen, finden besondere Berufshaftungspflichten Anwendung. Die Haftungsgrundlagen sollen an den Beispielen Arzthaftung und Anwaltshaftung untersucht werden.

a) Arzthaftung

Für die Arzthaftung (§§ 630a – 630h BGB) ist zu unterscheiden zwischen der Haftung wegen Behandlungsfehlern einerseits und wegen ärztlicher Aufklärungsversäumnisse andererseits. Anspruchsbegründend ist im ersten Fall der Behandlungsfehler, im zweiten Fall die mangelhafte Aufklärung über Risiken der Behandlung. Schäden aus Behandlungsfehlern sind nach § 280 BGB zu ersetzen.[37] Der angestellte Krankenhausarzt ist insoweit Erfüllungsgehilfe des haftenden Krankenhauses. Der Arzt selbst haftet bei Behandlungsfehlern persönlich nach § 823 Abs. 1 BGB.

Im Rahmen seiner Behandlung (sei es zur Diagnose oder zur Therapie) kann ein Arzt Systeme mit Künstlicher Intelligenz einsetzen. Lässt sich durch Verwendung von Systemen mit Künstlicher Intelligenz das Risiko von Fehldiagnosen und -behandlungen für Patienten senken, wird der Arzt sogar zu deren Verwendung aufgrund seiner berufsrechtlichen

[37] vgl. Wagner in MüKo, 8. Auflage, Vor § 630a BGB Rn. 11

Sorgfaltspflichten verpflichtet sein. Die ärztliche Sorgfaltspflicht erfordert den Einsatz von Diagnose- und Behandlungsmethoden nach neuestem Stand von Wissenschaft und Medizin, deren Nutzen nachgewiesen wurde. Schon allein der Verzicht auf deren Einsatz kann zu einer haftungsbegründenden Sorgfaltspflichtverletzung führen.

Im Übrigen gelten die allgemeinen Regelungen der Vertragshaftung nach § 280 BGB. Führt der Einsatz eines KI-Systems zu einer Verschlechterung des Patientenzustands, kommt bei Behandlungsfehlern eine Haftung des Krankenhauses oder des Arztes aus Behandlungsvertrag und bei einem Fehler des Produkts ergänzend eine Haftung des Herstellers aus Produkt- und Produzentenhaftung in Betracht. Eine Haftungslücke ist für Produkte, die Künstliche Intelligenz beinhalten, insoweit nicht erkennbar.

b) Anwaltshaftung

Auf Grund des Mandats (Geschäftsbesorgungsvertrag) haftet ein Rechtsanwalt seinen Mandanten für Schäden, die er aus Verschulden verursacht hat (sog. Anwaltshaftung), z.B. durch eine fehlerhafte oder unterlassene Beratung. Die Haftung greift auch ein, wenn der Anwalt im Rahmen der Leistungserbringung ein System mit Künstlicher Intelligenz einsetzt. Der Anwalt schuldet eine Aufklärung über die aktuelle Rechtslage zu einer bestimmten juristischen Fragestellung. Hierbei kann ihn KI unterstützen, für die Richtigkeit seiner Rechtsauskunft bleibt aber der Anwalt verantwortlich.

5.7 Haftung für Diskriminierung

Viele KI-Systeme sind darauf programmiert, Wertungsentscheidungen zu treffen bzw. zu unterstützen und Differenzierungen vorzunehmen. Solche Differenzierungen sind durch das geltende Recht teilweise vorgegeben. So darf z.B. eine Bank einen Kredit nur vergeben, wenn die Person und die Situation des Kreditnehmers eine Rückzahlung des Kredits erwarten lässt (§ 18a KWG). Vielfach wird aufgrund der Tatsache, dass diese Differenzierungen auf personenbezogenen Parametern beruhen können, befürchtet, dass KI-Systeme Diskriminierung Vorschub leisten. KI-gestützte Differenzierungen und Entscheidungen müssen sich jedoch an den geltenden Diskriminierungsverboten messen lassen, die sich insbesondere im AGG finden.[38] So sind nach § 2 Abs.1 Nr. 8 AGG Benachteiligungen aufgrund von Rasse, ethnischer Herkunft, Geschlecht, Religion, Weltanschauung, Behinderung, Alter oder sexueller Identität auch beim Zugang zu Dienstleistungen unzulässig.[39] Unzulässige Benachteiligungen können durch Unterlassungs- und Schadensersatzansprüche (§ 21 AGG) geltend gemacht werden. Ein

[38] Das AGG (= Allgemeines Gleichbehandlungsgesetz) dient der Umsetzung von vier Antidiskriminierungsrichtlinien: RL 2000/43/EG des Rats vom 29.6.2000 (ABl. EG L 180, 22), RL 2000/78/EG des Rats vom 27.11.2000 (ABl. EG L 303, 16), RL 2002/73/EG des Rats vom 23.9.2002 (ABl. EG L 269, 15) sowie RL 2004/113/EG vom 13.12.2004 (ABl. EG L 373, 37).
[39] Dieses Benachteiligungsverbot gilt nach der Gesetzesbegründung in BT-Drs. 16/1780, S. 32 auch beim Abschluss von Kredit- und Versicherungsverträgen.

Arbeitgeber haftet gemäß § 15 AGG einem abgelehnten Arbeitnehmer für eine Benachteiligung durch Diskriminierung. Außerdem ist bei Abschluss, Durchführung und Beendigung von Verträgen das zivilrechtliche Benachteiligungsverbot des § 19 AGG zu beachten. Wer sich auf eine unzulässige Benachteiligung beruft, muss nach § 22 AGG lediglich Indizien aufzeigen, die für eine solche Benachteiligung sprechen.

Soweit Bewertungsentscheidungen auf personenbezogenen Daten beruhen, müssen diese rechtskonform erhoben bzw. beschafft werden. Insoweit gelten die Vorgaben der DS-GVO zur Datenverarbeitung und zu den Transparenzforderungen nach Art. 22 DS-GVO, wonach bei Einsatz von KI-Systemen z.B. abgefragte Daten und die daraufhin durch ein KI-System ermittelten Muster offen zu legen sind. Die Anwendung dieser Regeln ist aber teilweise noch unklar und teilweise mit dem Wesen von KI schwer in Einklang zu bringen (z.B. Pflichten zur Transparenz und Erklärbarkeit des KI-Entscheidungsprozesses und deren Konkretisierung).

Desgleichen kann die Personalisierung von Preisen mit Diskriminierung und mit Nachteilen für die Verbraucher einhergehen.[40] Verstößt die Abfrage der hierfür erforderlichen personenbezogenen Daten gegen geltendes Recht (insbesondere Datenschutzrecht), kommt insoweit auch eine Haftung für die von Verbrauchern erlittenen Nachteile in Betracht.

5.8 Haftung für Datenschutzverstöße

Werden durch KI-Systeme personenbezogene Daten verarbeitet, ist das Datenschutzrecht zu beachten.[41] Bei Missachtung der datenschutzrechtlichen Vorgaben drohen sowohl dem Verantwortlichen für die Datenverarbeitung als auch dem Auftragsverarbeiter (z.B. IT-Provider) erhebliche Sanktionen. Zu beachten sind nicht nur die ordnungsrechtlichen Sanktionen nach Art. 83 DS-GVO, sondern auch die zivilrechtliche Haftungsnorm des Art. 82 DS-GVO.

Nach Art. 82 DS-GVO kann jeder Verstoß gegen eine Verpflichtung der DS-GVO zu einem deliktischen Schadensersatzanspruch führen. Schadensersatz kann also auch außerhalb eines bestehenden Vertragsverhältnisses jede Person geltend machen, die wegen eines Datenschutzverstoßes einen Vermögensschaden (z.B. Verweigerung eines Vertragsabschlusses wegen unzulässigen Profilings) oder einen immateriellen Schaden (z.B. öffentliche Bloßstellung, soziale Diskriminierung) erleidet.

