HSI
Hugo Sinzheimer Institute for
Labour and Social Security Law

The HSI is an institute of
the Hans-Böckler-Stiftung

WORKING PAPER

No. 17 · December 2022 · HSI-Working Paper

ARTIFICIAL INTELLIGENCE AND LABOUR LAW

Prof. Dr. Bernd Waas

Translated to English by Allison Felmy

Contents

A. Introduction **7**

B. Basics **8**

I. AI, Machine learning, artificial neural networks and big data 9

1. Terminology 9

2. Stages of development 11

a) Phase 1: Algorithms as rule-based procedures 11

b) Phase 2: Machine learning, artificial neural networks and data mining 12

c) Phase 3: Future development 13

II. General examples of application 14

C. The use of AI in working life **17**

I. Human resource analytics 18

II. AI and sensor technology 21

III. Application examples from different sectors 25

D. International level **29**

I. United Nations 29

1. ILO 29

2. UNESCO 31

II. Council of Europe 32

1. General 32

2. Council of Europe and artificial intelligence 33

a) Committee of Ministers 33

aa) Declaration of 13 February 2019 33

bb) Recommendation of 8 April 2020 34

cc) Recommendation of 7 March 2018 38

dd) Declaration of 17 March 2021 38

b) Work on an international treaty on AI 39

III. OECD 41

IV. G20 and others 43

V. Private initiatives: Institute of Electrical and Electronics Engineers (IEEE) 44

VI. Interim result 47

E. European Union **48**

I. The Communication on an Artificial Intelligence Initiative 48

II. The Work of the Expert Group on Artificial Intelligence 50

1. Ethics guidelines 50
2. Policy recommendations 52
3. White Paper on artificial intelligence 54
a) Basic contents 54
b) Results of the consultation 56
III. The Proposal of an "AI Law" 56
1. Overview of the draft regulation 57
a) Regulation of so-called high-risk AI systems 57
b) Requirements for high-risk AI systems 58
2. Criticism 59
a) Goals 59
b) Legal basis 60
c) Choice of Instrument 61
d) Relationship of the Regulation to the GDPR 62
e) Risk-based approach 63
aa) Unacceptable risk 64
bb) High risk 64
cc) Low risk 65
f) Standardisation and self-assessment of providers 66
g) Rights of the data subjects and legal protection 68
F. Germany 70
I. AI Strategy of the Federal Government 70
II. Data Ethics Committee 71
III. Enquete Commission 73
G. Problem areas under labour law 76
I. AI and the concept of the employee under German law 76
1. Possible starting points 76
2. The features of Section 611a BGB 77
a) Bound by instructions 77
b) Heteronomy 79
aa) Heteronomy outside of subjection to instructions? 79
bb) "Anticipation" of instructions through detailed contractual provisions 80
cc) Inclusion 82
c) The aspect of "direction" 83
aa) AI-mediated possibilities of direction 83
bb) In particular: "AI Nudging" 83
d) Summary 88

3. AI and platform employment 88

II. Possible legal capacity of AI.. 89

III. AI and exercise of the right to issue instructions................................ 93

1. Deficits in human decision-making.. 94

2. Inadmissibility of "machine decisions" 95

IV. Anti-discrimination law 100

1. The way AI works 101

2. Discrimination problems using the example of "AI recruiting"....... 106

a) Existence of "treatment" within the meaning of Section 3 AGG . 108

b) Subjective facts 108

c) Causality 109

d) Indirect discrimination 110

e) Fault.. 113

f) Burden of proof 114

3. Fundamental deficits of the current anti-discrimination law........... 115

a) Identifiability of acts of discrimination.. 115

b) Collective legal protection 116

c) The idea of prevention 117

V. Data protection 119

1. Basic terms.. 120

a) Personal data 120

b) Pseudonomysation 125

c) Profiling... 126

2. Principles for the processing of personal data 127

a) General meaning of the principles 127

b) The individual principles of Article 5 GDPR 129

aa) Lawfulness, fairness and transparency.................................. 129

bb) Purpose limitation 132

cc) Data minimisation 134

dd) Accuracy 135

ee) Storage limitation.. 136

ff) Interim result.. 137

3. Lawfulness of the processing of personal data 137

a) Lawfulness of the processing according to Article 6 GDPR 137

b) Prohibition of the processing of sensitive data 139

c) Specific regulations in Section 26 BDSG...................................... 140

4. Information obligations and rights of the data subjects.................. 142

a) Information obligations.. 142

b) Right to information ... 143

c) Further rights of data subjects ... 145

5. The prohibition of automated decisions ... 145

a) Purpose of the provision ... 146

b) Uncertain issues ... 148

6. Data protection through technology design and data protection-friendly default settings ... 150

7. Certification ... 152

8. Data protection impact assessment ... 153

9. Fundamental deficits of the current data protection law ... 154

VI. Occupational health and safety ... 156

VII Liability issues ... 162

1. Non-contractual liability ... 163

a) "Tortiousness" of the machine and "autonomy risk" ... 163

b) Individual questions of fault liability ... 164

aa) Causality ... 164

bb) Illegality and culpability ... 166

cc) AI as "vicarious agent” ... 167

c) Legal policy initiatives ... 168

aa) Recommendations of the expert group ... 168

bb) Resolution of the European Parliament ... 171

cc) Report of the Commission ... 174

2. Contractual liability ... 175

a) "Traffic" duties as contractual duties to protect ... 175

b) Digital aider and abettor liability? ... 176

VIII. Workplace co-determination ... 177

1. Recognised purposes of co-determination ... 177

a) Membership concept ... 179

b) Human dignity and personality rights ... 180

aa) Protection of human dignity ... 180

bb) Right of personality ... 183

cc) Summary ... 186

c) Collective representation of interests ... 186

d) Acceptance and expertise ... 188

2. New challenges ... 189

3. Co-determination and platform work ... 191

4. Individual regulations ... 192

a) Right of co-determination for technical equipment ... 192

aa) Prerequisites of the right of co-determination 192
bb) Right of initiative 194
b) Design of workplace, workflow and working environment 196
c) Selection guidelines 197
d) Vocational training 197
e) Involvement of experts 198
f) The regulatory instrument of the works agreement 200
5. Summary 202
H. Conclusion 203
Bibliography 207

A. Introduction

Today, we encounter artificial intelligence (AI) applications at every turn, albeit often unnoticed. AI research has developed at a rapid pace. What seemed almost unthinkable yesterday may already be an integral part of everyday life tomorrow.

AI is also becoming more and more prevalent in companies. AI applications offer firms but also their employees many opportunities; one need only think of workers' health protection. At the same time, however, the development brings considerable risks and challenges.

It is these risks and challenges in particular that this study addresses. The core questions are: What does AI mean for the protection of employees? What challenges do individual labour law and co-determination face? The fact that the study’s focus is on the risks and not on the opportunities should not be misunderstood. Since the paradigm of labour law is the protection of the employee – typically the weaker party – it should not be surprising if we focus on the question of whether and to what extent this is affected by AI.

The following sections, after a look at the basics (B.), will first look at the use of AI in working life (C.). This is already very advanced, especially in the USA. Initiatives will then be presented that focus in particular on the development of ethical principles for AI, but also in part on the regulation of AI. The analysis begins at the international level (D.), where the activities of the Council of Europe deserve special attention, then moves to the level of the European Union (E.), where the European Commission's proposal for an "AI law" will be discussed, and finally leads to the level of Germany (F.), where – in the form of some regulations of the *Betriebsrätemodernisierungsgesetz* – the first legislative activities have already been recorded. This is followed by a closer look at individual problem areas in labour law (G.). The study closes with a brief conclusion (H.).

B. Basics

A look at current articles on human resource management turns up numerous promises. One of these is that companies will increasingly be able to access an on-demand workforce.[1] In a survey of US executives conducted by the Harvard Business School and the Boston Consulting Group, 30% of respondents said they use digital platforms extensively to meet their needs for skilled workers. Almost 50% said they would use such platforms more in the future. 90% of respondents said they saw their use as a significant or very significant competitive advantage for their company. 60% saw it as possible or even very possible that their core workforce would be significantly reduced. To the same extent, managers indicated that they would prefer to "rent", "borrow" or "share" skilled workers with other companies in the future.[2] Remarkably, it is precisely the ability to "break work down into rigid, discrete components" in order to have it done on platforms that is "considered one of the most important indicators of whether a company is making the most of a blended workforce model".[3] What is propagated in the relevant literature is nothing less than "work without jobs",[4] that is, a "deconstruction" of work into individual tasks and, concomitantly, the identification and concrete deployment of employees on the basis of their skills and abilities rather than on the basis of their job descriptions.[5] The use of AI in particular is now expected to provide a further leap forward in development. The expectation is that it will allow tasks not only to be distributed among employees and platform workers, but also to be assigned to machines as needed. As a relevant article states:

"The way we have traditionally organized work and workers is becoming increasingly obsolete. We are moving toward a new *work operating system* that will deconstruct work into tasks and projects that may be assigned not only to employees but also to machines and contingent workers in talent marketplaces. In addition, workers will increasingly be identified not as holding a specific job but as possessing a set of skills and talents that can be applied wherever the organization may need them."[6]

[1] *Fuller/Raman/Bailey/Vaduganathan et al*, Building the on-demand workforce, 2020.

[2] *Fuller/Raman/Bailey/Vaduganathan et al*, Building the on-demand workforce, 2020, p. 2 f.

[3] *Fuller/Raman/Bailey/Vaduganathan et al,* Building the on-demand workforce, 2020, p. 22.

[4] Cf *Jesuthasan, Ravin /Boudreau, John*, Work Without Jobs - We need a new operating system built on deconstructed jobs and organisational agility, 5 Jan 2021, MIT Sloan Management Review; *Id*, Work without Jobs - How to Reboot Your Organisation's Work Operating System, 2022.

[5] Cf also *Jesuthasan/Boudreau*, Reinventing Jobs - A 4-Step Approach for Applying Automation to Work, 2018.

[6] *Jesuthasan/Boudreau*, Are You Ready to Lead Work Without Jobs? We're moving toward a system of work, design that will profoundly change the roles of organizational leaders, 8 April 2021, MIT Sloan Management Review.

This “work without jobs”, made possible by the use of AI and robotics, may still be in the future. But AI is already being used on a massive scale today, especially in the human resources sector. For example, AI applications not only enable comprehensive algorithmic surveillance of work performance and employee behaviour in general,[7] but also promise no less than to revolutionise human resource management by providing completely new tools. In the process, HR systems are increasingly transforming into data platforms that control every decision related to personnel.[8] In this study, this will be illustrated by examples from practical application. First, however, it is necessary to clarify what is meant by AI in the first place.

I. AI, Machine learning, artificial neural networks and big data

As will be seen shortly, this is where we encounter the first difficulties.[9]

1. Terminology

From a lay perspective, the term "artificial intelligence" would seem to have a clear content. But this is by no means the case. Rather, the term offers many different definitions, including some that are questionable. This is the case, for example, when one speaks of artificial intelligence whenever machines and/or computer program possess "human-like abilities or human-like intelligence". Apart from the fact that this leads to the question of what "intelligence" actually is, and the fact that the term "human-like" is anything but clearly delineated, the definition fails if only because there is widespread agreement that no machines or programs yet exist that work "like humans" even in the broadest sense.[10] But the definitions that focus on individual human abilities that an AI system should possess – visual perception, the ability to recognise speech, the ability to translate languages or the ability to make decisions – are not much better. After all, these definitions only describe applications. They do not actually explain or clarify anything.

[7] See only *Newlands*, Algorithmic Surveillance in the Gig Economy: The Organisation of Work through Lefebvrian Conceived Space, Organisation Studies 2020, 1.
[8] Cf. *Columbus*, How AI is shaping the future of work, 9 Jun 2022. https://venturebeat.com/2022/06/09/how-ai-is-shaping-the-future-of-work/.
[9] Instructive on this also *Herberger*, NJW 2018, 2825.
[10] See only *Bertoloni*, Artificial Intelligence and Civil Liability - Study requested by the JURI Committee, 2020, p. 18 with further references.

Instead of exploring the conceptual issues further, the definition used by the OECD will be used here, pragmatically: The Recommendation by the OECD's Council on Artificial Intelligence defines an "AI system" as a "machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments".[11]

As far as predictions in particular are concerned, reference should be made to the example of "predictive maintenance", which allows statements to be made about when maintenance must be carried out and an unplanned standstill of the machine or system can be avoided.[12] However, there are also reports of systems that are supposed to protect supply chains by predicting strikes.[13]

The definition takes the "learning capability" of AI systems into account by adding: "AI systems are designed to operate with varying levels of autonomy ". As indicated, this will be the working definition for the purposes of this study. One should, however, keep two things in mind: first, it is questionable (and also seems dangerous) to attribute "intelligence" to[14] AI systems without further qualification,[15] and second, restraint is always required when we are tempted to draw certain conclusions from the mere supposition that an AI system is "intelligent". In sum, an increasing scepticism towards the term "artificial intelligence" may be noted which critically examines the tendency toward anthropomorphism, fed partly by the concern that the "humanising" of AI expressed in the term could lead – notably on the part of the legislature – to an inappropriate handling of AI.[16] It should be noted, if only in passing, that there is often far more human intelligence behind "artificial intelligence" than one might think. For example, quite a few platforms specialise in providing companies with data sets marked by humans, which they can then use

[11] OECD, Council on Artificial Intelligence, OECD Legal Instruments, 2022, which continues: "AI systems are designed to operate with varying levels of autonomy".

[12] Cf. https://www.int.fraunhofer.de/de/geschaeftsfelder/corporate-technology-foresight/predictive-maintenance.html. In this respect, an interesting field of application is also opening up in the area of professional sport; cf. How AI Could Help Predict and Avoid Sports Injuries, Boost Performance – Computer vision, the technology behind facial recognition, will change the game in real-time analysis of athletes and sharpen training prescriptions, analytics experts say. https://www.wsj.com/articles.

[13] Cf. *Heimstädt/Dobusch*, Streik-Vorhersage mit Twitter-Daten, FAZ 11 April 2022, p. 16.

[14] See only *Fletcher/Larson*, Optimising Machines Is Perilous. Consider 'Creatively Adequate' AI - The future of artificial intelligence needs less data and can tolerate ambiguity, Jan 25, 2022: https://www.wired.com: "[...] we must banish the futurist delusion that AI is the smarter version of ourselves. AI's method of cogitation is mechanically distinct from human intelligence: Computers lack emotion, so they can't literally be courageous, and their logic boards can't process narrative, rendering them incapable of adaptive strategy. Which means that AI antifragility won't ever be human, let alone superhuman; it will be a complementary tool with its own strengths and weaknesses".

[15] On the term "artificial intelligence" being vague, also IEEE, Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System, 1st ed., 2019, p. 16, who instead speak of "autonomous and intelligent systems".