[40] Vgl. Antwort der Bundesregierung auf eine Kleine Anfrage der FDP-Fraktion vom 29.04.2019 in BT-Drs. 19/9772
[41] Vgl. z.B. zu den technischen und organisatorischen Anforderungen bei Entwicklung und Betrieb von KI das Positionspapier der DSK vom 06.11.2019

Ein Datenschutzverstoß kann in der Nichterfüllung von Transparenz- und Informationspflichten (Art. 12 ff. DS-GVO) liegen.[42] Der Verantwortliche muss danach z.B. die Logik eines Systems zur automatisierten Entscheidungsfindung erläutern können (Art. 13 Abs. 2 f, Art. 14 Abs. 2 g, Art. 15 Abs. 1 h DS-GVO). Umfang und Inhalt dieser Aufklärungspflicht sind zwar noch nicht konkretisiert, allerdings dürfte es für Verfahren des Machine Learning und des Deep Learning kaum immer gelingen, eine zutreffende Erklärung der technologischen Zusammenhänge zu geben, ohne die Betroffenen zu überfordern. Dies könnte dazu führen, dass diese Verfahren zur Bearbeitung personenbezogener Daten wegen des damit verbundenen rechtlichen Risikos ausscheiden. Hinzukommt, dass der Verantwortliche die Einhaltung der DS-GVO jederzeit nachweisen muss (Art. 5 Abs. 2 DS-GVO). Dies führt im Ergebnis zu einer Beweislastumkehr zugunsten des Geschädigten und erhöht gleichzeitig das Haftungsrisiko beim Einsatz von KI-Systemen.

5.9 Zusammenfassung

Das geltende Recht verfügt über ein dichtes Netz an verschiedenen Vorschriften für Herstellung und Einsatz technischer Produkte. Dabei stellt das Recht zum einen Konformitätsanforderungen insbesondere zur Produktsicherheit auf, denen ein Produkt entsprechen muss, um überhaupt zum Verkauf in der EU zugelassen zu werden. Zum anderen kommen für Schäden, die trotz dieser Konformitätsvorgaben beim Gebrauch technischer Produkte eintreten, verschiedene Haftungsvorschriften zur Anwendung. Konformitätsvorgaben und Haftungsvorschriften wurden zwar nicht spezifisch für KI-Anwendungen formuliert, sind aber ohne Abstriche von Herstellern, Betreibern und Nutzern von KI-Anwendungen zu beachten.

Der Hersteller haftet seinen Vertragspartnern aus Gewährleistung für die Mangelfreiheit seiner KI-Produkte und geschädigten Nutzern einer KI-Anwendung nach Produkt- und Produzentenhaftung. Zusätzliche Anforderungen an die Hersteller formuliert das Produktsicherheitsrecht, das zur Prävention von Rechtsgutsverletzungen dient.

Der Betreiber eines KI-Systems haftet dafür, dass es beim Einsatz des Systems nicht zu Datenschutzverstößen, Diskriminierungen, Personen- oder Sachschäden kommt. Setzt der Betreiber das KI-System für die Erfüllung seiner vertraglichen Verpflichtungen ein, kommen vertragliche Haftungsansprüche hinzu, die durch eine Berufshaftung ggf. noch verschärft werden. Der Nutzer haftet nach Deliktsrecht ebenfalls für eigenes Verschulden. Darüber hinaus ist bei ihm Mitverschulden bei der Bemessung eines möglichen Schadensersatzes zu berücksichtigen.

[42] Zu den Anforderungen des Datenschutzes bei Einsatz von KI-Systemen hat Bitkom bereits eine Untersuchung veröffentlicht, vgl. den Leitfaden „Machine Learning und die Transparenzanforderungen der DS-GVO“

Das geltende Recht richtet seine Haftungsverpflichtungen nicht zuletzt an der Gefahr für wichtige Rechtsgüter aus. Je größer die potenzielle Gefahr und je höherrangig das gefährdete Rechtsgut einzustufen sind, desto schärfer ist die Haftung und desto mehr Anspruchsgrundlagen und Haftungsverantwortliche stehen dem Geschädigten potenziell zur Verfügung. Ist z.B. die zeitgerechte Durchführung eines Vertrages beeinträchtigt, so kann die dadurch geschädigte Partei allein gegen den Vertragspartner Ansprüche aus Verzug geltend machen. Wird dagegen bei einem Autounfall der Beifahrer verletzt, kommen Haftungsansprüche gegen den Fahrer aus Vertrags- und Verschuldenshaftung, gegen den Unfallgegner aus Verschuldenshaftung, gegen den Fahrzeughersteller aus Produkthaftung, gegen die Halter der unfallbeteiligten Fahrzeuge bzw. deren Versicherer aus Gefährdungshaftung in Betracht. Deutlich wird diese abgestufte Haftungsverantwortung auch bei der Produzentenhaftung. Danach sind die Sorgfaltsanforderungen an den Hersteller umso höher je gefährlicher das Produkt und je wahrscheinlicher ein Schadenseintritt sind.

6 Vorgeschlagene Erweiterungen des geltenden Rechts

In der gegenwärtigen Diskussion wird vielfach in Zweifel gezogen, dass das geltende Haftungsrecht die Risiken von KI-Systemen noch angemessen adressiert und vollständig abdeckt. Daher werden verschiedene Konzepte zur Neu- und Umverteilung von Haftungsrisiken zwischen möglichen Beteiligten (Hersteller, Vermarkter, Betreiber, Nutzer) und zur Einführung neuer Haftungsgrundlagen vorgeschlagen. Einige der Konzepte werden nachfolgend betrachtet.

6.1 Anpassung der Produkthaftung

Das Produkthaftungsrecht in der EU ist durch die Produkthaftungsrichtlinie 85/374/EWG aus dem Jahr 1985 harmonisiert. Änderungen in diesem Bereich müssen daher zunächst auf EU-Ebene erfolgen. Die EU-Kommission hat denn auch bereits eine Evaluierung der geltenden Produkthaftungsrichtlinie vorgenommen[43] und eine Expertengruppe eingesetzt, die Vorschläge für mögliche Änderungen vor dem Hintergrund der KI-Entwicklung erarbeiten soll.[44] Im Folgenden setzt sich Bitkom mit einigen Überlegungen zur Änderung und Erweiterung der europäischen Produkthaftungsrichtlinie auseinander.

a) Erweiterung der Produktdefinition

Die Produkthaftung im geltenden Recht richtet sich im Wesentlichen am Warenhandel aus und definiert demzufolge ein Produkt als bewegliche Sache. Ob Software und KI-Anwendungen als Produkte in diesem Sinn anzusehen sind, ist umstritten. Zur

[43] Vgl. Bericht der Kommission über die Anwendung der Produkthaftungsrichtlinie aus dem Mai 2018
[44] Die Expertengruppe hat auch bereits im Mai 2019 einen ersten Bericht vorgelegt.

Klarstellung, aber auch zur Erweiterung des Produktbegriffs wird vorgeschlagen, KI-Anwendungen sowie digitale Inhalte und Dienstleistungen explizit in den Anwendungsbereich der Produkthaftungsrichtlinie einzubeziehen.[45]

Bitkom-Bewertung

Gegen die Ausweitung des etablierten Produktbegriffes auf KI-Systeme spricht, dass es bisher nicht zu Haftungslücken bei Einsatz und Nutzung solcher Systeme gekommen ist und sich das geltende Produkthaftungsrecht bewährt hat. Durch die Rechtsprechung ist bereits anerkannt, dass ein Hersteller nach den geltenden Grundsätzen der Produkthaftung verantwortlich ist, wenn der Schaden durch die Fehlfunktion einer in das Produkt integrierten Steuerungssoftware (embedded software) verursacht wurde.[46]

Der Erweiterung des Produktbegriffs auf eigenständig in Verkehr gebrachte Software („Stand-Alone-Software“) ist entgegen zu halten, dass die Verletzung von Rechtsgütern, die durch die Produkthaftung geschützt sind, durch ein rein immaterielles Gut ohne Beteiligung eines physischen Gegenstands kaum denkbar ist. Software kann sich nur als Komponente eines physischen Gegenstands und in Kombination mit diesem auf die physische Umwelt auswirken. Diesem Umstand trägt auch das Patentrecht dadurch Rechnung, dass es reinen Softwareprodukten die Patentfähigkeit verwehrt und eine technische Materialisierung von Software für einen Patentschutz fordert.

Gegen die Qualifizierung von Stand-Alone-Software als Produkt spricht des Weiteren, dass wesentliche Gefahren, z.B. Sicherheitslücken, nicht nur vom originären Hersteller der Stand-Alone-Software abhängen, sondern wesentlich auch vom Nutzer/Betreiber. Insbesondere kann der Hersteller der Software kaum vorhersehen oder kontrollieren, mit welchen anderen Softwarekomponenten der Betreiber/Nutzer die Software verbindet. Außerdem würde die Anwendung der Produkthaftung auf eigenständige Software den Urhebern von Open Source Software untragbare Haftungsrisiken aufbürden, die dazu führen würden, dass diese Art von Software nicht mehr angeboten wird.

Schließlich ist die Forderung nach Einbeziehung von Dienstleistungen in den Produktbegriff und damit in den Anwendungsbereich des Produkthaftungsrechts bisher nicht ausreichend begründet und geht zu weit. Insbesondere ist nicht ersichtlich, wie es außerhalb des Vertragsverhältnisses über eine Dienstleistung zu Schädigungen unbeteiligter Dritter kommen kann. Wird der Auftraggeber einer Dienstleistung geschädigt, kann er dafür bereits nach geltendem Vertragsrecht ausreichend Kompensation erlangen.

[45] Eine entsprechende Forderung findet sich z.B. im Gutachten der Datenethikkommission, S. 222 und Bericht der EU-Expertengruppe (vgl. Fn. 44), S. 42 f.
[46] So hat der BGH mit Urteil vom 16.06.2009, Az. VI ZR 107/08 eine Produkthaftung des Fahrzeugherstellers für Gesundheitsschäden bejaht, die durch einen Fehler in der Steuerungssoftware eines Airbags verursacht wurden.

Unabhängig von der Bestimmung des Produktbegriffs könnte überlegt werden, ob weitere Parteien, die am Betrieb oder an der Bereitstellung eines schadensstiftenden Produkts beteiligt waren, in eine verursachungsgerechte Aufteilung von Verantwortlichkeiten nach § 5 ProdHaftG bzw. Art. 5 der ProdHaft-RL 85/374/EWG einzubeziehen sind. Ansonsten hätte ein Hersteller, der den Schaden gegenüber den Geschädigten reguliert, u.U. nur vertragliche Rückgriffsansprüche gegen andere Beteiligte, die ebenfalls eine Schadensursache gesetzt haben.

b) Ausweitung der zeitlichen Herstellerverantwortung

Nach geltendem Produkthaftungsrecht ist der Hersteller verantwortlich für Fehler, die ihre Ursache vor Inverkehrbringen des Produkts haben. Es wird vorgeschlagen, die Haftung auszuweiten auf Fehler, die erst nach Inverkehrbringen des Produkts im Laufe der Nutzung entstehen (z.B. durch Cyberangriffe oder Updates).[47] Insoweit würde den Hersteller zukünftig eine aus Produkthaftung abgeleitete Pflicht treffen, ein KI-System nach Inverkehrbringen auf einem aktuellen Stand zu halten (Update-Pflicht) und diese Updates allen Nutzern – unabhängig von bestehenden Pflegeverträgen – zugänglich zu machen.

Bitkom-Bewertung

Eine Update-Pflicht für Anbieter digitaler Produkte wurde durch Art 7 Abs. 3 der Richtlinie (EU) 2019/771 bereits eingeführt, aber zu Recht auf Vertragsverhältnisse beschränkt. Denn nur soweit ein Vertragsverhältnis besteht und der Software-Hersteller über Daten zu den Softwarekunden verfügt, ist es ihm möglich, Updates in notwendigem Umfang auf Geräte der Kunden aufzuspielen. Außerdem dienen Updates von Software dazu, die Funktionalität des Programms aufrecht zu erhalten und so dem Nutzer den Nutzwert zu erhalten. Dieses Äquivalenzinteresse ist durch das Vertragsrecht geschützt. Die Produkthaftung dient dagegen dazu, nicht an einem Vertragsverhältnis beteiligte Dritte in der Integrität ihrer Rechtsgüter zu schützen. Hierfür sind aber nicht zwingend Software-Updates erforderlich, sondern eine Gefahr für Rechtsgüter Dritter könnte z.B. durch Warnungen, Rückruf oder – insbesondere bei ML-Systemen – durch Abschalten der Lernfunktion gewährleistet werden.

Nach der insoweit nicht eindeutigen Rechtsprechung des BGH kommt im Übrigen bereits auf der Grundlage des geltenden Rechts eine Update-Pflicht in Betracht, allerdings nicht auf Grundlage der Produkthaftung, sondern nur im Rahmen der deliktischen Produzentenhaftung bei zusätzlicher Verletzung einer Verkehrssicherungspflicht. So hat der BGH entschieden, dass ein Hersteller zur Nachbesserung seines Produkts verpflichtet sein kann, wenn nicht davon auszugehen ist, dass andere vom Hersteller im Rahmen

[47] Vgl. Bericht der EU-Kommission COM(2020) 64 final, S. 18 und Empfehlungen der Datenethikkommission, S. 221 f.

seiner Sorgfaltspflichten zu ergreifende Maßnahmen zu einer effektiven Beseitigung des Gefahrenpotenzials führen.[48]

Darüber hinaus ist zu bedenken, dass die Pflege von Software erheblichen Aufwand des Herstellers erfordert. Dieser Aufwand ist umso höher je komplexer und spezifischer die zu pflegende Software ist. Daher werden Softwarepflegeleistungen im Geschäftsverkehr zwischen Unternehmen regelmäßig nur auf Grundlage besonderer vertraglicher Regelungen und gegen Entgelt erbracht. Im Rahmen der Softwarepflege können die spezifische Einsatzumgebung eines Softwareprodukts, das Zusammenspiel mit anderen Hard- und Softwarekomponenten und der aktuelle technische Fortschritt sowie aktuelle Bedrohungen der Cybersicherheit angemessen und kundenindividuell berücksichtigt werden. Mit einer ähnlichen Updateverpflichtung aus Produkthaftung würde den Herstellern nicht nur ein wichtiges Geschäftsfeld wegbrechen, sondern eine solche Pflicht wäre für sie in der notwendigen Kundenindividualisierung schlicht nicht erfüllbar. Eine noch darüber hinausgehende gesetzliche Garantiepflicht für Fehlerfreiheit der Updates ließe den Aufwand weiter steigen und wäre für Hersteller nicht mehr kalkulierbar.

Letztlich ist dem Hersteller von Software und insbesondere von KI-Systemen eine verschuldensunabhängige Haftung nur insoweit zuzumuten als er auf die Geschehensabläufe direkten und unmittelbaren Einfluss hat. Eine solche Beherrschbarkeit des Geschehensablaufes ist aber nur auf der Grundlage eines Vertrages oder bis zum Inverkehrbringen eines Produkts gegeben.

c) Ausweitung des Anwendungsbereichs auf weitere Rechtsgüter

Aktuell ist nach den Regeln der Produkthaftung Schadensersatz nur zu leisten, wenn ein Mensch getötet, verletzt, in seiner Gesundheit beeinträchtigt oder wenn eine privat genutzte Sache beschädigt wird. Teilweise wird gefordert, die Ersatzpflicht auf immaterielle Schäden wie Datenverlust oder Datenschutzverletzungen bzw. Verletzung der informationellen Selbstbestimmung auszuweiten.[49]

Bitkom-Bewertung

Die Produkthaftung ist an vergleichsweise geringe Voraussetzungen geknüpft, um primär Verbraucher unabhängig von bestehenden Verträgen angemessen zu schützen. Zum Ausgleich dafür kann Schadensersatz nur bei Verletzungen besonders wichtiger Rechtsgüter gefordert werden. Diese Balance darf nicht durch eine schrankenlose Ausweitung der geschützten Rechtsgüter beseitigt werden.