[16] Cf. *Kostopoulos*, Decoupling Human Characteristics from Algorithmic Capabilities, 2021. It is also worth noting that human-like robots are predominantly white, a circumstance that often invokes fears; cf. . *Cave/Dihal*, The Whiteness of AI, Philosophy & Technology 2020, 685.

to train machine learning algorithms.[17] Amazon, for example, describes its Mechanical Turk platform as a "web service that provides an on-demand, scalable, human workforce to complete jobs that humans can do better than computers, for example, recognizing objects in photos" Occasionally, there is talk of artificial intelligence in this context: the artificial intelligence here only appears artificial;[18] the service is actually provided by "ghost employees".[19]

2. Stages of development

To further approach the phenomenon of AI, it may additionally be helpful to take a look at how artificial intelligence systems have developed in the past and may develop further in the future.[20]

a) Phase 1: Algorithms as rule-based procedures

In the development of artificial intelligence three phases are often distinguished in the literature. The first phase of AI was characterised by precise rule-based procedures, or algorithms. A computer can follow algorithms step by step to "decide" how to react "intelligently" to a certain situation. For this reason, this is often referred to as "symbolic AI".[21] This type of AI still has its place in today's environment. But it could not be the last word (which is why some also call it "good old-fashioned AI"), and this is due to its scope of application, which was limited from the outset. The use of this AI, namely, requires fairly static environments in which the rules are strict and the variables are unambiguous and quantifiable. This form of AI is not suitable for solving complex problems, because such systems require human experts who translate their knowledge into a code the computer can understand. This considerably limits the degree of the systems' "decision-making ability". While tasks can be carried out automatically, ultimately they can only be done under human direction. Any improvement to the system requires human

[17] See Vice report, March 8, 2021: Underpaid Workers Are Being Forced to Train Biased AI on Mechanical Turk.

[18] See also *Berg/Furrer/Harmon/Rani/Silberman*, Digital labour platforms and the future of work - towards decent work in the online world, 2018, p. 7; critically *Yeung,* A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework, Council of Europe, DGI(2019)05, p. 40: "Both the training for ML models, as well as the consequent human clean-up activities to weed out the models' externalities, are often concealed to maintain the mythology of seamless automation".

[19] *Wakefield*, AI: Ghost workers demand to be seen and heard, 28 March 2021: www.bbc.com. See also *Kaushik/Lipton/London*, Resolving the Human Subjects Status of Machine Learning's Crowdworkers. https://arxiv.org/pdf

[20] *Russell, Stuart/Norvig, Peter*: Artificial Intelligence - A Modern Approach, 4th ed., 2022, p. 35 ff.

[21] Cf. only *Boucher*, Artificial Intelligence: How does it work, why does it matter, and what can we do about it?, 2020, p. 2.

intervention. And even millions of "if-then-else" rules cannot solve problems where not only the variables change in real time, but also the rules.[22]

b) Phase 2: Machine learning, artificial neural networks and data mining

The second phase of AI, in contrast, comprises newer, "data-driven" approaches. This means AI systems are themselves capable of learning. In so-called machine learning (ML), IT systems are given the ability to recognise patterns and regularities on the basis of existing data sets and algorithms and to develop solutions independently. The insights gained from the data can be generalised and used for new problem solutions or for the analysis of new data. A special method of ML is so-called reinforcement learning. Here, a program learns a strategy independently through interaction with its environment and not by being "shown" which action is best in which situation.[23]

A special method of machine learning is so-called "reinforcement learning". Here, a program learns a strategy independently through interaction with its environment and not by being "shown" which action is best in which situation.[24] The term "deep learning" refers to a method of ML that uses artificial neural networks (ANN) with numerous intermediate layers between the input layer and the output layer, thereby forming an extensive internal structure.[25] The systems used in this process are based on artificial networks inspired by the functionality of the brain and modelled on neurons in the nervous system. Inputs are translated into signals, which are then routed through a network of artificial neurons. The more neurons and the more layers added to the network, the more complex the problems that can be solved. Deep learning requires networks with at least two hidden layers, each containing many neurons. With more layers, artificial neural networks can develop more abstract conceptualisations of problems by breaking them down into smaller sub-problems and providing more sophisticated answers.[26] The down side: The significantly increased complexity of the systems goes hand in hand with

[22] Cf. *Boucher*, Artificial Intelligence: How does it work, why does it matter, and what can we do about it?, 2020, p. 2 f.

[23] Cf. O'Gieblin, Prediction Engines Are Like Karma: You Get What You Stream, 18 Jun 2022: - https://www.wired.com/story/prediction-engines-are-like-karma-you-get-what-you-stream/.

[24] Cf. Reinforcement Learning Repository, University of Massachusetts, Amherst: https://all.cs.umass.edu/rlr/; cf. also *Russell, Stuart/Norvig, Peter*: Artificial Intelligence - A Modern Approach, 4th ed., 2022, p. 840 ff.

[25] In more detail *Russell, Stuart/Norvig, Peter*: Artificial Intelligence - A Modern Approach, 4th ed., 2022, p. 801 ff.

[26] See also *Boucher*, Artificial Intelligence: How does it work, why does it matter, and what can we do about it?, 2020, p. 3 ff.

significantly reduced explainability and transparency.[27] In this context, AI can undoubtedly help to find answers to countless research questions. However, the problem then arises that people remain unaware of how the AI arrived at its solution. Accordingly, AI must be further developed in such a way that it can not only provide answers, but also the basis for a "scientific understanding" of them.[28]

The developmental "leaps and bounds" occurring in AI technology could not be explained without the combination of AI and so-called data mining. The term designates an area of data processing that focuses on the automatic identification of patterns or anomalies in data sets. Big data refers to data sets that are so large or complex that they cannot be processed with conventional methods, or only with difficulty. It is important to realise that today there are, on the one hand, enormous amounts of data waiting to be analysed and, on the other hand, we have both the computer power and the storage capacity to realise this analysis. As far as the amount of data is concerned, it increased by more than thirty times between 2012 and 2020. AI drives the demand for data, which leads to technological innovations that aim at nothing other than collecting and evaluating new types of data.[29] Only recently does the realisation seem to be gaining more ground that more attention should be paid to the quality than the quantity of data.[30]

c) Phase 3: Future development

We are now in the midst of this second phase with the third phase yet to come. Therefore much is speculative here. Among the facets being discussed are a self-explanatory or contextual AI which may even be able to solve the aforementioned

[27] See also, for example, *Reichwald/Pfisterer,* CR 2016, 208 (212) with the consideration that with increasing "autonomy" of machines "neither the correctness or a certain behaviour can be guaranteed nor a decision taken can be retrospectively traced in the absence of logging".

[28] Optimistic Krenn/Pollice/Guo et al., On scientific understanding with artificial intelligence, p. 9. https://doi.org/10.48550/arXiv.2204.01467: "[…] we firmly believe that these research efforts can – within our lifetimes – transform androids into true agents of understanding that will directly contribute to one of the most essential aims of science, namely Scientific Understanding."

[29] See only *Kresge*, Data and Algorithms in the Workplace: A Primer on New Technologies, UC Berkeley Labor Center Working Paper, Technology and Work Program, November 2020; "[...] a crucial contributing factor to the success of the technology platforms [...] is their ability to analyse digital data streams and deploy that analysis in a technological system. The systems consist of data-mining, predictive analytics, and machine learning algorithms that identify and segment users into micro-categories of consumers for personalised, targeted advertising. Together, increased volumes of data and the development for new systems for analysis of that data constitute the core of the digital transformation of the economy". The author impressively describes the "ecosystem of employee data and data-driven technologies"; ibid, (12 ff.).

[30] *Brown*, Why it's time for "data-centric artificial intelligence", 7 June 2022, https://mitsloan.mit.edu/ideas-made-to-matter/why-its-time-data-centric-artificial-intelligence?utm_campaign=Artificial%2BIntelligence%2BWeekly&utm_medium=email&utm_source=Artificial_Intelligence_Weekly_279.

problem of (limited) explainability because it is "self-explanatory";[31] a further alignment with human intelligence (for example by creating "digital copies" of the human brain); and the creation of an "artificial consciousness".[32] When and whether AI will ever be able to detach itself from its respective context, that is, not only tackle specific problems but behave in a "generally intelligent" way, and how to achieve this,[33] is the subject of heated debate.[34] In any case, however, this is not a present-day problem.[35] Another future problem, the "distribution of tasks" between humans and machines, is also unresolved. It seems plausible that the foreseeable future will belong to a "hybrid intelligence" in which, in the best-case scenario, both humankind and machine will contribute their (complementary) strengths.[36]

II. General examples of application

A few examples are in order to illustrate that development is progressing rapidly and applications are within reach today that were not even in sight just a short time ago. Facebook has launched a project to teach an AI system to "understand" videos. Based on this, the system is then supposed to make recommendations on Instagram. What is remarkable here is that the software teaches itself to classify videos, analysing both the visual content and the audio content.[37] A group of Chinese researchers affiliated with the Alibaba Group and Tsinghua University have created what they claim is the largest Chinese-language AI system to date: It is a multimodal system trained on images as well as texts with about 100 billion

[31] See *Elton*, Self-explaining AI as an Alternative to Interpretable, in: Goertzel/Panov/Potapovm/Yampolskiy (eds.), Artificial General Intelligence, 13th International Conference, AGI 2020, St. Petersburg, Russia, September 16-19, 2020, Proceedings, 2020, p. 95.

[32] Cf. *Freed*, Report on "AI and Human Thought and Emotion", ibid. , p. 116; *Boucher*, Artificial Intelligence: How does it work, why does it matter, and what can we do about it?, 2020, p. 13 ff.

[33] See *Freed*, AGI Needs the Humanities, in: Goertzel/Panov/Potapovm/Yampolskiy (eds.), Artificial General Intelligence, 13th International Conference, AGI 2020, St. Petersburg, Russia, September 16-19, 2020, Proceedings, 2020, p. 107 with a discussion of various fields (drama, literature as a research field for imagination and metaphors, linguistics, music and hermeneutics) from which inspiration for novel "human-like" AI could be drawn.

[34] Cf. *Boucher*, Artificial Intelligence: How does it work, why does it matter, and what can we do about it?, 2020, pp. 13 and 17.

[35] However, "optimists" have recently been given a boost by an article in the renowned journal Artificial Intelligence. In it, the authors state that reinforcement learning will one day lead to replicating human cognitive abilities *and* achieving *artificial general intelligence.* This would create an AI that would be superior to humans in almost every cognitive task; cf. *Silver/Singh/Precup/Sutton*, Reward is enough, in: Artificial Intelligence, October 2021, 103535; critically, *Vamplew et al.,* Scalar reward is not enough: A response to Silver, Singh, Precup and Sutton (2021): https://arxiv.org/pdf/2112.15422.pdf.

[36] See only *Dellermann/Ebel/Söllner/Leimeister*, Hybrid Intelligence, 2018.

[37] Fortune March 12, 2021: Facebook reveals A.I. that is already improving Instagram video recommendation.

variables. It was trained on 1.9 terabytes of images and more than 292 gigabytes of text.[38] The use of robots made of flexible components ("soft-bodied robots") could receive a boost from a newly developed deep-learning technique: MIT researchers have used a neural network to plan where sensors should be placed so the robots can take the best shape for a particular task.[39] Developers of responsive soft robots have drawn inspiration from ketchup bottles.[40] The research company OpenAI has developed an algorithm that uses reinforcement learning to achieve "superhuman" results on all 55 classic Atari 2600 games. The significance of this development goes far beyond the actual field of application because the researchers have succeeded in equipping the application with a kind of memory: The system creates an archive of all previous actions and places it has explored. It then selects a place to return to and explore further, using a kind of rule of thumb for the states that seem most promising. [41]

At the same time, it seems as if many AI applications are becoming "mass-produced". At any rate, the number of so-called no code AI platforms, which denotes software that enables people without special coding knowledge to create algorithms, is increasing rapidly. For example, the company Primer recently developed a programming-free platform called Automate that enables non-specialists to train an AI system in about 20 minutes so that it can perform its tasks with an accuracy that approaches human levels. Because the system works with a powerful, pre-trained AI algorithm that only needs to be tuned to a client's specific needs, the company says it can deliver good results with as few as 10 to 20 examples if they are chosen carefully.[42]

Some developments are unreservedly positive. OpenAI, for example, has developed methods that can look into the inner workings of artificial neural networks and thus supposedly make their notoriously opaque decision-making more interpretable. Incidentally, this has revealed that individual neurons in a large neural network can encode a particular concept, a finding that parallels what neuroscientists have discovered in the human brain.[43]

[38] *Junyang/Men/Yang*: M6: A Chinese Multimodal Pretrainer, 2021: https://arxiv.org/abs/2103.00823.
[39] SciTechDaily March 29, 2021: MIT's New Artificial Intelligence Algorithm Designs Soft Robots That Sense.
[40] *Amolf*, Responsive soft robots inspired by sputtering ketchup bottle, 8 July 2022. https://techxplore.com/news/2022-07-responsive-soft-robots-sputtering-ketchup.html?utm_campaign.
[41] *Ecoffet/Huizinga/Lehman/Stanley/Clune*, First return, then explore, Nature 2021, 580.
[42] https://primer.ai/products/primer-automate.
[43] *Kahn*, What they've found will surprise you, Fortune 4 March 2021.

Some other things seem at least ambivalent, some give cause for thought and some, cause for concern. Consider for example the rapidly developing so-called neurotechnology,[44] in which the brain becomes the interface for communication between human consciousness and information technology communication systems. In research, imaging techniques are used to understand the functioning of the human brain and to identify the neural correlates of mental states and behaviour. Clinical applications of brain imaging, as well as other neurotechnologies, can significantly improve the well-being of patients with neurological diseases and offer the prospect of new preventive, diagnostic and therapeutic tools. But at the same time, they open up the possibility of being able to "read minds". For example, scientists were able to determine from decoded brain activity with 70 per cent accuracy which of two alternative activities the subjects would perform.[45] Another study showed that movement patterns of people recorded by their smartphones can be used to diagnose early signs of Alzheimer's disease.[46] These and other findings have led researchers to call for the guarantee of a number of so-called neuro-rights.[47] Specifically, the so-called Morningside Group, composed of neuroscientists from around the world, calls for the recognition of (1) the right to identity, or the ability to control both one's physical and mental integrity; (2) the right to agency, freedom of thought and free will to choose one's actions; (3) the right to mental privacy, or the ability to keep one's thoughts from disclosure; (4) the right to fair access to mental augmentation, or the ability to ensure that the benefits of sensory and mental enhancements through neurotechnology are equitably distributed across the population; and (5) the right to protection from algorithmic bias, or the ability to ensure that technologies do not introduce bias.[48]

[44] On the increasingly strong mutual influence of neurotechnology or neuroscience on the one hand and AI on the other, see *Ienca*, Brain Machine Interfaces, Artificial Intelligence and Neurorights: https://brain.ieee.org/.

[45] *Haynes/Sakai/Rees/Gilbert/Frith/Passingham*, Reading Hidden Intentions in the Human Brain, Current Biology 2007, 323.

[46] *Nieto-Reyes/Duque/Montaña/Lage*, Classification of Alzheimer's Patients through Ubiquitous Computing Sensors 2017, 1679.