[48] BGH, Urteil vom 16.12.2008, Az. VI ZR 170/07 „Pflegebettenentscheidung“
[49] Vgl. Verbraucherzentrale Bundesverband: „Smarte Haftung für smarte Produkte“, Oktober 2019 und Empfehlungen der Datenethikkommission, S. 222

Eingriffe in den Datenschutz und Verletzungen von Persönlichkeitsrechten mithilfe digitaler Werkzeuge sind bereits außerhalb der Produkthaftung im Datenschutzrecht, im Antidiskriminierungsrecht und in der Deliktshaftung sanktioniert. Zu Schutzlücken kommt es insoweit also nicht. Es wäre nicht sinnvoll, die insoweit schon scharfen Sanktionen für Unternehmen zu verdoppeln. Ein notwendiger risikoorientierter Haftungsansatz bedeutet auch, dass nicht nach denselben Voraussetzungen für jegliche Schäden in mehrfacher Weise zu haften ist. Es muss also dabei bleiben, dass Datenverluste nur zu einer Produkthaftung führen, wenn in dem Datenverlust gleichzeitig eine Sachbeschädigung des Datenträgers zu sehen ist.

d) Änderung der Beweislastverteilung

Die Zuordnung eines schadenstiftenden Fehlers zum Verursacher, die Beweispflicht des Geschädigten und dementsprechend die Durchsetzung von Schadensersatzansprüchen im Produkthaftungsrecht kann zu Schwierigkeiten führen, insbesondere dann, wenn komplexe KI-Systeme aus zahlreichen Komponenten bestehen und am Betrieb des Systems mehrere Personen beteiligt sind. Daher hat die EU-Kommission Beweiserleichterungen zugunsten des Geschädigten bzw. sogar eine Beweislastumkehr zulasten des Herstellers ins Gespräch gebracht. Die Umkehr der Beweislast für das Vorliegen eines Produktfehlers und dessen Ursache für den eingetretenen Schaden soll eintreten in Fällen, in denen der Hersteller einschlägige Sicherheitsvorschriften nicht eingehalten hat.[50]

Bitkom-Bewertung

Zunächst ist darauf hinzuweisen, dass Schwierigkeiten beim Nachweis begründeter Schadensersatzansprüche nicht nur bei KI-Systemen, sondern bei jeder Art von Produkten auftreten können. Möglicherweise bestehende Nachweisschwierigkeiten vergrößern sich nicht dadurch, dass KI-Systeme in einem Produkt integriert sind. Insoweit ist für die Beweislastverteilung keine Ausnahme- oder Sonderregelung für KI erforderlich.

Auch kommen dem Geschädigten bereits nach der Rechtsprechung zum geltenden Recht Beweiserleichterungen wie Anscheinsbeweis und Fehlervermutung zugute. Der Geschädigte muss nicht die konkrete Fehlerart benennen oder technische Hintergründe des Fehlers erläutern. So ist regelmäßig schon dann ein Produktfehler anzunehmen, wenn ein Produkt trotz sachgemäßen Gebrauchs zu einem Personen- oder Sachschaden geführt hat. Entspricht das Produkt nicht den einschlägigen gesetzlichen Sicherheitsvorschriften oder technischen Normen, so begründet dies ebenfalls eine widerlegbare Vermutung der Fehlerhaftigkeit. Die vielfach beklagte Undurchschaubarkeit von Prozessen des

[50] Bericht der Europäischen Kommission COM(2020) 64 final, Seite 17, Report from the Expert Group on Liability and New Technologies – New Technologies Formation, S. 48 f.

maschinellen Lernens wirkt sich also vor allem zum Nachteil des Herstellers aus, da dieser zu beweisen hat, dass für einen eingetretenen Schaden kein Fehler des von ihm in Verkehr gebrachten Systems verantwortlich war. Deshalb setzen sich die Hersteller seit langem und in eigenem Interesse dafür ein, Methoden und Prüfverfahren zur Erhöhung der Transparenz zu entwickeln.

6.2 Anpassung der Produzentenhaftung

Ähnlich wie bei der Produkthaftung sieht die EU-Kommission auch bei der verschuldensabhängigen Produzentenhaftung nach § 823 BGB Schwierigkeiten des Geschädigten, das Vorliegen der Anspruchsvoraussetzungen zu beweisen und entsprechend seine Haftungsansprüche gegen den Hersteller durchzusetzen. So sind Geschädigte mit den Sorgfaltsanforderungen an KI-Hersteller regelmäßig nicht ausreichend vertraut und haben unzureichenden Einblick in den Herstellungsprozess von KI-Produkten, um Sorgfaltspflichtverstöße des Herstellers nachweisen zu können. Auch insoweit kommt daher aus Sicht der EU-Kommission eine Änderung der Beweislastverteilung zulasten der Hersteller in Betracht.[51]

Bitkom-Bewertung

Der Nachweis eines schuldhaften Verhaltens obliegt im Regelfall dem Geschädigten. Abweichend davon hat der BGH jedoch im Bereich der Produzentenhaftung den Grundsatz entwickelt, dass der Hersteller im Wege der Beweislastumkehr darzulegen und zu beweisen hat, dass ihn in Bezug auf die Fehlerhaftigkeit des Produktes, die zu dem Schaden beim Verbraucher geführt hat, kein Verschulden trifft.[52]

Auch können faktisch vorhandene technische Möglichkeiten die Sorgfaltspflichten bei der Produktion und bei der Produktionsüberwachung erhöhen. Neue Technologien ermöglichen auch neue Formen der Produktbeobachtung. Bei der sogenannten „integrierten Produktbeobachtung“ wird durch eine kontinuierlichen Selbstüberwachung und einer Verbindung zum Internet erreicht, dass kritische Anomalien eines Systems direkt und sofort an den Hersteller gemeldet werden können, darüber hinaus sind sogar Gegenmaßnahmen über den gleichen Weg denkbar. Schon heute kann also aus den Verkehrssicherungspflichten eine „integrierte Produktbeobachtung“ abzuleiten sein. Dabei kommt es jedoch auf den Einzelfall an, insbesondere auf das Gefahrenpotenzial des Systems, die berechtigten Erwartungen des Verkehrs und den Aufwand für den Hersteller.

[51] KI-Weißbuch der EU-Kommission, S. 18 f.; Report from the Expert Group on Liability and New Technologies – New Technologies Formation, S. 52 f.
[52] BGH, Urteil vom 19.11.1991, Az. VI ZR 171/91

6.3 Anpassung des Produktsicherheitsrechts

Die einschlägigen Vorschriften zur Produktsicherheit, die insbesondere in den sog. CE-Richtlinien[53] niedergelegt sind, haben den Zweck, Gesundheits- und Sicherheitsrisiken technischer Produkte zu eliminieren, bevor sich diese Risiken durch Inverkehrbringen der jeweiligen Produkte realisieren können. Zur Implementierung der gesetzlichen Anforderungen in Produkten, Systemen und Anlagen haben sich die in europäischen technischen Normen (z.B. CEN; CENELEC) im Einklang mit internationalen Normen (ISO, IEC) festgelegten Prozesse zur Risikobeurteilung und die Prinzipien der Sicherheitsintegration zur Risikominderung als wirksam erwiesen.

a) Konkretisierung des Anwendungsbereichs

Aus Sicht der EU-Kommission kann es zu Rechtsunsicherheiten kommen bei der Frage, ob bestimmte Sicherheitsanforderungen auch für Software oder für Dienstleistungen, die mit Hilfe von KI-Systemen erbracht werden, gelten. Relevant wird diese Frage z.B., wenn die KI-Software nicht vom Hersteller eines Produkts, sondern von einem Dritten zugeliefert wird.[54]

Bitkom-Bewertung

Die Anforderungen des Produktsicherheitsrechts gelten unabhängig von spezifischen Technologien und decken neue Technologien grundsätzlich ab, auch wenn ihre Geltung nur teilweise explizit auch für Software festgelegt ist (z.B. in der Verordnung für Medizinprodukte (EU) 2017/745).

b) Robustheit und Genauigkeit

Das Weißbuch der EU-Kommission für Künstliche Intelligenz enthält Überlegungen, den Herstellern Vorgaben für die Entwicklung von KI-Anwendungen mit hohem Risiko zu machen, damit diese Systeme technisch solide und präzise funktionieren.[55] Mögliche Risiken des Systems sollen vorab bewertet werden. Nicht näher spezifizierte Maßnahmen sollen Schäden möglichst gering halten, eine genaue, robuste und in den Ergebnissen reproduzierbare Arbeitsweise der Systeme gewährleisten sowie die Widerstandsfähigkeit gegen Manipulationsversuche begründen. Aus den ethischen Grundsätzen für KI lässt sich die Forderung ableiten, dass bei risikogeneigtem KI-Einsatz zur Gefahrbegrenzung eine „Abschaltvorrichtung“ bzw. ein „Not-Aus“ vorzusehen ist.