[47] Cf only *Dayton*, Call for human rights protections on emerging brain-computer interface technologies - Industry self-regulation is not enough, say AI researchers: nature index 16 March 2021; *Ienca/Andorno*, Towards new human rights in the age of neuroscience and neurotechnology, Life Sciences, Society and Policy, 2017: https://doi.org/10.1186/s40504-017-0050-1 2017; also *Yuste/Genser/Herrmann*, It's Time for Neuro-Rights - New Human Rights for the Age of Neurotechnology: Horizons, 2021, 154; cf. Genser/Herrmann/Yuste, International Human Rights Protection Gaps in the Age of Neurotechnology, 2022.

[48] *Yuste/Genser/Herrmann*, It's Time for Neuro-Rights - New Human Rights for the Age of Neurotechnology: Horizons, 2021, 154.

C. The use of AI in working life

AI is increasingly being used in working life.[49] This is also true in Germany.[50] For example, the head of Stepstone recently reported in the FAZ that now "practically all hiring processes take place digitally in some form".[51] The use of AI is diverse and may also include the use of so-called bots to conduct job interviews.[52] However, the use of AI has particularly flourished in the USA. Here, there are three main areas of application: Human resource analytics, which aims, for example, at hiring employees and evaluating their performance; algorithmic management, which includes workforce scheduling and the coordination and control of employee activities; and finally, task automation, which includes the use of robots. It is worth noting that the use of AI often goes hand in hand with other technologies. This is obvious for the latter area, which involves the interplay of AI and robotics. However, the interaction of AI and sensor technology is also particularly "fruitful". This is then no longer just about optimising production processes using sensors to measure and monitor temperatures, vibrations, pressure, fill levels, humidity, speed, weight, acceleration, inclination, etc and then employing AI to analyse the data,[53] but about the combined use of sensors and AI to monitor and "control" workers.

In the following, we will first take a closer look at the broad field of human resource analytics. Then we will focus on "AI and sensor technology". Examples of applications from different sectors will then round off the picture.

[49] Cf. on this most recently *De Stefano/Wouters*, AI and digital tools in workplace management and evaluation – An assessment of the EU's legal framework, May 2022, p. 10 ff.
[50] Cf only *Thieltges,* ZfP 2020, 3 (19 ff.).
[51] FAZ 31 Dec 2021, p. 28: "Praktisch alle Einstellungsprozesse finden jetzt in irgendeiner Form digital statt".
[52] Cf. *Johnson*, 7 effective uses of AI in recruitment. https://www.unleash.ai/artificial-intelligence/7-effective-uses-of-ai-in-recruitment/. The author describes the advantages as follows: "Robots, also known as bots, are now trained to conduct physical interviews as part of the hiring process. These bots use both natural language processing (NLP) and interview analytics to assess the candidate's suitability by skimming their soft skills and personality traits. The use of bots to conduct physical interviews is beneficial to recruiters, as they guarantee consistency in the interview process since the same interview experience is meant to provide equal experiences to all candidates."
[53] https://www.industrie-energieforschung.de/forschen/kuenstliche-intelligenz.

I. Human resource analytics

The term human resource (people) analytics usually describes digital applications based on large collections of data for purposes of measuring, analysing and forecasting the performance of employees, designing workplaces or identifying and developing talent.[54] What is generally considered to be the advantage of human resource analytics is that the relevant applications are evidence-based. This is why, according to a widely expressed view, they are clearly preferable to decisions based on "gut feeling".[55] Research shows that 98% of all companies in the Fortune 500 use applicant tracking systems in their hiring processes (sourcing, screening, interviewing and selection/rejection).[56] AI applications promise to make hiring processes more objective and at the same time significantly less time-consuming. One such application has been developed by the company HireVue. The software extracts up to 25,000 data points from video interviews, examines visual and verbal cues and compares word choice, facial movements, body language and tone of voice to infer certain personality traits and thus identify the best candidates for a given job.[57] Human resource analytics seems to be enjoying triumphant success on the whole. According to a recent study, corporate HR departments are now even more data-driven than finance departments, with AI and machine learning being key drivers.[58]

Human resource analytics allows the targeted use of data and data analyses in human resource management in conjunction with other company data to support human resource management decisions and processes.[59] In this context, AI is used to identify cause-and-effect relationships in the company (such as causes of high fluctuation) and to forecast future developments and events.[60] This is done on

[54] See *Moore*, Data subjects, digital surveillance, AI and the future of work, 2020, p. 18; see also *Collins/Fineman/Tsuchida,* People analytics: Recalculating the route, Global Human Capital Trends, Deloitte Insights, February 28, 2017, with the following objectives of HR analytics: "to measure, report and understand employee performance, aspects of workforce planning, talent management and operational management".

[55] Cf. only *Huff/Götz*, NZA Supplement 2019, 73.

[56] Thus *Sánchez-Monedero/Dencik/Edwards,* What Does It Mean to 'Solve' the Problem of Discrimination in Hiring?, 2019.

[57] See *McGuire*, There's no going back: how AI is transforming recruitment, Personnel Today 20 January, 2021; *Ajunwa,* An Auditing Imperative for Automated Hiring Systems, Harvard Journal of Law & Technology 2021, 1 (19).

[58] HR Gazette "People Analytics: 10 Trends to Watch in 2020": https://hr-gazette.com/people-analytics-10-trends-to-watch-in-2020.

[59] *Huff/Götz*, NZA Supplement 2019, 73 (73): "helfen datengestützte Erkenntnisse, wirksame Strategien für kritische Herausforderungen zu entwickeln und das Personalmanagement agil und zielwirksam an den Erfordernissen im Unternehmen auszurichten"; cf. on the whole also *Chalutz Ben-Gal*, Human Resources Based Organizational Data Mining (HRODM): Themes, Trends, Focus, Future (2020).

[60] Sceptical *Nowotny*, In AI we trust: power, illusion and control of predictive algorithms, 2021. The trust in the "predictive power" of algorithms is often matched by the idea of an "ethical AI", but sceptical recently e.g.*Gill,* AI & SOCIETY 2022, 411 (in her review of the aforementioned book): "In seeking certainty in algorithmic predictions, we are in danger of

the basis of comprehensive software solutions that can permeate the entire company and make information available to users without prior knowledge. The promise of human resource analytics is that "data-driven insights help develop effective strategies for critical challenges and align human resource management with business needs in an agile and purpose-effective way".[61] Human resource analytics come in many guises. But there are two unifying elements: (1) the search for new pools of quantitative data that correlate with business and employment success, and (2) the use of such data to replace subjective human decisions with (supposedly) objective decisions.[62]

One example of the application of HR analytics is so-called data-enhanced leadership, in which managers are evaluated by their employees under various categories. The quality of leadership can also be viewed in a cause-and-effect context, so that statements can be made about which behaviour of a manager has triggered which reaction among the employees, leading for instance to increased job satisfaction. The output of results can be aggregated for one's own area of responsibility. However, it is just as conceivable that the qualities of the managers can be viewed and evaluated "from above" and, in particular, in comparison.[63]

Another application example is fluctuation analyses and fluctuation forecasts. Such analyses can be used to calculate the fluctuation rate and the fluctuation costs for different employee groups. Perhaps even more significant is the fluctuation forecast. Here, based on a cause-and-effect relationship learned from past data, the system can calculate an individual future fluctuation risk for each employee, and this in real time due to its capacity for linking with new data. This makes it immediately recognisable which employees have a particularly high risk of fluctuation (so-called regretted leavers). Since the premature departure of top performers would be particularly painful and costly for the company, it is worth

‘renouncing the inherent uncertainty of the future and replacing it with the dangerous illusion of being in control’. There is also a tacit assumption and misplaced confidence that smart AIs would ultimately take care of the unresolved ethical, transparency and accountability conficts when we are able to develop computational tools ‘to assess the performance and output quality of deep learning algorithms and to optimise their training’. The danger is that ‘we end up trusting the automatic pilot while flying blindly in the fog’, becoming part of a fine-tuned and inter-connected predictive system, thereby diminishing our motivation and ability to stretch the boundaries of imagination.”

[61] *Huff/Götz*, NZA Supplement 2019, 73 (73). However, it seems that the expectations associated with the use of people analytics are not always fulfilled; cf. only *Marabelli/Vaast/Carlile,* Making Lemonade: Dealing with Analytics Surveillance in the Workplace, Academy of Management Annual Meeting, 2020.

[62] Thus *Bodie/Cherry/McCormick/Tang*, The Law and Policy of People Analytics, Saint Louis U. Legal Studies Research Paper No. 2016-6, 1 (11 f.).

[63] *Huff/Götz*, NZA Supplement 2019, 73 (77).

taking timely preventive measures.[64] However, there is a recognisable risk of abuse here in particular.[65]

Notably, some applications of human resource analytics are at the interface with what is known as gamification of work in the USA. The phenomenon can also be observed in Germany, in increasing measure.[66] Here, work is enriched with playful elements, which may be a welcome change not least from the employee's point of view, but is not without its dangers. According to the literature: “In people analytics, games are being used for their predictive power, often to quantify or measure particular skills or aptitudes or to screen job candidates.[67] The stream of responses provided by a job candidate in a computer game could tell an employer how that candidate would respond to a work challenge. At the same time, having a game as part of a job interview could perhaps encourage the candidates to play, have fun, relax and perhaps let their guard down. The hope is that the candidates may show their “true colors” instead of the stilted and perhaps narrow affect that a candidate typically shows in an in-person interview”.[68]

The start up Knack, for example, has developed games that prompt participants to make countless decisions, actions and reactions. This so-called "micro-behaviour" is then analysed by algorithms with the goal of producing a meaningful potential analysis. [69]

A process called organisational network analysis (ONA) or relationship analytics is rapidly gaining importance in the context of human resource analytics. This involves determining the strength, frequency and type of interactions of people in a professional network in order to then develop patterns of their collaboration. AI makes it possible to identify communication and collaboration systems in the

[64] *Huff/Götz*, NZA Supplement 2019, 73 (77).

[65] So also *Huff/Götz*, NZA Supplement 2019, 73 (78): "Thus, with the wrong conception, it is not only recognisable for the employer who is particularly valuable in the workforce and is 'on the way out', but also who is unproductive, dissatisfied and thus disagreeable in the eyes of superiors."See also *Renan Barzilay*, Data Analytics at Work: A View From Israel on Employee Privacy and Equality in the Age of Data-Driven Employment Management, Comparative Labor Law & Policy Journal 2019, 421.

[66] Cf. in this respect only BAG, NZA 2021, 552 on the legal status of a crowdworker (at para. 50), according to which the defendant company "stimulate[d] the 'play instinct' of the users through the promise of experience points and the associated benefits with the aim of inducing them to take up regular employment" (“durch die Inaussichtstellung von Erfahrungspunkten und den damit verbundenen Vorteilen den ‘Spieltrieb’ der Nutzer an[regte] mit dem Ziel, diese dadurch zu einer regelmäßigen Beschäftigung zu bewegen”).

[67] So-called recruitainment; cf. on this also Gamification goes Recruiting – Wird der neue Job in Zukunft erspielt?, euroforum 22 Aug 2017.

[68] *Bodie/Cherry/McCormick/Tang*, The Law and Policy of People Analytics, 1 (13 f.).

[69] *Bodie/Cherry/McCormick/Tang*, The Law and Policy of People Analytics, 1, (16).

company and analyse them on a real-time basis.[70] This makes it possible, for example, to show how strongly company departments are networked with each other, how high the knowledge transfer actually is (for instance through file uploads), what the mood in the company is like and where the opinion leaders or experts in the company can be found.[71]

Two things are noteworthy about all this: first, human resource analytics is no longer a separate task for a few specialists, but is embedded in all processes and procedures in the company. Second, the analytics team is always "up to speed" because the relevant data is constantly updated and re-evaluated. AI is then used to predict fraud patterns, uncover trust networks between employees, illuminate interactions between employees, show correlations between coaching and employee engagement, and also determine patterns for employee time management. It is not uncommon for the justifications given in favour of using AI to have the ring of benevolent care, as when the analysis of employees' travel data and other data is intended to "improve their energy levels, well-being and performance".[72]

II. AI and sensor technology

AI also has a wide range of applications in the processing of what is called sociometric data. For example, wearable sensors can be used to collect different types of information about people’s behaviour, ranging from the duration of their conversations (with the respective parts of the conversation), voice pitch and gestures (arm and hand movements, nodding, facial expressions)[73] to their spatial

[70] One of the companies offering ONA advertises it as follows: "In today's hyper-connected workplaces, organisational relationship networks are opening up new data insights into how employees communicate, collaborate and influence each other to get their work done. Research shows the average employee sustains around 130 work relationships at any one time. These networks are a key asset for individuals and organisations alike; however, until now these relationships have not been visible or accessible to the organisation. TrustSphere's People Analytics solutions leverage Relationship Analytics and Organizational Network Analysis across enterprise communication and collaboration systems on a real-time basis. Without ever looking at content, TrustSphere analyses digital interactions across the organisation. Proprietary algorithms generate a range of actionable insights that enable HR and Talent Management teams to make better data-driven decisions around": https://www.trustsphere.com/ona-for-people-analytics.

[71] Wikipedia "People Analytics".

[72] See *Fineman*: People analytics: Recalculating the route, Deloitte Insights, 2017.

[73] Much attention has been paid in the press to the case of an Amazon driver who resisted the installation of a camera in his vehicle that permanently registered facial expression and body movements: For this Amazon van driver, AI surveillance was the final straw, news.trust, 27 Mar 2021; cf also more recently: Amazon Delivery Drivers Forced to Sign 'Biometric Consent' Form or Lose Job, Vice report, 23 Mar 2021.

position. On the basis of the collected data, developers intend for it to be possible to forecast the success of teams. So-called data signatures can also be developed, which can be used to describe, for instance, a natural, charismatic leader. Such a data signature is then integrated into an app as a benchmark so that managers can check "in real time" whether they are displaying the required communication and leadership behaviour.[74] The game app from the provider Knack can be used to collect data on individual decision-making behaviour (time spent considering, order of action, type of problem solving, etc). This data is then used to measure the characteristics of the participants (such as creativity and social intelligence) and, if necessary, to predict the innovative capacity of individual employees.[75]

Interestingly, quite a few applications take advantage of the trend to increase health and well-being. Especially by employing (acceleration and infrared) sensors, data can be collected that may be related to physical and emotional well-being and physical and mental health, but may certainly also have significance in terms of increasing the productivity of employees.[76] These applications often use graphical user interfaces (dashboards) to visualise the data. There can be many reasons for establishing common dashboards, but is certainly also a suitable means of ensuring competition between employees. Today there are already applications that can determine calorie consumption depending on activity and heart rate. Google is reported to be working on a lens that can determine a person's blood sugar level from tear fluid. Tiny LED lights surrounding the lens are supposed to indicate when the blood sugar level has reached a certain threshold. Smart lenses would also have the ability to take one reading per second, providing information about changing blood glucose levels. Another fast-growing area of research and development is the detection and measurement of emotions. This involves determining people's oxytocin levels to determine what they are "really feeling" (sentiment analysis).[77] AI technologies are moving in the same direction, registering emotional reactions "in real time" by "decoding facial expressions, analysing speech patterns, monitoring eye movements and measuring neurological immersion levels".[78] The employee experience is a particular trend

[74] *Kaiser/Kraus*, ZfO 2014, 379 (380).

[75] *Kaiser/Kraus*, ZfO 2014, 379 (380).

[76] Cf. *Moore*, Data subjects, digital surveillance, AI and the future of work, 2020, p. 21 with a reference to the so-called "Qantified Self" movement. This is a network of users and providers of methods as well as hardware and software solutions with the help of which, for example, environmental and personal data can be recorded, analysed and evaluated. A central goal is to gain knowledge about personal, health and sporting issues as well as personal habits; cf. wikipedia "Quantified Self".

[77] See *Purdy/Zealley/Maseli, O.* (2019). The Risks of Using AI to Interpret Human Emotions. Harvard Business Review 18/11/19.

[78] See *Moore*, Data subjects, digital surveillance, AI and the future of work, 2020, p. 23.

which many consider to be promising. The idea is that by combining and analysing both qualitative and quantitative employee data from a variety of sources, companies will learn what makes their employees "tick". The results are often described as "win-win", since not only the commitment, well-being, performance and employee loyalty are increased, but also the profitability of the company.[79]

In general, it seems that AI is increasingly being used to uncover (inner) characteristics of people. It is no longer just about determining the "trustworthiness" of a person, for example based on their payment habits and other financial information. Rather, it is very generally using communication data and relationships in social networks on- and offline.[80] Among other things, the literature reports on experiments by researchers and developers to detect emotional states from keyboard strokes;[81] to derive sensitive information (including health status) from telephone metadata;[82] to detect emotions and develop psycho-demographic "profiles" based on data from the online network Twitter;[83] to identify criminal tendencies[84] and genetic diseases using automated facial recognition;[85] to determine sexual orientation based on Facebook contact lists[86] and to identify psychological traits from "digital footprints" such as "likes" or posts on Twitter;[87] to determine various personality traits, such as sexual orientation, ethnicity, religious and political attitudes, age, gender or intelligence from "likes" on Facebook[88] and to recognise sexual orientation, especially homosexuality, from pictures of

[79] HR Gazette "People Analytics: 10 Trends to Watch in 2020": https://hr-gazette.com/people-analytics-10-trends-to-watch-in-2020.

[80] See *Wei/Yildirim/ Van den Bulte*, Credit Scoring with Social Network Data, Marketing Science 2016, 234. The authors see the risk that people may change their social behaviour in order to increase their creditworthiness; ibid, 250. These and all subsequent examples in *Orwat*, Discrimination Risks through the Use of Algorithms, 2020, p. 11.

[81] *Epp/Lippold/Mandryk*, Identifying emotional states using keystroke dynamics, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2011, p. 715.

[82] *Mayer/Mutchler/Mitchell*, Evaluating the privacy properties of telephone metadata; in: Proceedings of the National Academy of Sciences 2016, 5536 (5540). The authors conclude that it is not uncommon for highly sensitive insights to be derived from telephone metadata when combined with data from other, easily accessible sources.

[83] *Volkova/Bachrach*, On Predicting Sociodemographic Traits and Emotions from Communications in Social Networks, in: Cyberpsychology, Behavior, and Social Networking 2015, 726 (735).

[84] *Wu/Zhang*, Automated inference on criminality using face images 2016, 4038 (4052). Based on extensive experiments and rigorous cross-validations, the authors conclude that data-driven face classifiers are able to make reliable inferences on criminality through supervised machine learning.

[85] *Gurovich et al.* , Identifying facial phenotypes of genetic disorders using deep learning; in: Nature medicine 2019, 60.

[86] *Jernigan/Mistree*: Gaydar: Facebook friendships expose sexual orientation; in: First Monday, 2009, No. 10

[87] *Matz/Netzer*, Using Big Data as a window into consumers' psychology; in: Current Opinion in Behavioral Sciences 2017, 7.

[88] *Kosinski/Stillwell/Graepel/Thore*, Private traits and attributes are predictable from digital records of human behaviour; in: Proceedings of the National Academy of Sciences, 2013, 5802 (5805): "We show that a variety of people's personal traits, from sexual orientation to intelligence, can be automatically and accurately inferred from their Facebook likes. The similarity between Facebook Likes and other widely used types of digital records, such as browsing histories, search queries, or purchase histories, suggests that the potential to uncover users' characteristics is likely not limited to Likes. Furthermore, the wide variety of traits predicted in this study suggests that with appropriate training data, it may be possible to uncover other traits as well".

people.[89] A report has recently caused a stir, according to which Chinese researchers have succeeded in developing software that can recognise the loyalty of members of the Communist Party by their facial expressions.[90]
Whether and to what extent AIs can really make such statements in a responsible manner, or whether they are "charlatanism", is judged varyingly in the literature.[91]

Coming back to the specific area of working life, numerous AI applications are used in the context of monitoring work performance and increasing productivity.[92] One example of this is systems for electronic performance monitoring. This includes monitoring emails, tapping phones, tracking computer content and usage times, video surveillance and GPS tracking. The data collected is intended to provide indications of employee productivity, but also provides information on location, email use, intensity of website surfing, printer use, phone use and tone of voice, as well as movements during a conversation.[93] The advantage of using AI is described as being that data can be used in real time: "The days of annual performance reviews are numbered. Instead, managers are increasingly using real-time data analytics to identify the drivers of their employees' performance and thus obtain immediately actionable information for feedback, promotions, compensation, skills development and career planning".[94]

[89] *Kosinski/Wang*, Deep neural networks are more accurate than humans at detecting sexual orientation from facial images; in: Journal of Personality and Social Psychology, 2018, 246. In their paper, the authors believe they have shown that faces contain much more information about sexual orientation than could be perceived or interpreted by the human brain.

[90] China Boasts of ‘Mind-reading’ Artificial Intelligence that Supports ‘AI-tocracy’. https://www.voanews.com/a/china-boasts-of-mind-reading-artificial-intelligence-that-supports-ai-tocracy-/6651986.html?tpcc.

[91] Cf. *Narayanan*, How to recognise AI snake oil (set of slides): https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf.

[92] Cf. also *Spencer/Cole/Joyce/Whittaker/Stuart*, Digital Automation and the Future of Work, European Parliamentary Research Service 2021, p. 38 f. In the UK, three in five workers said in a survey that they had been monitored at work in the past year, with surveillance of devices and phone calls believed to have increased sharply during the pandemic; cf. *Reece*, Workers say no to increased surveillance since COVID-19.
https://www.tuc.org.uk/blogs/workers-say-no-increased-surveillance-covid-19.

[93] See only *Moore*, The Quantied Self in Precarity - Work, Technology and What Counts, 2018, p. 146 ff.; see also *Ball*, Electronic Monitoring and Surveillance in the Workplace – Literature review and policy recommendations, 2021.

[94] HR Gazette "People Analytics: 10 Trends to Watch in 2020": https://hr-gazette.com/people-analytics-10-trends-to-watch-in-2020.

III. Application examples from different sectors

In the US, AI is now found in almost all industries.[95] AI is particularly widespread in call centres, which use so-called labour management systems. Here, the pandemic has led to many employees working from home, where they are comprehensively monitored by AI. Video surveillance systems are also often used, to which employees are required to contractually agree.[96] If misconduct is detected, such as the use of a mobile phone outside working hours, the system sends notifications to the supervisor, who can then intervene immediately. Another frequently-used program records the conversations of call centre employees and customers. Based on an analysis of customer sentiment and employee behaviour, the system provides real-time behavioural guidance to employees on a computer dashboard, encouraging them to be more empathetic, more efficient in their conversations or more confident. Supervisors have permanent access to this dashboard. This also includes a "customer experience score" based on the employee's performance metrics such as call efficiency, sales and customer satisfaction. The provider Cogito promotes its system of so-called augmented intelligence as follows: "AI can read *honest signals* conveyed by voices within conversation to suggest behavioural changes to keep the interaction successful and productive". This involves analysing a conversation millisecond by millisecond for over 200 different vocal and non-verbal signals, which are then analysed and matched with insights from millions of conversations from the company's own database.[97]

Warehouses and distribution centres are also an important field of application for AI.[98] Here, AI measures the productivity of employees. The data is collected by wearable trackers worn by each employee. These measure, for example, scanning rates, the number of incorrect scans and the duration of interruptions between scans. Some systems show productivity results in the form of rankings.

[95] The following examples are taken from a recent study by the UC Berkeley Labor Center; see *Bernhardt/Kresge/Suleiman*, Data and Algorithms at Work - The Case for for Worker Technology Rights, November 2021.

[96] See *Solon*, Big Tech Call Centre Workers Face Pressure to Accept Home Surveillance, NBC News, 8 Aug 2021: https://www.nbcnews.com/tech/tech-news/big-tech-call-center-workers-face-pressure-accepthome-surveillance-n127622.

[97] Cogito Corporation, Augmented Intelligence in the Contact Centre: The Why, What, and How, 2020: https://cogitocorp.com.

[98] See *Bernhardt/Kresge/Suleiman*, Data and Algorithms at Work - The Case for for Worker Technology Rights, November 2021, p. 7 f.

Supervisors receive real-time productivity metrics from employees. Systems can also send automatic messages to HR. AI systems are also reported to automatically issue terminations in given cases.[99] Other systems assign specific tasks to employees. So-called "lead-me" carts direct workers from one storage location to another, determine the pace of work and give instructions on what product and what quantity of items to pick at a particular station.[100] There is also a degree of gamification in these areas. For example, there are "fitbit competitions" organised (allegedly) for motivational purposes, though it is not uncommon for one shift to compete against another.[101] In Amazon warehouses, screens are installed next to the employees' workstations that run simple games. Putting together orders and moving items are translated into virtual movements in a game. So, for example, the faster someone selects items and puts them in a box, the faster their car drives along a virtual track.[102]

AI is also widely used in retail and grocery shops. AI systems analyse data to predict customer demand and make decisions about the most efficient deployment of staff. This involves adjusting schedules as new data becomes available. One of the systems used in this area estimates the sales productivity of each employee and creates schedules based on the corresponding values. In doing so, the program also allows for employee preferences to be taken into account. But it is reported that the corresponding functions are often not enabled.[103] One of the largest grocery delivery services allows customers to monitor workers as they assemble and scan items and also communicate with them. The customer receives notifications about estimated delivery times. The customer can also directly evaluate the workers' performance. The system measures the accuracy of the employees, the speed with which they fulfil orders and the degree to which they stick to a given script in chats with customers. This information is merged with customer ratings. Employees receive regular notifications about their performance. If speed and quality standards are not met, this can easily be sanctioned.[104] In a sense, the monitoring of employees is outsourced to the customers, which

[99] *Lecher*, How Amazon Automatically Tracks and Fires Warehouse Workers for Productivity, The Verge, August 25, 2019. https://www.theverge.com/2019/4/25/18516004/amazon-warehouse-fulfillment-centers-productivity-firing-terminations.

[100] See also *Dzieza*, Robots Aren't Taking Our Jobs - They're Becoming Our Bosses. The Verge, February 27, 2020: https://www.theverge.com/2020/2/27/21155254/automation-robots-unemployment-jobs-vs-human-google-amazon.

[101] *McCrea*, Labor Management Systems (LMS), The New Age of Employee Engagement, Logistics Management, 3 June 2020: https://www.logisticsmgmt.com.

[102] *Vincent*, Amazon Turns Warehouse Tasks into Video Games to Make Work "Fun," The Verge, May 22, 2019. https://www.theverge.com/2019/5/22/18635272/amazon-warehouse-working-conditions-gamification-video-games.

[103] See *Bernhardt/Kresge/Suleiman*, Data and Algorithms at Work - The Case for for Worker Technology Rights, November 2021, p. 9.

[104] See also *Bhuiyan*, Instacart shoppers say they face unforgiving metrics: It's a very easy job to lose, Los Angeles Times. August 27, 2019.

arguably proves to be extremely effective.[105] Another platform translates performance metrics into a "pay algorithm", but the calculation of pay reportedly often remains opaque and not infrequently appears to be inequitable.[106]

In transport, the use of AI serves primarily to monitor employees. Thus many trucks are equipped with sensors that measure location, braking and acceleration patterns, frequency of lane changes, speed and seat belt habits. In addition, dashcams and audio recording technologies record the driver's activities in the truck cab. This data is then further analysed using image processing systems, facial analysis and object recognition. Video surveillance in particular has long outgrown simple recording by a camera. This then reads in a trade journal as follows: “Connectivity, complex algorithms and deep data sets are now transforming these products into powerful safety and efficiency tools. Video safety systems now provide an array of features to reduce driver behavior risk
including real-time coaching, sharper images that move over higher speed broadband, facial recognition, improved analytics, and easier integration with traditional GPS-based tracking systems”.[107]

In the construction industry, so-called geofencing and geolocation technologies are increasingly being used to determine locations. These systems work via apps that are installed on employees' mobile phones, use the phone's GPS function and automatically clock their owners in and out when they enter and leave the construction site. Safety monitoring systems are also used; these analyse video footage to monitor the proper wearing of protective equipment. Other systems are also used to prevent accidents. For example, some systems continuously track the movements of workers on the construction site. If the system detects a hazard, the person concerned can be warned immediately by vibrations on a wristband.[108]

The application of AI in the USA is taking place in a "regulatory vacuum".[109] The situation in Germany and Europe is different. But the point here was only to show

[105] See *Levy/Barocas*, Refractive Surveillance: Monitoring Customers to Manage Workers, International Journal of Communication 2018, 1166.

[106] See again *Bernhardt/Kresge/Suleiman*, Data and Algorithms at Work - The Case for for Worker Technology Rights, November 2021, p. 9.

[107] *Clinton*, Smarter Video Telematics Wave Arrives, Automotive Fleet, 19 Mar 2019: https://www.automotive-fleet.com/327438/wave-of-smarter-video-telematics-solutions-arrives.

[108] Cf. on the whole *Bernhardt/Kresge/Suleiman*, Data and Algorithms at Work - The Case for for Worker Technology Rights, November 2021, p. 13 f. with further references.

[109] Thus explicitly *Bernhardt/Kresge/Suleiman*, Data and Algorithms at Work – The Case for Worker Technology Rights, November 2021, p. 18. However, there are signs - at least in part - that the legislature could take countermeasures. The California Fair Employment and Housing Council (FEHC) recently presented a draft regulation that specifically targets the use of "automated decision-making systems" in hiring and employment.

what AI applications are already capable of today. A few things stand out: Some applications are definitely useful. Nevertheless, the fact remains that employees are "under constant surveillance". It is also conspicuous that many applications amount to transferring tasks from superiors to customers. Finally, it is noteworthy that quite a few apps make use of the human play instinct. This is especially true of apps that are used quite frequently in hiring processes.

D. International level

In the meantime, there are countless initiatives dealing with the development of ethical principles in dealing with AI, some of which aim to regulate AI applications. The following is a brief overview of these initiatives.