[53] Eine Übersicht über geltende CE-Richtlinien findet sich z.B. hier.
[54] KI-Weißbuch der EU-Kommission, S. 16 f.
[55] KI-Weißbuch der EU-Kommission , S. 24 f.

Bitkom-Bewertung

Technische Systeme müssen in Gefahrensituationen durch einen „Not-Halt“ in einen gefahrfreien Zustand versetzt werden können (vgl. z.B. Ziff. 1.2.4 des Anhangs I zur Maschinenrichtlinie 2006/42/EG). Auch müssen Maschinen so konstruiert sein, dass Fehler in den Steuerungsaggregaten nicht zu Gefahrensituationen führen (Ziff. 1.2.1 des Anhangs I zur Maschinenrichtlinie 2006/42/EG). Anforderungen der Robustheit sind also bereits Bestandteil des geltenden Rechts. Insoweit müssen keine neuen Anforderungen an KI-Systeme formuliert werden, sondern es muss lediglich sichergestellt werden, dass die bestehenden Anforderungen auf KI-Systeme Anwendung finden.

c) Konformitätsbewertung

Die EU-Kommission ist der Auffassung, dass KI-Anwendungen mit hohem Risiko eine objektive, vor Marktzulassung vorzunehmende Konformitätsbewertung durchlaufen sollten. In Prüf- oder Zertifizierungsverfahren würde dann kontrolliert, ob eine KI-Anwendung mit hohem Risiko Vorgaben des geltenden Rechts sowie obligatorische Qualitätsmerkmale für die Anwendung selbst (z.B. Robustheit und Genauigkeit) und für den Entwicklungsprozess erfüllt. Für nicht sonderlich riskante KI-Anwendungen schlägt die Kommission ein freiwilliges Gütesiegel vor, mit dem die Einhaltung gewisser Anforderungen nachgewiesen werden kann.[56]

Bitkom-Bewertung

Hochrisikoanwendungen unterliegen bereits nach dem geltenden Recht erhöhten Sicherheitsanforderungen, die vor dem Inverkehrbringen des Produkts abgeprüft werden. Die geltenden Produktsicherheitsvorgaben sind am Einsatzzweck des jeweiligen Produkts ausgerichtet und auf diesen abgestimmt. Sie gelten auch, wenn in den jeweiligen Produkten KI zum Einsatz kommt.

Die aktuell bestehenden Markteinführungsregularien sind bereits sehr weitreichend und aufwändig. Zusätzliche Konformitätsbewertungen können die Markteinführung innovativer KI-Anwendungen erheblich verzögern oder gar verhindern. Um den Aufwand einer umfangreichen Konformitätsbewertung zu vermeiden, könnten Unternehmen dazu tendieren, im Zweifel auf KI-Anwendungen zu verzichten und stattdessen "traditionelle" Systeme zu nutzen. Dies würde dazu führen, dass KI weniger Verbreitung findet, auch wenn von weniger intelligenten Algorithmen kaum weniger Gefahren ausgehen. Zudem könnte sich der mit der Konformitätsbewertung verbundene Mehraufwand insbesondere für kleine und mittlere Unternehmen negativ auswirken. Fixe Regulierungskosten fallen bei ihnen überproportional ins Gewicht.

[56] KI-Weißbuch der EU-Kommission, S. Seite 27 f.

Zu bedenken ist darüber hinaus, dass sich heuristisch programmierte Algorithmen mehrmals am Tag ändern können, und lernende Systeme ihre Muster mit der Aufnahme weiterer Trainingsdaten und Echtdaten kontinuierlich anpassen. Ein Prüfprozess könnte daher allenfalls eine Bestandaufnahme für einen bestimmten Zeitpunkt liefern, deren Wert fraglich wäre. Nimmt ein KI-Prüfprozess einen längeren Zeitraum in Anspruch oder ist er mit größerem Aufwand verbunden, kann er angesichts dieser lediglich bedingten Präventionswirkung allenfalls für Hochrisiko-Anwendungen verhältnismäßig sein. Insoweit sollte sich die Prüfung darauf beschränken, ob die Ergebnisse des Lernprozesses von kontinuierlich lernenden Systemen sowie deren Datenerfassung noch den gesetzlichen und technischen Sicherheitsanforderungen genügen.

6.4 Einführung einer Algorithmenkontrolle

Eine besondere Form der Konformitätsbewertung stellt die Algorithmenkontrolle dar. Der Sachverständigenrat für Verbraucherfragen hat in seinem Gutachten 2016 gefordert, dass Algorithmen nach standardisierten Offenlegungspflichten gegenüber einer staatlichen Digitalagentur offen gelegt werden sollen, sodass zumindest stichprobenartig überprüft werden kann, ob die Algorithmen den Vorgaben des Verbraucherrechts, des Datenschutzrechts, des Antidiskriminierungsrechts und der digitalen Sicherheit entsprechen.[57] Bei dieser Art von Kontrolle sollen also nicht wie bei der Produktsicherheit mögliche Gesundheitsrisiken im Fokus der Betrachtung stehen, sondern die Konformität mit Vorschriften zum Persönlichkeits- und Vertraulichkeitsschutz.

Bitkom-Bewertung

Marktzulassungsprüfungen und Funktionsüberprüfungen kennt das Recht bereits für Produkte mit besonderem Gefährdungspotenzial (z.B. Kraftfahrzeuge). Entsprechende Kontrollen sind auch durchzuführen, wenn in das jeweilige Produkt KI-Funktionalitäten integriert sind. Marktzulassungsprüfungen müssen aber auf Produkte mit hohem Gefährdungspotenzial beschränkt bleiben. Eine solche Beschränkung wäre schon angesichts der schieren Anzahl der zu prüfenden KI-Systeme und Algorithmen nicht zu vermeiden. Für Datenschutzzwecke finden sich inzwischen auch ausreichende Offenlegungspflichten in Art. 13, 14 und 15 der DS-GVO.

Die Einführung einer Kontrolle zur Verhinderung von Diskriminierungen ist auch deswegen kritisch zu sehen, weil algorithmische Systeme entwickelt werden, um Differenzierungen zu ermöglichen und Unterscheidungen vorzunehmen. Für die Frage, wann die Schwelle von einer rechtlich erforderlichen Differenzierung zu einer rechtlich missbilligten Diskriminierung überschritten wird, fehlen allgemein nutzbare Kriterien. So

[57] Vgl. Sachverständigenrat für Verbraucherfragen: Gutachten „Verbraucherrecht 2.0“ aus dem Dezember 2016, S. 67

muss z.B. die Begünstigung von Frauen bei der Prämienberechnung von Kfz-Haftpflichtversicherungen keine Geschlechterdiskriminierung sein, sondern kann durch ein statistisch ermitteltes vorsichtigeres Fahrverhalten von Frauen begründet sein.

Schließlich sind Algorithmen für Computersysteme in der Regel durch Urheberrechte oder als Geschäftsgeheimnis geschützt. Diese Rechtspositionen sind bei Entscheidungen über eine Offenlegung zu berücksichtigen.

6.5 Einführung einer Gefährdungshaftung für KI

Um Geschädigten die Durchsetzung ihrer Ansprüche auf Ersatz von Schäden durch KI zu erleichtern, wird eine Ausweitung der Gefährdungshaftung für KI-Systeme mit besonderem Risikoprofil ins Gespräch gebracht.[58] Auf ein Verschulden oder eine Fehlerhaftigkeit des Systems kommt es für eine solche Haftung nicht an. Für Hersteller bzw. Betreiber von KI-Systemen entfällt insoweit die Möglichkeit, sich dadurch von der Haftung zu befreien, dass sie die Beachtung sämtlicher Sorgfaltspflichten und die Einhaltung einschlägiger technischer Standards nachweisen. Die Einführung einer Gefährdungshaftung für KI-Systeme wird damit begründet, dass bei manchen KI-Systemen insbesondere im Bereich des Deep Learning nicht mehr im Einzelnen nachvollziehbar sei, wie das System auf sein Ergebnis kommt. Auch sei das Lernverhalten solcher Systeme vom Hersteller nicht mehr im Einzelnen kontrollierbar und die Ergebnisse vom Hersteller nicht mehr beherrschbar.