I. United Nations

As one would expect, the United Nations has a prominent role.[110]

1. ILO

The work of the International Labour Organisation (ILO), a specialised agency of the United Nations, has the objective to promote social justice and guarantee universal human and social rights. Governments, trade unions and employers work together in the ILO, in particular, drawing up legally binding conventions that have come to regulate labour law with almost global coverage.

The ILO has also been dealing with the future issues of labour law for some time. A corresponding initiative was proposed by the Director-General of the ILO in 2013 as one of the seven initiatives for the centenary of the ILO. Its first task was to initiate a series of national dialogues.[111] These were followed by the report of an independent commission. The initiative culminated in the adoption of the Centenary Declaration on the Future of Work in 2019.[112]

[110] Cf. also for an overview International Telecommunication Union, United Nations Activities on Artificial Intelligence (AI), 2021: https://www.itu.int.
[111] See ILO, Synthesis Report of the National Dialogues on the Future of Work, 2017.
[112] Cf. https://www.ilo.org/global/topics/future-of-work/lang--en/index.htm.

The 2019 report of the Global Commission on the Future of Work includes reflections on the role of AI.[113] The panel advocates a "human-in-command" approach. This should ensure that "the final decisions affecting work are taken by human beings, not algorithms". The dignity of workers must be protected.[114] In addition, the Commission calls in particular for protection of workers' data and protection against discrimination.[115] It is noteworthy that the expert panel also associates hopes with the use of AI: "Technology, including artificial intelligence, robotics and sensors, carries with it countless opportunities to improve work: the extraction of knowledge through the use of data mining can assist labour administrations to identify high-risk sectors and improve labour inspection systems; digital technologies such as apps and sensors can make it easier for companies and social partners to monitor working conditions and labour law compliance in supply chains; blockchain technology – which provides transparency and security through encrypted blocks and decentralized databases – could guarantee the payment of minimum wages and facilitate the portability of skills and social protection for migrant workers, or the payment of social security for those working on digital labour platforms”.[116] In this respect, it is also noteworthy that not only governments, but also workers' and employers' organisations are explicitly called upon to "invest" in digital technologies.[117]

As for the Centenary Declaration on the Future of Work adopted at the Labour Conference, it includes a commitment to "harnessing the fullest potential of technological progress and productivity growth, including through social dialogue, to achieve decent work and sustainable development, which ensure dignity, self-fulfilment and a just sharing of the benefits for all".[118]

[113] Global Commission on the Future of Work, Work for a brighter future, 2019.

[114] Report, 2019, p. 43.

[115] Report, 2019, p. 44. Incidentally, the recommendation formulated in this context of "developing an international governance system for digital labour platforms that obliges platforms (and their clients) to respect certain minimum rights and protections" is also noteworthy.

[116] Report, 2019, p. 43 f. On blockchain technology, see also *Kritikos*, What if blockchain could guarantee ethical AI?, 2020.

[117] Report, p. 44; specifically on the work of trade unions, the report states: "Workers' organisations need to adopt innovative organising techniques - including the use of digital technology to organise labour. Workers across diverse workplaces and countries can be organized through digital means and engage in new forms of connected action. Digital technology provides workers' organisations with the potential to connect with workers outside traditional workplaces and offer new services, such as the mining of data to design effective strategies and the sharing of information about crowdworking platforms or portable benefits"; ibid, p. 42.

[118] Cf. ILO Centenary Declaration for the Future of Work (at II. A. (ii)).

2. UNESCO

Outside the ILO, the United Nations is also dealing with the topic of AI. In particular, UNESCO (United Nations Educational, Scientific and Cultural Organization) should be mentioned here, likewise a specialised agency of the United Nations, with the task of contributing to the preservation of peace and security by promoting international cooperation in education, science, culture and communication.

On the basis of a draft submitted by a group of experts[119] and following an extensive consultation process, UNESCO Member States in November 2021 adopted Recommendations on the ethics of artificial intelligence,[120] which can be considered the first global understanding on common rules.

One section of the recommendations is dedicated to the area of "economy and labour". It states, among other things, that "Member States should work with private sector companies, civil society organizations and other stakeholders, including workers and unions to ensure a fair transition for at-risk employees". It also calls on Member States to "encourage and support researchers to analyse the impact of AI systems on the local labour environment in order to anticipate future trends and challenges". These studies should "investigate the impact of AI systems on economic, social and geographic sectors, as well as on human-robot interactions[121] and human-human relationships, in order to advise on reskilling and redeployment best practices".[122] UNESCO's recommendations are thus primarily aimed at the consequences that the use of AI has on the employability of workers.

In addition, the recommendations contain the explicit recognition of a right to privacy and the demand for adequate data protection[123] – and thus some items that may appear less than spectacular from a European perspective, but are quite remarkable in international terms.

[119] Ad Hoc Expert Group (AHEG) for the preparation of a draft text of a recommendation on the ethics of artificial intelligence, SHS/BIO/AHEG-AI/2020/4 REV.2, Paris, 7 Sept 2020: https://unesdoc.unesco.org/ark:/48223/pf0000373434.

[120] UNESCO, Recommendation on the ethics of artificial intelligence.

[121] In the meantime, a whole branch of research has been established to deal with them; see only *Kim*, Working With Robots: Human Resource Development Considerations in Human-Robot Interaction, Human Resource Development Review 2022, 48.

[122] Recommendations No. 117 and 118.

[123] Recommendations 32 and 33.

II. Council of Europe

1. General

Within the Council of Europe, too, the discussion of AI issues takes up a good deal of space.[124] The Council is of course a European and thus regional international organisation.[125] Headquartered in Strasbourg, it currently has 47 Member States with a total of 820 million citizens. According to Article 1 of its Statute, the Council of Europe has the task of "achieving a greater unity between its members". Though it is a forum for debate on general European issues, binding international treaties are also concluded within the framework of this organisation. Prominent among these is the European Convention on Human Rights (ECHR), whose guardian is the European Court of Human Rights (ECtHR). The European Social Charter (ESC), for which special monitoring procedures exist, is also relevant from a labour law perspective. Important bodies of the Council of Europe are, apart from the Secretary General,[126] the Committee of Ministers of the Council of Europe, a decision-making body,[127] in which the Member States are represented by their foreign ministers or their permanent representatives; and the Parliamentary Assembly, an advisory body[128] to which the parliaments of the Member States send representatives. Also worth mentioning is the Commissioner for Human Rights,[129] an independent institution of the Council of Europe with the task of promoting the protection of human rights in the Member States and pointing out possible deficits in this area to the Council's bodies.

[124] See for example Council of Europe, Council of Europe work on Artificial Intelligence, SG/Inf(2019)21 of 02.07.2019.

[125] Overview page: https://www.coe.int/en/web/artificial-intelligence/home.

[126] https://www.coe.int/en/web/secretary-general/home.

[127] https://www.coe.int/web/cm.

[128] https://pace.coe.int/en/. The Parliamentary Assembly also elects the Secretary General, the Commissioner for Human Rights and the judges of the European Court of Human Rights.

[129] https://www.coe.int/en/web/commissioner.

2. Council of Europe and artificial intelligence

Given that the Council of Europe's work is in great measure committed to safeguarding human rights, it is not surprising that these rights are at the heart of the organisation's engagement with the challenges posed by AI. [130]

a) Committee of Ministers

The Committee of Ministers of the Council of Europe has dealt with the matter several times in recent years.[131]

aa) Declaration of 13 February 2019

In February 2019, the Committee of Ministers adopted a Declaration "on the manipulative capabilities of algorithmic processes".[132] Here it acknowledges that "[a]dvanced technologies play a pivotal role in maintaining the efficiency and public service value of digitisation, in strengthening individual autonomy and self-determination, and in enhancing human flourishing by creating optimal conditions for the exercise of human rights".[133] And yet, the Committee also sees dangers associated with these technologies, including for individuals who, "because of their particularly large digital footprint, are especially exposed to new forms of data-driven surveillance".[134] One of these dangers is the possibility of "micro-targeting of individuals based on profiles".[135] Also, “[D]ata-driven technologies and systems are designed to continuously achieve optimum solutions within the given parameters specified by their developers. When operating at scale, such optimisation processes inevitably prioritise certain values over others, thereby

[130] See also Council of Europe, Algorithms and Human Rights - Study on the human rights dimensions of automated data processing techniques and possible regulatory implications, Council of Europe study DGI(2017) 12 prepared by the committee of experts on internet intermediaries (MSI-NET), 2018.

[131] Cf also most recently Recommendation CM/Rec(2021)8 of the Committee of Ministers to Member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling (Adopted by the Committee of Ministers on 3 November 2021 at the 1416th meeting of the Ministers' Deputies).

[132] Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes (Adopted by the Committee of Ministers on 13 February 2019 at the 1337th meeting of the Ministers' Deputies), Decl (13/02/2019)1.

[133] Declaration No. 3.

[134] Declaration No. 5.

[135] Declaration No. 6.

shaping the contexts and environments in which individuals, users and non-users alike, process information and make their decisions".[136] In this context, the Committee of Ministers expressly details the manipulative capabilities of AI, stating: "Contemporary machine learning tools have the growing capacity not only to predict choices but also to influence emotions and thoughts and alter an anticipated course of action, sometimes subliminally. The dangers for democratic societies that emanate from the possibility to employ such capacity to manipulate and control not only economic choices but also social and political behaviours, have only recently become apparent. In this context, particular attention should be paid to the significant power that technological advancement confers to those – be they public entities or private actors – who may use such algorithmic tools without adequate democratic oversight or control".[137] The Declaration continues: "Fine grained, sub-conscious and personalised levels of algorithmic persuasion may have significant effects on the cognitive autonomy of individuals and their right to form opinions and take independent decisions. These effects remain underexplored but cannot be underestimated".[138]

bb) Recommendation of 8 April 2020

In April 2020, the Committee of Ministers adopted a Recommendation on the impact of algorithmic systems on human rights.[139] The Committee, which also uses this instrument to formulate what it terms "Guidelines", here recommends inter alia that Member States "[R]eview their legislative frameworks and policies, as well as their own practices with respect to the procurement, design, development and ongoing deployment of algorithmic systems to ensure that they are in line with the guidelines set out in the appendix to this recommendation; promote their implementation in all relevant areas and evaluate the effectiveness of the measures taken at regular intervals, with the participation of all relevant stakeholders";[140] "ensure, through appropriate legislative, regulatory and supervisory frameworks related to algorithmic systems, that private sectoractors engaged in the design, development and ongoing deployment of such systems comply with the applicable laws and fulfil their responsibilities to respect human

[136] Declaration No. 7.
[137] Declaration No. 8.
[138] Declaration No. 9.
[139] Recommendation CM/Rec(2020)1 of the Committee of Ministers to Member States on the human rights impacts of algorithmic systems (Adopted by the Committee of Ministers on 8 April 2020 at the 1373rd meeting of the Ministers' Deputies).
[140] Recommendation No. 1.

rights in line with the UN Guiding Principles on Business and Human Rights[141] and relevant regional and international standards";[142] and "endow their relevant national supervisory, oversight, risk assessment and enforcement institutions with the necessary resources and authority to investigate, oversee and co-ordinate compliance with their relevant legislative and regulatory framework, in line with this recommendation".[143]

Looking at the guidelines themselves, the first point of interest is that they aim not only at obligations of states, but also at responsibilities of private sector actors. With regard to the latter, the Guidelines explicitly state that "Private sector actors engaged in the design, development, sale, deployment, implementation and servicing of algorithmic systems, whether in the public or private sphere, must exercise due diligence in respect of human rights." They have, further, "the responsibility to respect the internationally recognised human rights and fundamental freedoms of their customers and of other parties who are affected by their activities". According to the Committee of Ministers, this responsibility exists "independently of States' ability or willingness to fulfil their human rights obligations". In terms of content, the Committee of Ministers calls for private sector actors to take "continuing, proactive and reactive steps to ensure that they do not cause or contribute to human rights abuses and that their actions, including their innovative processes, respect human rights".[144]

The Committee of Ministers does not fail to recognise that "[o]perating typically by detecting patterns in large datasets, algorithmic systems offer the potential to improve the performance of services (particularly through increased precision, targeting and consistency), provide new solutions, and deliver returns in efficiency and effectiveness of task and system performance". Thus these systems have "led to immense improvements in the categorisation and searchability of digital information and have facilitated important advances in fields such as medical diagnostics, transportation and logistics, enabling the broader and faster sharing of information globally and making possible new forms of co-operation and co-ordination. As a result, they permeate many aspects of contemporary human life".[145] However, there are dangers that should not be underestimated: "[T]here are also significant human rights challenges attached to the increasing reliance on

[141] UN Guiding Principles on Business and Human Rights, UN doc A/HRC/17/31.
[142] Recommendation No. 3.
[143] Recommendation No. 4.
[144] See C.1.1. Under C.1.3. the horizontal effect of human rights is again explicitly addressed.
[145] See A.3.

algorithmic systems in everyday life, such as regarding the right to a fair trial; the right to privacy and data protection; the right to freedom of thought, conscience and religion; the right to freedom of expression; the right to freedom of assembly; the right to equal treatment; and economic and social rights. The functionality of algorithmic systems is frequently based on the systematic aggregation and analysis of data collected through the digital tracking at scale of online and offline identity and behaviour of individuals and groups. In addition to the intrusion on individuals' privacy and the increasing potential of highly personalised manipulation,[146] tracking at scale can have a serious adverse effect on the exercise of human rights, which must be considered throughout the entire life cycle of an algorithmic system, from the proposal stage onward".[147] Here the Committee of Ministers highlights the potential fallibility of many AI systems, cautioning that "While it is often argued that the costs are offset by gains in rationalisation and accuracy, it is important to note that most algorithmic systems are based on statistical models in which errors form an inevitable part, sometimes with feedback loops that maintain, replicate and reinforce pre-existing biases, errors and assumptions. Although it may seem as if larger datasets provide better chances of finding recurrent patterns and correlations, accuracy rates do not automatically increase with the size of the dataset. As a result of the large number of people affected by algorithmic systems, the number of errors in the form of false positives and false negatives, and of people who are affected by these errors and inbuilt bias, will also expand, triggering additional interferences with the exercise of human rights in multiple ways (…)".[148]

The content of the individual guidelines cannot be discussed in detail here. However, it should be noted that some of them are surprisingly specific. The section on the obligations of states calls for informational self-determination,[149] transparency,[150] identifiability[151] and the existence of sufficient legal remedies (contestability).[152] A further demand is for a human rights impact assessment;[153]

[146] In fact, "hyper-personalisation" is one of the most important trends in marketing; cf. only Deloitte, Omnia AI - Connecting with Meaning, Hyper-personalizing the customer experience using data, analytics and AI: https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/deloitte-analytics/ca-en-omnia-ai-marketing-pov-fin-jun24-aoda.pdf.

[147] See A.4.

[148] See A.5.

[149] Cf. B.2.1.

[150] Cf. B.4.1.

[151] Cf. B.4.2.

[152] Cf. B.4.3.

[153] Cf. B.3.1.

this is also one of the central demands of the human rights commissioners of the Council of Europe.[154] The guidelines are also relatively specific with regard to the responsibilities of private sector actors. They call for protection against discrimination,[155] the requirement of consent to the use of AI,[156] but also transparency, accountability and the availability of effective legal remedies.[157] They also call for private sector actors to "actively engage in participatory processes with consumer associations, human rights advocates and other organisations representing the interests of individuals and affected parties, as well as with data protection and other independent administrative or regulatory authorities, on the design, development, ongoing deployment and evaluation of algorithmic systems, as well as on their complaint mechanisms".[158]

As significant as the Recommendation and its Guidelines are, it is also clear that the world of work plays only a limited role. On the one hand, the preamble emphasises up front the need to "ensure that racial, gender and other societal and work-related inequalities that have not yet been eliminated in our societies are not intentionally or inadvertently perpetuated by algorithmic systems".[159] Also, the Recommendation explicitly addresses the importance of using AI in recruitment and other selection processes.[160] Finally, states are called upon to "incentivise technological innovation in line with existing human rights, including social rights and internationally recognised labour and employment standards".[161] On the other hand, however, this is not further elaborated in the guidelines.

[154] See Commissioner for Human Rights, Unboxing Artificial Intelligence: 10 steps to protect Human Rights; Council of Europe, 2019, p. 7.

[155] See C.1.4: "Private sector actors that design, develop or implement algorithmic systems should follow a standard framework for human rights due diligence to avoid fostering or entrenching discrimination throughout all life-cycles of their systems. They should seek to ensure that the design, development and ongoing deployment of their algorithmic systems do not have direct or indirect discriminatory effects on individuals or groups that are affected by these systems, including on those who have special needs or disabilities or who may face structural inequalities in their access to human rights".

[156] Cf. C.2.1.: "Private sector actors should ensure that individuals who are affected by their algorithmic systems are informed that they have the choice to give and revoke their consent regarding all uses of their data, including within algorithmic datasets, with both options being equally easily accessible. Users should also be given the possibility to know how their data are being used, what the real and potential impact of the algorithmic system in question is, how to object to the processing of their data, and how to contest and challenge specific outputs. Consent rules for the use of tracking, storage and performance measurement tools of algorithmic systems must be clear, simply phrased and complete, and should not be hidden in the terms of service".

[157] Under C.4.

[158] Under C.4.5.

[159] At the same time, it is described as desirable to "correct these imbalances through the use of appropriate technologies".

[160] Under A.8.

[161] Under B.6.3.

cc) Recommendation of 7 March 2018

Although not central to the issue of AI, another recommendation of the Committee of Ministers also deserves interest. This is the Recommendation on the roles and responsibilities of internet intermediaries.[162] The term "internet intermediaries" is used to describe actors who "facilitate interactions on the internet between natural and legal persons by offering and performing a variety of functions and services".[163] For these actors specifically the Recommendation urges the need to respect human rights. It explicitly emphasises that "[l]aws, regulations and policies applicable to internet intermediaries, regardless of their objective or scope of application, including commercial and non-commercial activities, should effectively safeguard human rights and fundamental freedoms, as enshrined in the European Convention on Human Rights, and should maintain adequate guarantees against arbitrary application in practice".[164] States, it then goes on to say, have the "obligation to protect human rights and fundamental freedoms in the digital environment. All regulatory frameworks, including self- or co-regulatory approaches, should include effective oversight mechanisms to comply with that obligation and be accompanied by appropriate redress opportunities".[165]

dd) Declaration of 17 March 2021

In addition, the Committee of Ministers has adopted a statement on the risks of computer-assisted or AI-assisted decision-making in the social safety net.[166] In it, the Committee urgently warns of the dangers of AI: ”The unregulated development of such computer-assisted or automated decision-making systems, coupled with a lack of transparency and insufficient public scrutiny, and their incorporation into the administration of social services, pose risks. These systems can, if not developed and used in accordance with principles of transparency and legal certainty, amplify bias and increase risks. This may lead to higher negative impact for members of the community who are in a situation of vulnerability. Under such circumstances,

[162] Recommendation CM/Rec(2018)2 of the Committee of Ministers to Member States on the roles and responsibilities of internet intermediaries (Adopted by the Committee of Ministers on 7 March 2018 at the 1309th meeting of the Ministers' Deputies).

[163] Recommendation, Preamble No. 4.

[164] Recommendation No. 1.1.2.

[165] Recommendation 1.1.3.

[166] Declaration by the Committee of Ministers on the risks of computer-assisted or artificial-intelligence-enabled decision making in the field of the social safety net (Adopted by the Committee of Ministers on 17 March 2021 at the 1399th meeting of the Ministers' Deputies), Decl(17/03/2021).

they can replicate entrenched discrimination patterns, including as regards women, and can affect people in low-skilled and poorly paid jobs”.[167]

b) Work on an international treaty on AI

Some time ago, a committee, the Ad Hoc Committee on Artificial Intelligence (CAHAI), was set up with the task of examining the prospects of introducing an international agreement on AI.[168] If such an agreement were to be reached, it would undoubtedly be of considerable importance for the European Union and Germany. It should be noted that Council of Europe Convention No. 108 for the Protection of Individuals with regard to Automatic Processing of Personal Data[169] is widely regarded as the precursor to the European General Data Protection Regulation (GDPR). Convention No. 185 on Cybercrime, the so-called "Budapest Convention", has set standards (even worldwide) for regulations in this area.[170]

In December 2020, the Committee presented a report[171] which contains two things: an overview of the various studies that have been produced under the Committee on the impact of AI systems on human rights, the rule of law and democracy; and a collection of experiences from various states that could be used in the development of an international legal framework for the use of certain AI systems. Furthermore, also in December 2020, the Committee presented a feasibility study with a view to developing a legal framework for the development, design and application of AI. In addition to the promise held by AI, the study also describes the risks it brings from a human rights perspective. The report also makes notable findings on working life: “AI systems are increasingly used to monitor and track workers, distribute work without human intervention and assess and predict worker potential and performance in hiring and firing situations. In some situations, this can also have detrimental consequences for workers’ right to decent pay, as their

[167] Declaration by the Committee of Ministers on the risks of computer-assisted or artificial-intelligence-enabled decision making in the field of the social safety net (Adopted by the Committee of Ministers on 17 March 2021 at the 1399th meeting of the Ministers' Deputies), Decl(17/03/2021), p. 1.

[168] A precise description of the task at https://rm.coe.int/leaflet-cahai-en-june-2020/16809eaf12: "examine the feasibility and potential elements, on the basis of broad multi-stakeholder consultations, of a legal framework for the development, design and application of artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law".

[169] https://rm.coe.int/1680078b37.

[170] https://rm.coe.int/special-edition-budapest-convention-en-2022/1680a6992e. See also Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108), Report on Artificial Intelligence, (Convention 108), Report on Artificial Intelligence, Artificial Intelligence and Data Protection: Challenges and Possible Remedies, 2019.

[171] Council of Europe, Towards Regulation of AI Systems - Global perspectives on the development of a legal framework on Artificial Intelligence (AI) systems based on the Council of Europe's standards on human rights, democracy and the rule of law, Compilation of contributions DGI (2020)16 prepared by the CAHAI Secretariat, 2020.

pay can be determined by algorithms in a way that is irregular, inconsistent and insufficient. Furthermore, AI systems can also be used to detect and counter the unionisation of workers. These applications can jeopardise the right to just, safe and healthy working conditions, dignity at work as well as the right to organise. The discrimination capacity of AI systems that assess and predict the performance of job applications or workers can also undermine equality, including gender equality, in matters of employment and occupation".[172]

In conclusion, the feasibility study finds that no international legal instrument exists to date that is specifically tailored to the challenges of AI. It also concludes that the current level of protection of existing international and national instruments is fragmentary. It continues: "An appropriate legal framework will likely consist of a combination of binding and non-binding legal instruments, that complement each other. A binding instrument, a convention or framework convention, of horizontal character, could consolidate general common principles – contextualised to apply to the AI environment and using a risk-based approach – and include more granular provisions in line with the rights, principles and obligations identified in this feasibility study. This international legal instrument could then be "combined with additional binding or non-binding sectoral Council of Europe instruments to address challenges brought by AI systems in specific sectors".[173]

In addition, the Committee launched an extensive consultation process with the aim of gathering voices on a possible legal framework for AI. The consultation involved governments and public administrations, international organisations, businesses, civil society, academia and the technical community.[174]

At the last plenary session in November-December 2021, the Committee then adopted the statement on "Possible Elements of a Legal Framework for Artificial Intelligence, Based on the Council of Europe Standards on Human Rights, Democracy and the Rule of Law". These were submitted to the Committee of Ministers for further consideration.[175]

[172] CAHAI, Feasibility Study, CAHAI(2020)23, Strasbourg, 17 December 2020, p. 10.
[173] CAHAI, Feasibility Study, CAHAI(2020)23, Strasbourg, 17 December 2020, p. 56.
[174] For more information, see: https://www.coe.int/en/web/artificial-intelligence/cahai-multi-stakeholder-consultation; see also CAHAI, Analysis of the Multi-Stakeholder Consultation, CAHAI(2021)07.
[175] https://www.caidp.org/resources/coe-ai-treaty/.

III. OECD

Initiatives to regulate AI have also been developed by the Organisation for Economic Co-operation and Development (OECD). The OECD's activities have since led to a recommendation, which in turn is based on preliminary work by a group of experts.[176] Adopted in 2019,[177] the Recommendation, which is the first intergovernmental standard for AI, is of interest for two reasons: First, it lists five principles for the responsible use of trustworthy AI: inclusive growth, sustainable development and well-being; people-centred values and fairness; transparency and explainability; robustness, safety and security; and accountability. Second, it provides concrete guidance for national policies and international cooperation: investing in AI research and development; fostering a digital ecosystem for AI; shaping an enabling policy environment for AI; building human capacity and preparing for labour market transformation; and international cooperation for trustworthy AI.

In terms of principles, it states under "Human-centred values and fairness" that "AI actors should respect the rule of law, human rights and democratic values, throughout the AI system lifecycle. These include freedom, dignity and autonomy, privacy and data protection, non-discrimination and equality, diversity, fairness, social justice, and internationally recognised labour rights".[178] As for the recommendations on national policies and international cooperation, under the theme of "building human capacity and preparing for labour market transformation", it states that governments should "work closely with stakeholders to prepare for the transformation of the world of work and of society. They should empower people to effectively use and interact with AI systems across the breadth of applications, including by equipping them with the necessary skills". Governments should also "take steps, including through social dialogue, to ensure a fair transition for workers as AI is deployed, such as through training programmes along the working life, support for those affected by displacement, and access to new opportunities in the labour market". Finally, governments should "also work closely with stakeholders to promote the responsible use of AI at work, to enhance

[176] AI Group of experts at the OECD (AIGO).
[177] Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, Adopted on: 22/05/2019.
[178] Under 1.2a).

the safety of workers and the quality of jobs, to foster entrepreneurship and productivity, and aim to ensure that the benefits from AI are broadly and fairly shared".[179]

In February 2021, the OECD presented a "Framework for the Classification of AI systems" developed by its expert group. This combines the characteristics of AI systems with the OECD's 2019 AI principles, classifying AI systems and AI applications along the following dimensions: Human & Planet, Economic Context, Data & Input, AI Model and Task & Output. The framework aims to help policymakers, regulators, legislators and others to assess the opportunities and risks of different types of AI systems so as to inform their AI strategies and ensure policy coherence across borders.[180] It is also worth noting that the OCED has established the OECD AI Policy Observatory as a platform for multidisciplinary studies on AI in particular.[181]

OECD recommendations are not legally binding. But they are not legally irrelevant. Illustrative of this is a complaint filed some time ago with the US Federal Trade Commission (FTC) against HireVue, a leading provider of job interview software. The complainant alleged various violations of the OECD Recommendation, and invoked the OECD Principles which it cited among the "established public policies" as defined by the Federal Trade Commission Act (FTA).[182][183]

[179] Under 2.4.

[180] Cf. https://wp.oecd.ai/app/uploads/2022/02/Classification-2-pager-1.pdf.

[181] https://www.oecd.ai/.

[182] 15 U.S.C. § 45(n): "[...] In determining whether an act or practice is unfair, the Commission may consider established public policies as evidence to be considered with all other evidence [...]".

[183] https://epic.org/wp-content/uploads/privacy/ftc/hirevue/EPIC_FTC_HireVue_Complaint.pdf. In January 2021, the company announced that it would no longer use facial recognition software: https://epic.org/hirevue-facing-ftc-complaint-from-epic-halts-use-of-facial-recognition/.

IV. G20 and others

Other international actors have also recognised the challenges of AI systems. In this respect, the group of the twenty most important industrialised and emerging countries (G20) should be mentioned in particular. In 2019, the G20 adopted a declaration on artificial intelligence which builds on the principles developed by the OECD.[184] At the G20 summit in Riyadh, the group expressly committed to a "human centred approach".[185] In the final declaration of the G20 Summit in Rome, the group reaffirmed its intent to implement the G20 principles on AI.[186]

In September 2021, a group of renowned scientists addressed the governments of the G20 Member States, but also the G20 as such, with a series of recommendations. The recommendations to the Member States include the following: "define human-centric AI in terms of meaningful human control, transparency, explainability, fairness, justice, inclusiveness, sustainability, and education [...], "interpret AI systems as a support to human decision-making, not a replacement. Do not recognise machines as moral agents and do not give them an electronic personality or identity"; "apply a multi-stakeholder approach to all decisions regarding AI"; "when regulating AI, impose conditions on the uses of AI (and not AI per se)". The recommendations to the G20 include: "define a standard glossary including all aspects of human-centric AI" and "set up an independent and multi-disciplinary AI ethics committee, including representatives of all 20 country-level AI ethics committees".[187]

[184] G20 Ministerial Statement on Trade and Digital Economy: https://www.mofa.go.jp/files/000486596.pdf.

[185] Leaders' Declaration G20 Riyadh Summit November 21 - 22, 2020, No. 19.

[186] Cf. https://www.consilium.europa.eu/media/52730/g20-leaders-declaration-final.pdf. At the G7 level, the so-called *Charlevoix Common Vision for the Future of AI should be* mentioned, which also contains a commitment to a "human-centred" approach: https://www.mofa.go.jp/files/000373837.pdf.

[187] *Casalone et al.* , Human-centric AI: From Principles to Actionable and Shared Policies, September 2021.