Die Europäische Kommission erwägt zudem, für Produkte, die der neu einzuführenden Gefährdungshaftung unterliegen, eine Versicherungspflicht einzuführen, um den Geschädigten ähnlich wie bei Kraftfahrzeugen eine ausreichende Haftungsmasse für seine Ersatzansprüche zur Verfügung zu stellen.

Bitkom-Bewertung

Eine allgemeine Betreiber- oder Gefährdungshaftung (z.B. in Form einer Halterhaftung) findet sich bereits im geltenden Recht für Gefahrenquellen mit besonders hohen Risiken für Leben und Gesundheit von Menschen. KI-Systeme können jedoch nicht allgemein als besonders risikogeneigt oder Gefahr erhöhend angesehen werden. Im Gegenteil: KI-Systeme werden genutzt, um menschliche Unzulänglichkeiten (z.B. die begrenzte Fähigkeit, viele Daten in einer sehr kurzen Zeit zu verarbeiten) auszugleichen und damit Risiken und Schäden zu reduzieren. Eine allgemeine Betreiberhaftung für sämtliche KI-

[58] Überlegungen hierzu finden sich z.B. im Bericht der EU-Kommission COM(2020) 64 final, S. 19 f., in den Empfehlungen der Datenethikkommission, S. 220, aber auch bereits in der Entschließung (2015/2103(INL) des Europäischen Parlaments vom 16.02.2017, da selbstlernende Systeme und autonome Roboter nicht mehr als Werkzeuge eines menschlichen Akteurs (Hersteller, Eigentümer, Betreiber oder Nutzer) angesehen werden könnten.

Produkte ginge daher zu weit. Vielmehr kommt es für die Beurteilung des Gefährdungspotenzials auf den Einsatz eines Systems im Einzelfall und auf den Risikobeitrag des Systems an. Das Haftungsrisiko für Hersteller oder Betreiber eines Systems sollte sich nicht nur deswegen erhöhen, weil das System mit Künstlicher Intelligenz arbeitet.

Wenn eine Gefährdungshaftung für Hochrisiko-Anwendungen eingeführt werden sollte, muss der Geltungsbereich einer solchen Haftung eng definiert und die Geltung für einzelne Anwendungen gut begründet sein. In der Folge ist dann zu gewährleisten, dass der Geltungsbereich nicht beliebig ausgedehnt wird.

Hohe Versicherungskosten für eine Pflichtversicherung könnten verschiedene Betreiber von der Nutzung von Innovationen abhalten.

6.6 Trainingsdatenregulierung

In ihrem Weißbuch zu Künstlicher Intelligenz schlägt die EU-Kommission vor, „geeignete Maßnahmen“ zu ergreifen, um sicher zu stellen, dass Trainingsdaten für KI-Anwendungen mit hohem Risiko im Einklang mit EU-Bürgerrechten und im Einklang mit einschlägigen Vorgaben zur Datensicherheit erhoben, zusammengestellt und verarbeitet werden.[59] Die Anforderungen und Auflagen sollen dazu dienen, die EU-Sicherheitsvorschriften, Diskriminierungsverbote und den Schutz der Privatsphäre für die Trainingsdaten durchzusetzen. Um die Einhaltung dieser Anforderungen im Nachgang kontrollieren zu können, sollen zu den verwendeten Datensätzen Aufzeichnungen erstellt und in bestimmten Fällen die Datensätze selbst aufbewahrt werden.

Bitkom-Bewertung

Auch wenn die EU-Kommission weder die „geeigneten Maßnahmen“ für Trainingsdaten noch etwaige Konsequenzen bei deren Nichtbeachtung konkretisiert hat, so ist doch absehbar, dass die Nichtbeachtung dieser Anforderungen sanktioniert werden dürfte und damit zu einer Haftungsrelevanz führt. Aufgrund der nur sehr vagen Skizzierung der möglichen Maßnahmen lässt sich nicht abschließend beurteilen, ob sie über die Anforderungen des geltenden Rechts überhaupt hinausgehen, sinnvoll und zielführend sind oder nicht vielmehr Herstellern und Betreibern von KI-Systemen unnötige zusätzliche Lasten aufbürden.

Diskriminierende Entscheidungen, Datenschutzverstöße und die damit einhergehende Verletzung individueller Rechte führen bereits aktuell zu einer Haftung (vgl. oben), auch wenn bei der Entscheidungsfindung ein KI-System Anwendung findet. Angesichts der

[59] Vgl. KI-Weißbuch der EU-Kommission , S. 22 f.

insoweit bereits scharfen Sanktionen ist nicht ersichtlich, wie weitere Vorschriften in diesem Bereich den Rechtsschutz verbessern können. Im Gegenteil: eine Pflicht zur Speicherung von Datensätzen für Zwecke der Nachkontrolle steht in Konflikt mit dem Datenschutzrecht, das dem Ansatz der Datensparsamkeit folgt und weitgehende Löschpflichten vorsieht. Außerdem hat die DS-GVO zum Schutz der Privatsphäre im Datenbereich die nationalen Datenschutzbehörden berufen. Von deren Vorgaben abweichende Anforderungen könnten daher keinen Bestand haben.

Letztlich kann aus den verwendeten Trainingsdaten auch dann nicht zwingend auf eine Diskriminierung durch ein ML-System geschlossen werden, wenn in den Trainingsdaten diskriminierende Tendenzen („biases“) angelegt sind. Denn mit ausreichender Expertise und Sorgfalt können Unzulänglichkeiten in den Lerndatensätzen im Trainingsprozess ausgemerzt werden, sodass das System dennoch Ergebnisse innerhalb eines akzeptablen Spektrums liefert.

6.7 Verschärfung von Dokumentationsanforderungen

Aus Sicht der EU-Kommission führen Komplexität und fehlende Transparenz vieler KI-Systeme zu Schwierigkeiten bei der Rekonstruktion und Überprüfung von deren Lern- und Entscheidungsmechanismen. Dies wiederum könne die wirksame Rechtsdurchsetzung erschweren, nicht zuletzt die Durchsetzung von Haftungsansprüchen. Daher sollen Hersteller über die Programmierung und das Training von Algorithmen für KI-Systeme mit hohem Risiko sowie über die verwendeten Trainingsdaten Aufzeichnungen führen und in besonderen Fällen die verwendeten Trainingsdaten selbst für Kontrollzwecke aufbewahren. Außerdem soll der Hersteller Informationen über Einsatzzweck, Einsatzgrenzen, Einsatzbedingungen, Funktionsweise und Fähigkeiten solcher KI-Systeme bereitstellen. Nicht zuletzt sollen Personen darüber informiert werden, wenn sie mit KI-Systemen in Kontakt treten.[60]

Bitkom-Bewertung

Dokumentationsanforderungen für Software sind bereits im geltenden Recht etabliert. Die Auslieferung einer Bedienungsanleitung für eine Software (auch als Dokumentation oder Handbuch bezeichnet) gehört zu den vertraglichen Leistungspflichten eines Softwareanbieters.[61] Bereits aus Eigeninteresse wird jeder Softwarehersteller darüber hinaus seinen Entwicklungsprozess und dessen Qualitätssicherung in gewissem Umfang protokollieren und dokumentieren, um unberechtigten Schadensersatzforderungen entgegentreten zu können. Fehlt eine solche Dokumentation gänzlich, wird regelmäßig

[60] Vgl. KI-Weißbuch der EU-Kommission, S. 23 f.
[61] BGH, Urteil vom 04.11.1992, Az. VIII ZR 165/91 für den Softwareverkauf und BGH, Urteil vom 20.02.2001, Az. X ZR 9/99 für die Erstellung einer Software nach Werkvertragsrecht

schon von einem Organisationsverschulden im Sinne der Produzentenhaftung oder sogar von einer Vernachlässigung anerkannter technischer Standards auszugehen sein.

Transparenz ist kein Selbstzweck und auch bei menschlichen Entscheidern kann ein Entscheidungsprozess nicht bis ins letzte Detail nachvollzogen werden. Letztlich kommt es nur darauf an, ob das Ausgabeergebnis eines KI-Systems Schäden verursacht oder individuelle Rechte verletzt. Für diese Beurteilung ist die Kenntnis des Entscheidungsmodells einer KI erforderlich, nicht aber die Kenntnis sämtlicher Einzelheiten der Programmierung, des Lernverfahrens und sämtlicher Trainingsdaten. Aus den Daten lässt sich das Modell nicht rekonstruieren, da es vor allem von der gewählten Parametrisierung abhängt. Die Katalogisierung, Speicherung und Verwaltung von Daten und Datenmodellen erfordert regelmäßig einen erheblichen Aufwand. In Anwendungsbereichen, die mit geringen Gewinnspannen arbeiten, ist die Verwendung von KI möglicherweise nicht mehr wirtschaftlich, wenn diese mit zusätzlichen Dokumentationsanforderungen einhergeht. Außerdem muss die Anforderung an Transparenz mit schützenswerten Interessen der Hersteller (Schutz von Geschäftsgeheimnissen) abgewogen und zum Ausgleich gebracht werden.

Das geltende Recht sieht beim Betrieb bestimmter Technologien in gewissem Umfang eine Protokollierung vor (z.B. in § 63a StVG für das teilautomatisierte Fahren). Eine ähnliche Protokollierung wäre auch für KI-Systeme denkbar, sodass bei Auswertung einer Schadenssituation mögliche Fehler und Schadensursachen besser ermittelt werden können. Verhältnismäßig wäre dies aber nur für Hochrisiko-Anwendungen.

6.8 Menschliche Aufsicht

Aus Sicht der EU-Kommission müssen KI-Anwendungen für Anwendungsbereiche mit hohem Risiko einer besonderen menschlichen Kontrolle unterworfen werden. Daher überlegt die EU-Kommission, menschliche Aufsicht und Überwachung des KI-Systems während der Entwurfsphase und während des Betriebs vorzuschreiben mit der Möglichkeit, in Echtzeit einzugreifen und das System notfalls zu deaktivieren. Auch könnten Ergebnisse eines KI-Systems erst dann endgültig Wirksamkeit erlangen, nachdem sie durch einen Menschen freigeben oder nicht von einem Menschen widerrufen wurden.[62]

Bitkom-Bewertung

Das Erfordernis einer menschlichen Kontrolle von KI leitet sich aus den ethischen Anforderungen ab, die von der Hochrangigen Expertengruppe der EU für KI formuliert wurden. Diese ethische Grundausrichtung wird von den Bitkom-Mitgliedsunternehmen

[62] KI-Weißbuch der EU-Kommission, S. 25

grundsätzlich geteilt. Es ist aber zu bezweifeln, ob sich dieser Grundsatz vollumfänglich in Rechtspflichten überführen lässt. Denn in der Folge wären autonome KI-Systeme, z.B. autonome Fahrzeuge, rechtswidrig und eine Gefährdungshaftung für autonome KI-Systeme überflüssig. In vielen Bereichen, insbesondere in der Industrie, ist der Einsatz von KI lediglich zur Unterstützung, nicht zur Ersetzung menschlicher Entscheidungen vorgesehen.

6.9 Einführung einer Rechtspersönlichkeit für KI-Systeme

Um Haftungslücken und unübersichtliche Haftungszurechnungen beim Einsatz von KI-Systemen und autonomen Robotern zu vermeiden, wird verschiedentlich vorgeschlagen, für solche System den Status einer elektronischen Person zu schaffen. Haftungsrisiken, die auf elektronische Personen zurückzuführen sind, sollen durch eine verpflichtend abzuschließende Haftpflichtversicherung abgedeckt werden.[63]

Bitkom-Bewertung

Bitkom steht Überlegungen zur Einführung einer „elektronische Person“ ablehnend gegenüber. Es kann nicht im Sinne von Gesetzgeber und Gesellschaft sein, auf diese Weise der Delegation von Verantwortung auf KI-Systeme Vorschub zu leisten. Solange (natürliche oder juristische) Personen für die Folgen des Einsatzes von KI direkt haftbar bleiben, besteht ein starkes Eigeninteresse des möglicherweise Haftenden, die KI und ihre Ergebnisse zu kontrollieren und Schadensrisiken zu minimieren. Insgesamt darf die Verantwortung für die Folgen eines Technologieeinsatzes nicht an die Technologie delegiert werden.

Es ist auch nicht erkennbar, wie der Umweg über eine im Schadensausgleich zusätzlich zwischengeschaltete elektronische Rechtspersönlichkeit einen Schadensausgleich erleichtern oder Haftungsprobleme beseitigen könnte. Denn nach wie vor müssten die Voraussetzungen für einen Haftungsanspruch dargelegt und nachgewiesen werden. Und es müsste auch weiterhin jemanden geben, der die E-Person im Rechtsverkehr vertritt oder zumindest die Versicherung für die E-Person abschließt und die Prämienlast trägt. Wenn ein insoweit Verantwortlicher identifiziert ist, kann man diese (natürliche oder juristische) Person auch direkt in die Verantwortung nehmen.

Haftungsverantwortlichkeiten lassen sich bereits durch Haftungsgrundlagen und Zurechnungsnormen des geltenden Rechts verursachungsgerecht zuweisen. Neue Regelungen für elektronische Personen werden sich kaum bruchlos und widerspruchsfrei in das Geflecht der geltenden Haftungsregelungen einfügen lassen, die im Grundsatz auf Willensfreiheit, Pflichtverletzung und Verschulden beruhen.

[63] Ein solcher Vorschlag findet sich z.B. in der Entschließung (2015/2103(INL) des Europäischen Parlaments vom 16.02.2017

7 Bitkom-Position

Der Haftungsschutz für neue Technologien darf nicht hinter dem bisher etablierten Schutzniveau zurückbleiben. Die vielen derzeit vorgeschlagenen und diskutierten Rechtsänderungen für KI würden bei gleichzeitiger Umsetzung jedoch in eine nicht gerechtfertigte Überregulierung münden, die eine erstickende Wirkung auf die weitere KI-Entwicklung in der EU hätte. Im Rahmen der Regulierung von KI werden vielfach Fragen diskutiert, die keine Spezifika von KI betreffen, sondern generell für digitale Technologien noch nicht abschließend beantwortet sind. Aus Sicht des Bitkom ist allenfalls eine punktuelle Nachjustierung des geltenden Rechts gerechtfertigt, um Risiken von KI angemessen zu adressieren.

Aus den vorstehenden Ausführungen leitet Bitkom folgende Empfehlungen ab:

- Systeme Künstlicher Intelligenz sind zunächst nichts anderes als Computer-Anwendungen (Software). Dabei ist aber streng zwischen Expertensystemen und Systemen des maschinellen Lernens zu unterscheiden. Die **allgemein mit KI in Verbindung gebrachten besonderen Risiken** (z.B. Intransparenz der Entscheidungswege, Gefahr diskriminierender Entscheidungen, Unvorhersehbarkeit eines konkreten Einzelergebnisses) **können ausschließlich beim maschinellen Lernen auftreten, soweit dieses unkontrolliert, insbesondere unabhängig vom Hersteller nach Inverkehrbringen des Produktes erfolgt**. Es ist daher einer sachangemessenen Rechtsfortbildung nicht dienlich, sondern führt nur zu Verwirrung und Rechtsunsicherheit, wenn Rechtsänderungen am Begriff der Künstlichen Intelligenz ansetzen. Stattdessen müsste – wenn überhaupt – über Rechtsänderungen für Systeme des maschinellen Lernens und des Deep Learning nachgedacht werden. Letztlich lässt sich trotz aller Bemühungen Künstliche Intelligenz nur sehr grob und näherungsweise definieren. Diese Definitionsschwierigkeiten sowie die Vielfalt möglicher Anwendungsfelder, unterschiedlicher Techniken und jeweils unterschiedlich zu gewichtender Einflussfaktoren führen zu der Erkenntnis, dass der Begriff der Künstlichen Intelligenz kein geeigneter Anknüpfungspunkt für eine rechtliche Bewertung oder weitere, noch zu entwickelnde Regelungen darstellt.
- Eigenschaften wie **Komplexität, Autonomie oder Konnektivität** sind **nicht spezifisch kennzeichnend für KI-Systeme**. Zwar kann Künstliche Intelligenz in Systemen eingesetzt werden, die autonom arbeiten, mit anderen Systemen vernetzt sind oder eine gewisse technische Komplexität aufweisen. Diese Eigenschaften sind aber nicht typisch für KI und können auch bei Systemen auftreten, die ohne KI arbeiten. Aus diesem Grund käme es zu Fehlregulierungen, wenn sämtliche KI-Systeme einer Regulierung unterworfen würden, die sich an Komplexität, Autonomie oder Konnektivität eines Systems orientiert.
- Lern- und Entscheidungsvorgänge von **KI-Systemen laufen in vorgegebenen Bahnen** und innerhalb vorgegebener Grenzen ab. Dies gilt selbst für das Deep Learning. Auch