V. Private initiatives: Institute of Electrical and Electronics Engineers (IEEE)

Non-governmental actors have also developed initiatives aimed at developing ethical principles for dealing with AI. In this respect, the Institute of Electrical and Electronics Engineers (IEEE) deserves special interest. The IEEE is a New York-based, global professional association of engineers, predominantly in the fields of electrical engineering and information technology. The organisation's goals include the standardisation of techniques, hardware and software.[188]

Guidelines on the use of AI have also been developed under the umbrella of this organisation.[189] Eight principles form their core: respect for human rights; enhancement of human well-being; preservation of data agency; demonstrable effectiveness, transparency, accountability; awareness of misuse potential; and competence.[190]

The Guidelines cannot be examined in any detail here, embedded as they are in a relatively profound discussion of ethical issues. Only the recommendations made with regard to respect for human rights will be presented here: It is recommended that "[g]overnance frameworks, including standards and regulatory bodies" be established "to oversee processes ensuring that the use of A/IS [autonomous and intelligent systems] does not infringe upon human rights, freedoms, dignity, and privacy, and of traceability". There is also a need to "translate existing and forthcoming legal obligations into informed policy and technical considerations". Autonomous and intelligent systems should always be subordinate to human judgement and control. Finally, "[f]or the foreseeable future A/IS should not be granted rights and privileges equal to human rights".[191] In all of this, it is made clear that the human rights legal framework is "the floor and not the ceiling" for the standards to which those who create autonomous and intelligent systems must adhere.[192]

[188] https://www.ieee.org.

[189] *Möslein*, RDi 2020, 34, points out that such guidelines can find their way into the applicable law via the general clauses of civil law (Secs. 138, 242, 826 BGB).

[190] IEEE, Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System, 1st ed., 2019, p. 4: "A/IS creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation".

[191] IEEE, Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System, 1st ed., 2019, pp. 19 f.

[192] IEEE, Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System, 1st ed., 2019, p. 78.

The consequences of the increased use of autonomous and intelligent systems for the relationship between employers and employees, as well as labour relations in general, are explicitly, albeit briefly, addressed in these Guidelines: “The impact of A/IS on the workplace and the changing power relationships between workers and employers requires ethical guidance. Issues of data protection and privacy via big data in combination with the use of autonomous systems by employers are increasing, where decisions made via aggregate algorithms directly impact employment prospects. The uncritical use of A/IS in the workplace, and its impact on employee-employer relations, is of utmost concern due to the high chance of error and biased outcome”.[193]

Furthermore, in a section that is particularly worth reading, the authors emphasise the importance of the willingness and ability to empathise with others: "Collaboration requires enough commonality of collaborating intelligences to create empathy – the capacity to model the other’s goals based on one’s own". At the same time, however, the importance of autonomy is underlined: "According to scientists within several fields, autonomy is a psychological need. Without it, humans fail to thrive, create, and innovate. Ethically aligned design should support, not hinder, human autonomy or its expression".[194] Relatively concrete recommendations are derived from these insights:
“It is important that human workers’ interaction with other workers not always be intermediated by affective systems (or other technology) which may filter out autonomy, innovation, and communication. Human points of contact should remain available to customers and other organizations when using A/IS. Affective systems should be designed to support human autonomy, sense of competence, and meaningful relationships as these are necessary to support a flourishing life. Even where A/IS are less expensive, more predictable, and easier to control than human employees, a core network of human employees should be maintained at every level of decision-making in order to ensure preservation of human autonomy, communication, and innovation. Management and organizational theorists should consider appropriate use of affective and autonomous systems to enhance their

[193] IEEE, Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System, 1st ed., 2019, p. 47. In this respect, an orientation towards the EU's "concept of responsible research and innovation (RRI)" is explicitly recommended; cf. for example *Lindner/Goos/Güth/Som/Schröde*, "Responsible Research and Innovation als Ansatz für die Forschungs-, Technologie- und Innovationspolitik – Hintergründe und Entwicklungen, Büro für Technikfolgen Abschätzung im Deutschen Bundestag, Hintergrundpaper No. 22, 2016.
[194] IEEE, Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System, 1st ed., 2019, p. 102.

business models and the efficacy of their workforce within the limits of the preservation of human autonomy".[195]

As worthy of note as these statements are, it is at the same time striking that the interests of workers are otherwise only marginally illuminated: At one point the report states: "Employees should be empowered and raise ethical concerns in day-to-day professional practice".[196] Another passage addresses the change in work tasks and forms of employment: "For example, rather than carrying out a task themselves, workers will need to shift to supervision of robots performing that task". Other concerns relate to the change in "traditional employment structures, with an increase in flexible, contract-based temporary jobs without employee protection, and a shift in task composition away from routine/repetitive and towards complex decision-making". In this context, the authors call for two things: opportunities for further education and retraining[197] and that in future not only unemployment be measured, but above all the extent of underemployment be determined.[198]

[195] Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System, 1st ed., 2019, p. 102; cf. also *Yeung*, A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework, Council of Europe, DGI(2019)05, Committee of experts on human rights dimensions of automated data processing and different forms of artificial intelligence (MSI-AUT), 2018, p. 37: "In addition to [...] concerns about the use of AI technologies to imitate human behaviour are diffuse but often deeply-felt anxieties that our collective life may become increasingly "dehumanised", as tasks previously performed by humans are automated. Many fear that values and qualities that we cherish, including the value of real human interaction, of genuine empathy, compassion and concern, may be replaced by the relentless efficiency and consistency of AI driven services".

[196] IEEE, Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System, 1st ed., 2019, p. 132. See also *Friedmann*, Ethical concerns with replacing human relations with humanoid robots: an ubuntu perspective. https://doi.org/10.1007/ s43681-022-00186-0.

[197] IEEE, Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System, 1st ed., 2019, p. 132; cf. also ibid, p. 153.

[198] IEEE, Ethically Aligned Design, - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System 1st ed., 2019, p. 153.

VI. Interim result

At the international level, there are a number of different initiatives to regulate AI applications. Thus the above report is anything but conclusive. What seems particularly promising is the attempt currently being made within the framework of the Council of Europe to reach a (regional) agreement under international law for the development, design and application of AI. It should be borne in mind that the challenges of AI can hardly be met at the level of the nation-state alone. Rather, what is needed is an international, and preferably a global, regulatory strategy.[199]

The focus of the initiatives is on developing ethical standards for AI: data protection, accountability, safety, transparency and explainability, fairness and non-discrimination, human control over the technology, professional responsibility and promotion of human values.[200] One can make out a certain convergence of principles.[201] But in any case, the task remains to translate the existing concepts, which are by necessity relatively abstract, into concrete guidelines.[202]

[199] This is also the claim of *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 275.

[200] See *Fjeld/Achten/Hilligoss/Nagy/Srikumar*, Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI, Berkman Klein Center for Internet & Society at Harvard University, Research Publication No. 2020-1 based on a comprehensive review of supranational, governmental and private initiatives.

[201] See *Gasparotti*, Ethics Guidelines on Artificial Intelligence A comparison of EU and OECD guidelines, cepInput 07, 2019.

[202] See also *van Wynsberghe*, Artificial Intelligence: From ethics to policy, 2020.

E. European Union

AI has played a significant role at the level of the European Union for some time. The approaches developed here will be presented in the following. There is no need to expound on the significance of the EU for the situation in Germany. Suffice it here to point out that, for obvious reasons, only standards that apply throughout the Union have a chance of being taken into account globally.[203]

In its mid-term review of the Digital Single Market strategy, presented in May 2017, the Commission was already taking pains to stress the importance of building on Europe's scientific and industrial strengths, as well as its innovative start-ups, to become a leader in the development of AI technologies, AI platforms and AI applications.[204] In October 2017, then, the European Council stated that the EU urgently needed to respond to emerging trends such as AI "while at the same time ensuring a high level of data protection, digital rights and ethical standards", and called on the Commission to "put forward a European approach to artificial intelligence".[205]

I. The Communication on an Artificial Intelligence Initiative

Then, in 2018, the European Commission presented a European strategy on AI in a communication[206] which stated its aim to boost the EU's technological and industrial capacities and promote the spread of AI in both the private and public sectors through, among other things, investment in research, innovation and better

[203] See also *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 275: "In the globalised digital world, only harmonised regulatory standards across the EU will be able to set standards that are sufficiently effective in the long term".

[204] Communication from the Commissionto the European Parliament, the European Economic and Social Committee and the Committee of the Regions v. 10.5.2017, COM(2017) 228 final.

[205] European Council, Brussels, EuCO 14/17 of 19.10.2019, p. 7: http://data.consilium.europa.eu/doc/document/ST-14-2017-INIT/en/pdf.

[206] Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, Artificial Intelligence for Europe v. 25.4.2018, COM(2018) 237 final.

access to data. It made clear, however, that the same time the socio-economic changes brought about by AI should be accompanied by improvements such as modernising education and training systems, anticipating changes in the labour market, supporting labour market transitions and adapting social security systems. In all of this, the Commission was keen to ensure an appropriate ethical and legal framework based on the values of the Union and in line with the Charter of Fundamental Rights of the EU, and proposed a European AI Alliance for the development of AI ethics guidelines, which should in turn contribute to this same framework.[207] Interestingly, the alignment of the AI strategy with fundamental values such as data protection was seen by the European Commission from the outset as (among other things) a potential competitive advantage. The Communication expressly states: "The EU must […] ensure that AI is developed and applied in an appropriate framework which promotes innovation and respects the Union's values and fundamental rights as well as ethical principles such as accountability and transparency. The EU is also well placed to lead this debate on the global stage".[208]

In its Communication the Commission announced the development of "ethical guidelines for AI" involving "all relevant decision-makers". These should address "issues such as the future of work, fairness, safety, security, social inclusion and algorithmic transparency " and at the same time more generally "look at the impact on fundamental rights, including privacy, dignity, consumer protection and non-discrimination".[209] At the same time, it made clear that there are limits to possible "self-regulation".[210]

The need for an international orientation of the European AI strategy was also addressed by the Commission, with specific reference to the G7/G20, the United Nations and the OECD. The Commission considered that the EU, "based on its values and fundamental rights", could "make a unique contribution to the worldwide debate on AI".[211] The Communication was accompanied by two publications, a

[207] Communication, p. 3 f.
[208] Communication, p. 3.
[209] Communication, p. 18.
[210] Communication, p. 18.
[211] Communication, p. 22.

Commission report on security and liability issues[212] and a Communication on a European Data Strategy[213].

II. The Work of the Expert Group on Artificial Intelligence

1. Ethics guidelines

In June 2018, the European Commission appointed 52 experts from academia, civil society and industry to a High Level Expert Group on AI.[214] The "AI HLEG" presented a first draft of guidelines on the ethics of implementing AI in December 2018. This draft was revised following a consultation period, and amended guidelines published in April 2019.[215] In them the group elaborates four "ethical imperatives" that must be adhered to: respect for human autonomy; prevention of harm; fairness; and explainability.[216] The authors admonish AI practitioners to "[a]cknowledge and address the potential tensions between these principles".[217] The expert group formulates the following specific requirements: 1) human agency[218] and oversight,[219] 2) technical robustness[220] and safety, 3) privacy and data governance, 4) transparency,[221] 5) diversity,[222] non-discrimination and fairness, 6) societal and environmental well-being, and 7) accountability.[223] One demand made by the AI HLEG regarding transparency – under the aspect of "communication" – is that "AI systems should not represent themselves as humans to users". Humans, instead, "humans have the right to be informed that they are

[212] Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee on the security and liability implications of artificial intelligence, the internet of things and robotics v. 19.02.2020, COM(2020) 64 final.
[213] Communication from the Commission to the European Parliament, the Council and the European Economic and Social Committee - A European Data Strategy v. 19.2.2020, COM(2020) 66 final.
[214] This High Level Expert Group also formed the steering group for the European AI Alliance, a forum for broad public discussion of all aspects of AI development and its impact on the economy and society: https://ec.europa.eu/digital-single-market/en/european-ai-alliance.
[215] High Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 2019.
[216] Guidelines, p. 12.
[217] Guidelines, p. 13.
[218] Guidelines, p. 16: "The overall principle of user autonomy must be central to the system's functionality."
[219] Guidelines, p. 16: "Human oversight helps ensuring that an AI system does not undermine human autonomy or causes other adverse effects. Oversight may be achieved through governance mechanisms such as a human-in-the-loop (HITL), human-on-the-loop (HOTL), or human-in-command (HIC) approach".
[220] Guidelines, p. 16.
[221] Guidelines; p. 18.
[222] Guidelines, p. 18 ("consideration and involvement of all affected stakeholders throughout the process" and "ensuring equal access through inclusive design processes as well as equal treatment").
[223] Guidelines, p. 17 f.

interacting with an AI system. This entails that AI systems must be identifiable as such".[224]

To ensure the fulfilment of these requirements, the expert group essentially recommends focusing on technical methods: "architectures for trustworthy AI", conceptually integrated "ethics and rule of law (X-by-design)", "explanation methods", "testing and validating"[225] and "quality-of-service parameters".[226] Concerning possible architectures, the authors consider that "[r]equirements for Trustworthy AI should be 'translated' into procedures and/or constraints on procedures, which should be anchored in the AI system's architecture".[227] The group's take on requirements of "ethics and rule of law by design" is that companies are responsible for "identifying the impact of their AI systems from the very start, as well as the norms their AI system ought to comply with to avert negative impacts".[228] In the context of "explanation methods ", the AI HLEG states, " we must be able to understand why [the system] behaved a certain way and why it provided a given interpretation", given that "sometimes small changes in data values might result in dramatic changes in interpretation".[229] The enumerated "non-technical methods" include: regulation, codes of conduct, standardisation, certification, "accountability via governance frameworks", "education and awareness to foster an ethical mind-set", "stakeholder participation and social dialogue", and "diversity and inclusive design teams". Regarding "accountability" specifically, the AI HLEG concludes that companies "should set up governance frameworks, both internal and external, ensuring accountability for the ethical dimensions of decisions associated with the development, deployment and use of AI systems". This could "include the appointment of a person in charge of ethics issues relating to AI systems, or an internal/external ethics panel or board ". [230]

There is no specific consideration of workers' interests in the guidelines. In fact, workers are only mentioned in passing, as when "asymmetries of power or information, such as between employers and workers" are acknowledged[231] and reference is made to "potentially vulnerable persons and groups, such as workers,

[224] Guidelines; p. 18.
[225] See Guidelines, p. 21-22. In this respect, it is primarily a matter of using "sufficiently realistic data" and monitoring "throughout the entire life cycle".
[226] Cf. Guidelines, p. 22 (definition of "appropriate quality of service indicators").
[227] Guidelines, p. 21.
[228] Guidelines, p. 21.
[229] Guidelines, p. 21 with the example of confusing a school bus with an ostrich.
[230] Guidelines, p. 23.
[231] Guidelines, p. 2 (at footnote 2) with a reference to "articles 24 to 27 of the Charter of Fundamental Rights of the EU (EU Charter) dealing with the rights of the child and the elderly, the integration of persons with disabilities and workers' rights".

women, persons with disabilities, ethnic minorities, children, consumers or others at risk of exclusion".[232] There is also the requirement that AI systems "should support humans in the working environment, and aim for the creation of meaningful work".[233] In relation to the possibility to contest and seek redress against decisions made by AI systems, reference is made to the "right of association and to join a trade union in a working environment, as provided for by Article 12 of the EU Charter of Fundamental Rights".[234] Under requirement 5, diversity, non-discrimination and fairness, in the context of "stakeholder participation", the AI HLEG advises: “In order to develop AI systems that are trustworthy, it is advisable to consult stakeholders who may directly or indirectly be affected by the system throughout its life cycle. It is beneficial to solicit regular feedback even after deployment and set up longer term mechanisms for stakeholder participation, for example by ensuring workers information, consultation and participation throughout the whole process of implementing AI systems at organisations”.[235] Finally, it points out that AI "could help governments, unions and industry with planning the (re)skilling of workers [and] could also give citizens who may fear redundancy a path of development into a new role".[236]

In April 2019, the European Commission adopted a Communication in which it explicitly welcomed the seven core demands of the AI HLEG (prioritising human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental well-being; and accountability).[237]

2. Policy recommendations

Also in 2019, the policy recommendations of the AI High-Level Expert Group were published.[238] Here the group speaks, among other things, of the need to "enable workers made redundant or faced with the threat of redundancy due to automation and increased AI take-up" to "seek new forms of employment as the structure of

[232] Guidelines, p. 11.
[233] Guidelines, p. 12.
[234] Guidelines, p. 15 (at footnote 32).
[235] Guidelines, p. 19.
[236] Guidelines, p. 33.
[237] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Building trust in Human Centric Artificial Intelligence of 8 Apr 2019, COM(2019) 168 final.
[238] High-Level Expert Group on Artificial Intelligence, Policy and Investment Recommendations for Trustworthy AI, 2019.

the labour market is reshaped in response to the turn to increased reliance on digital services and processes".[239]

Among other things, the recommendations of the expert group – complementing the ethics guidelines developed by the group – focus on the legal framework for AI.[240] The group advocates for a "risk-based and multi-stakeholder approach". The paper states: “The character, intensity and timing of regulatory intervention should be a function of the type of risk created by an AI system. In line with an approach based on the proportionality and precautionary principle, various risk classes should be distinguished as not all risks are equal. The higher the impact and/or probability of an AI-created risk, the stronger the appropriate regulatory response should be. 'Risk' for this purpose is broadly defined to encompass adverse impacts of all kinds, both individual and societal. For specific AI applications that generate 'unacceptable' risks or pose threats of harm that are substantial, a precautionary principle-based approach should be adopted instead”.[241] Furthermore, the AI HLEG suggests an evaluation and, if necessary, a modification of the current EU law,[242] without, however, addressing labour (protection) law regulations.