wenn Einzelergebnisse des Systems evtl. schwer zu prognostizieren oder nachzuvollziehen sind, kann das System nicht aus den vorgegebenen Bahnen ausbrechen.

- Bei der Rechtsetzung sind ethische Normen und Grundsätze zu berücksichtigen. **Ethische Vorgaben** für Künstliche Intelligenz **eignen sich** aber **nicht** dazu, vollumfänglich **(„1:1“) in Rechtsnormen transformiert zu werden**.
- Das geltende Haftungsrecht verfolgt richtigerweise einen **technologieneutralen Ansatz**. Das bedeutet, die Haftung knüpft an der Schädigung eines Rechtsguts an, ohne dass es dabei auf die konkrete Technologie ankäme, die bei der Schadensentstehung involviert war. Von diesem Ansatz darf der Gesetzgeber nicht abweichen, indem er besondere Haftungsvorschriften für KI einführt und damit die KI-Technologie benachteiligt. Das Haftungsrecht hat eine verlässliche Kompensation für eingetretene Schäden zu gewährleisten unabhängig davon, welche Technologie zur Entstehung des Schadens beigetragen hat.
- Das geltende Recht deckt Risiken, die durch KI-Systeme begründet werden könnten, zur Genüge ab. Individuelle Rechte und Interessen möglicher KI-Anwender werden durch Diskriminierungsverbote, Produktsicherheitsvorschriften und Haftungsnormen geschützt. **Weder im Haftungsrecht noch bei der Marktzugangskontrolle ist es sinnvoll, neue Vorschriften zu schaffen, die nur für Künstliche Intelligenz gelten**. Allerdings ist bei einer Anpassung der geltenden Haftungs- und Marktzugangsvorschriften darauf zu achten, dass das geltende Haftungsrecht KI-Systeme mit besonders hohen Risiken für Leben, Leib oder Gesundheit von Menschen angemessen erfassen und keine Haftungslücken auftreten.
- Künstliche Intelligenz wird vielfach in unkritischen Anwendungen, z.B. zur Verwaltung des Arbeitsspeichers eines Computers, bei der Terminplanung, bei der Planung von Lagerbeständen und Wareneinkauf, eingesetzt. Sie kann aber auch in besonders kritischen Anwendungen Verwendung finden, deren Nutzung mit erheblicher Gefahr für Leben, körperliche Unversehrtheit und Gesundheit von Menschen verbunden ist. Daher ist sowohl im Haftungsrecht als auch im Produktsicherheits- und Marktzugangsrecht ein **risikoorientierter Regulierungsansatz notwendig**, wonach sich rechtliche Anforderungen nur dann risikoangemessen erhöhen, wenn dadurch ein nachgewiesenermaßen besonders hohes Gefahrenpotenzial kritischer Anwendungen begrenzt werden soll. Auch erhöhte Anforderungen müssen aber sinnvoll, sachangemessen, erfüllbar und verhältnismäßig sein.
- Eine Anwendung ist nicht deswegen als besonders kritisch einzustufen, weil sie KI-Komponenten umfasst. Vielmehr sind zur **Risikobewertung kritischer Anwendungen mehrere Kriterien kumulativ zu prüfen und zu erfüllen**. Dazu gehören: Grad der Entscheidungsautonomie und Grad der Lernautonomie eines Systems, Reichweite der Gefahr (beschränkt auf ein Vertragsverhältnis, innerhalb einer Lieferkette oder im öffentlichen Raum), die zumutbaren Schutzmöglichkeiten des eventuell gefährdeten Personenkreises sowie die Schutzwürdigkeit der gefährdeten Rechtsgüter. Der Grad der Entscheidungsautonomie hängt davon ab, ob eine KI-Anwendung nur als reine

Informationsquelle dient, eine Unterstützungsfunktion für menschliche Entscheidungen bietet oder ob es – in Teilen – eine menschliche Aktivität ersetzt oder völlig autonom ohne direkte menschliche Kontrolle entscheiden kann. Der Grad der Lernautonomie hängt davon ab, ob sich das Modell eines ML-Systems im Produktivbetrieb a) vollumfänglich b) nur für begrenzte, nicht sicherheitsrelevante Parameter oder c) überhaupt nicht verändern kann. Entsprechend dieser Einteilung könnten für die jeweiligen KI-Systeme "Aktionsgrenzen" definiert werden, innerhalb derer sich eine KI "bewegen" darf.

- Es kann **nicht der Anspruch an KI-Systeme sein, fehlerfrei zu agieren**. Dies ist nach derzeitigem Stand technisch nicht zu leisten. Entsprechend darf eine Sanktion den Hersteller oder Betreiber eines KI-Systems nicht schon dann treffen, wenn ein Status der Fehlerlosigkeit nicht erreicht wird. Allerdings kann KI-Systemen eine besondere, über bisher bekannte Technologien hinausgehende Gefahrneigung kann nicht nachgesagt werden. Eher im Gegenteil: indem sie menschliche Fehler (z.B. verspätete Reaktion, Subjektivität) vermeidet **kann KI menschliche Fehleinschätzungen und damit Schadensrisiken reduzieren**. Wenn diese Möglichkeiten realisiert werden sollen, dürfen KI-Verantwortliche nicht mit Haftungsrisiken belastet werden, die über die Haftungsrisiken bei anderen Technologien hinausgehen.
- Ein funktionsfähiges Haftungssystem muss ein **sorgfältiges Gleichgewicht** herstellen **zwischen Ansprüchen von Individuen auf Schadensausgleich** und Schadensprävention einerseits **und zumutbaren Anforderungen an Hersteller und Betreiber** einer Technologie sowie an die Technologie selbst andererseits. In Situationen, in denen einer Person eine rechtskonforme Entscheidung unmöglich ist, weil dafür eine Abwägung zwischen zwei gleichrangigen Rechtsgütern notwendig wäre (Dilemmasituationen), befreit das Recht die Person von einer Haftung. Dieser Grundsatz muss auch für Künstliche Intelligenz gelten.
- Das Recht muss an einigen Stellen auf technische Wertungen und Standards zurückgreifen (z.B. bei der Bestimmung eines Fehlers bei der Produkthaftung oder eines Mangels im Gewährleistungsrecht). Anders als eine isolierte Regulierung von KI ist die EU-weite **Abstimmung von technischen Qualitätsanforderungen** an KI und ihren Entwicklungsprozess **wünschenswert**, z.B. Standards für Sicherheit und Robustheit solcher Systeme in Hochrisiko-Anwendungen.
- Für den Einsatz von KI-Systemen beim Kontakt mit Verbrauchern oder mit der Öffentlichkeit **(B2C)** ergeben sich **andere Herausforderungen** und Anforderungen **als** im Geschäftsverkehr mit Unternehmen (**B2B**). Während für Privatpersonen ein hoher einheitlicher Schutzstandard erforderlich ist, sollten die Vorgaben im B2B-Bereich flexibel sein und in gewissem Umfang durch vertragliche Absprachen modifiziert werden können.