Fundamental to the recommendations is a so-called human-centric approach. In this sense, the AI HLEG also calls for “a process of representation, consultation and, where possible, co-creation, where workers are involved in the discussion around AI production, deployment or procurement process in order to ensure that the systems are usable and that the worker still has sufficient autonomy and control, fulfilment and job satisfaction. This implies informing and consulting workers when developing or deploying AI, as set out in the existing texts adopted by the European institutions and the social partners”.[243] The recommendations continue: “Workers (not only employees but also independent contractors) should be involved in discussions around the development, deployment or procurement of algorithmic scheduling and work distribution systems, to ensure compliance with health and safety legislation, data policy, working time legislation and work-life balance legislation. Social dialogue plays a key role to enable this”.[244]

[239] Recommendations, p. 36.
[240] For example, Recommendations, p. 37: "This section complements the Guidelines by providing guidance on appropriate governance and regulatory approaches beyond voluntary guidance".
[241] Recommendations, p. 37 f. ; fn. 53 goes on to state (with reference to Council of Europe, Revised draft study of the implications of advanced digital technologies (including AI systems for the concept of responsibility within a human rights framework, 2019): “This includes not only tangible risks to human health or the environment, but also includes intangible risks to fundamental rights, democracy and the rule of law, and other potential threats to the cultural and socio-technical foundations of democratic, rights-respecting, societies.”
[242] Recommendations, p. 38 f.
[243] Recommendations, p. 13.
[244] Ibid.

3. White Paper on artificial intelligence

Building on the work of the AI HLEG, the European Commission in February 2020 presented a White Paper on AI to initiate a broad public consultation.[245]

a) Basic contents

In it, the Commission once again forcefully describes the risks associated with the use of AI.[246] On the issue of non-discrimination, for example, the Communication states: “Bias and discrimination are inherent risks of any societal or economic activity. Human decision-making is not immune to mistakes and biases. However, the same bias when present in AI could have a much larger effect, affectingand discriminating many people without the social control mechanisms that govern human behaviour. This can also happen when the AI system ‘learns’ while in operation. In such cases, where the outcome could not have been prevented or anticipated at the design phase, the risks will not stem from a flaw in the original design of the system but rather from the practical impacts of the correlations or patterns that the system identifies in a large dataset”.[247]

At the same time, the Commission identifies specificities of AI systems that make the enforcement of fundamental rights more difficult: “The specific characteristics of many AI technologies, including opacity ('black box-effect'), complexity, unpredictability and partially autonomous behaviour, may make it hard to verify compliance with, and may hamper the effective enforcement of, rules of existing EU law meant to protect fundamental rights. Enforcement authorities and affected persons might lack the means to verify how a given decision made with the involvement of AI was taken and, therefore, whether the relevant rules were

[245] White Paper on Artificial Intelligence – A European Approach to Excellence and Trust, COM(2020) 65 final of 19 Feb 2020; cf also *Unger*, ZRP 2020, 234; *Jüngling*, MMR 2020, 440; *Gasparotti/Harta*, European Strategy on Artificial Intelligence An Assessment of the EU Commission's Draft White Paper on AI, 2020. On Basic Questions of Regulation at National and European Level *Hacker*, NJW 2020, 2142.

[246] Elsewhere, the Commission explicitly recognises that "workers and employers are directly affected by the design and use of AI systems in the workplace." The involvement of the social partners will therefore "be a crucial factor in ensuring a human-centred approach to AI at work"; White Paper, p. 7.

[247] White Paper, p. 11 f.

respected. Individuals and legal entities may face difficulties with effective access to justice in situations where such decisions may negatively affect them".[248]

Like the AI HLEG, the Commission advocates a "risk-based approach". This is "important to help ensure that the regulatory intervention is proportionate". However, there is a need for "clear criteria to differentiate between the different AI applications, in particular in relation to the question whether or not they are 'high-risk'".[249] The Commission advocates a two-step assessment. First, it should be determined whether AI is used in an area in which significant risks are to be expected due to the nature of the activities typically undertaken.[250] The second criterion is whether the "AI application in the sector in question is, in addition, used in such a manner that significant risks are likely to arise". The latter reflects the understanding that "not every use of AI in the selected sectors necessarily involves significant risks".[251] However, there could be "exceptional instances where, due to the risks at stake, the use of AI applications for certain purposes is to be considered as high-risk as such – that is, irrespective of the sector concerned and where the below requirements below would still apply". In this respect, the Commission specifically mentions the area of anti-discrimination law: "In light of its significance for individuals and of the EU acquis addressing employment equality, the use of AI applications for recruitment processes as well as in situations impacting workers' rights would always be considered "high-risk" and therefore the below requirements would at all times apply. Further specific applications affecting consumer rights could be considered".[252]

However, the "risk-based approach" favoured by the Commission proved to be controversial from the beginning. Some criticise the definition of "high-risk" as not clear enough to create legal certainty. Clarification is required as to when the risks associated with the use of an AI application are to be considered "significant".[253] Others question whether a meaningful distinction can be made between low-risk and high-risk applications and suggest instead that a risk management approach

[248] White Paper, p. 12.
[249] White Paper, p. 17.
[250] White Paper, p. 17.
[251] White Paper, p. 18.
[252] White Paper, p. 18. Another example cited is the use of AI applications for the purposes of remote biometric identification.
[253] Cf. EU White Paper on Artificial Intelligence, cepAnalysis No. 4/2020, p. 4.

should be adopted, whereby the party best positioned to control or mitigate the risks would be deemed legally responsible.[254]

b) Results of the consultation

The consultation opened by the Commission with the White Paper, which ran from 19 February to 14 June 2020, drew a large number of comments. Overall, the need for action was almost universally affirmed. A large majority of respondents felt that there were gaps in the legislation or that new legislation was needed.[255] As a result, the Commission announced plans for a regulatory proposal.[256]

III. The Proposal of an "AI Law"

In April 2021, the European Commission presented a proposal for a regulation establishing harmonised rules on AI.[257] This law on artificial intelligence is intended to set standards worldwide. Meanwhile, the regulatory debate has not stood still. It has gained momentum in the USA, for example.[258] In the Commission's draft regulation, the term "artificial intelligence" is defined broadly, too broadly in the opinion of many critics.[259] The draft regulation does not venture outside the classic regulatory structures of product safety law, which means that essentially, technical

[254] Cf *Bertoloni*, Artificial Intelligence and Civil Liability - Study requested by the JURI Committee, 2020, p. 99 ff.
[255] Cf COM(2021) 206 final, p. 7 f.
[256] In addition, on 9 Mar 2021, the Commission presented a Communication entitled "2030 Digital Compass": the European way for the Digital Decade", which highlights inter alia the benefits of AI for manufacturing workers: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, COM(2021) 118 final, p. 9: " Manufacturing: thanks to 5G connectivity, devices in factories will be even more connected and collect industrial data. Artificial Intelligence will instruct robots in real time, making them increasingly collaborative, improving workers' jobs, safety, productivity and wellbeing. Manufacturers will be able to enhance predictive maintenance and produce on demand, based on consumers' needs, with zero stocks, thanks to digital twins, new materials and 3D printing."
[257] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Brussels, 21 Apr 2021, COM(2021) 206 final.
[258] For example, the Algorithmic Accountability Act of 2022 was recently introduced in the US Senate: https://www.wyden.senate.gov/news/press-release.
[259] See only *Bomhard/Merkle*, RDi 2021, 276 (277); *Ebers/Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (529); cf. *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung*, How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 55, who even argue for an extension of the regulation beyond AI.

standards, certifications and risk and quality management systems are the tools relied upon to contain the risks posed by AI systems.[260]

1. Overview of the draft regulation

a) Regulation of so-called high-risk AI systems

The draft focuses on so-called high-risk AI systems. Which systems are considered high-risk AI systems is determined by Article 6(2) of the Draft Regulation in conjunction with Annex III. In the present context, Annex III No. 4, which deals with "employment, workers management and access to self-employment", is of particular interest: According to No. 4(a) AI systems " intended to be used for recruitment or selection of natural persons, notably for advertising vacancies, screening or filtering applications, evaluating candidates in the course of interviews or tests " are high-risk systems in the sense of Article 6(2) of the Draft Regulation. According to No. 4(b), such high-risk AI systems also include those "intended to be used for making decisions on promotion and termination of work-related contractual relationships, for task allocation and for monitoring and evaluating performance and behavior of persons in such relationships". The Commission argues that such systems "may appreciably impact future career prospects and livelihoods of these persons". In the recruitment process and in the assessment, promotion or retention of people in work-related contractual relationships[261] such systems could "perpetuate historical patterns of discrimination, for example against women, certain age groups, persons with disabilities, or persons of certain racial or ethnic origins or sexual orientation". AI systems used to monitor the performance and behaviour of these individuals could "also impact their rights to data protection and privacy".[262]

[260] Cf. only *Spindler*, CR 2021, 361 (361); see also *Wiebe*, BB 2022, 899. An overview of current Union law on "AI in the workplace" in *Adams-Prassl*, Regulating algorithms at work: Lessons for a ‘European approach to artificial intelligence’, ELLJ 2022, 30; see also *De Stefano/Wouters*, AI and digital tools in workplace management and evaluation – An assessment of the EU’s legal framework, May 2022, p. 35ff.

[261] These should include not only employment relationships, but also those contractual relationships that relate to "services through platforms"; see Recital 36.

[262] Recital 36.

b) Requirements for high-risk AI systems

Chapter 2 of Title III of the Regulation contains a series of requirements that such systems must meet. These include an obligation to establish a risk management system (Art. 9(1) Draft Regulation). AI systems that use techniques in which models are trained with data must be developed on the basis of training, validation and testing data sets that meet quality criteria detailed in Article 10 of the Draft Regulation.[263] The technical documentation for a high-risk AI system must be drawn up before the system is placed on the market or put into service and kept up to date (Art. 11(1) Draft Regulation). High-risk AI systems must be designed and developed so as to facilitate the automatic recording of events (logs) during the operation of the high-risk AI system (Art. 12(1), first sentence, Draft Regulation). High-risk AI systems must be designed and developed in such a way that their operation is sufficiently transparent to enable users to interpret and make appropriate use of the system output (Art. 13(1), first sentence, Draft Regulation).[264] Also, high-risk AI systems must be designed and developed in such a way that they can be effectively supervised by natural persons during the period the AI system is in operation, such as via appropriate human-machine interface tools (Art. 14(1) Draft Regulation). Finally, high-risk AI systems are to be designed and developed in such a way that they achieve an appropriate level of accuracy, robustness and cybersecurity with regard to their intended use and function consistently in this respect throughout their life cycle (Art. 15(1) Draft Regulation).[265] In addition, Chapter 3 contains a number of provisions that establish "horizontal" obligations, especially for providers of high-risk AI systems.[266] These include, in particular, the obligation to establish a quality management system that ensures compliance with the provisions of the Regulation (Article 17(1) Draft Regulation).[267] For users, the regulation contains only few obligations.[268]

[263] It is noteworthy that Art. 10(5) Draft Regulation is intended to facilitate the detection of bias with the aim of avoiding indirect discrimination; cf. on this *Veale/Zuiderveen Borgesius*, CRi 2021, 97 (103).

[264] Particularly with respect to Arts 12 and 13 of the Draft Regulation, *Roos/Weitz*, MMR 2021, 844 (847) call for the requirements to be concretised by the legislature.

[265] The regulations are described as "institutionalised proceduralisation" by *Valta/Vasel*, ZRP 2021, 142 (144).

[266] Cf. COM(2021) 206 final, p. 16.

[267] According to some critics, the regulation displays a number of substantive weaknesses. However, these will not be discussed in detail here; instead, cf *Ebers/Hoch/Rosenkranz/Ruschemeier/Steinrötter,* RDi 2021, 528 (533 ff.).

[268] On the – very limited – obligations of the users of AI systems, see Spindler, CR 2021, 361 (369); cf also Valta/Vasel, ZRP 2021, 142 (144): "Im Gewand der Produktregulierung versteckt und unvollkommen ausgeführt finden sich Anforderungen an die Nutzer künstlicher Intelligenz, an das erforderliche, aber nicht ohne Weiteres leistbare Maß

2. Criticism

The draft raises a number of questions which will be critically discussed below.[269] The proposal submitted by the Commission will be taken as a basis. It is true that changes are emerging in the legislative process, for example with regard to the definition of AI or the legal obligations of users. However, it currently seems uncertain whether and to what extent the respective demands will find their way into the legal text.

a) Goals

According to the Commission, the draft aims to achieve several objectives: to "ensure that AI systems placed on the Union market and used are safe and respect existing law on fundamental rights and Union values"; to "ensure legal certainty to facilitate investment and innovation in AI"; to "enhance governance and effective enforcement of existing law on fundamental rights and safety requirements applicable to AI systems"; and finally to "facilitate the development of a single market for lawful, safe and trustworthy AI applications and prevent market fragmentation".[270]

What is strange about this is that the draft repeatedly emphasises the reference to fundamental rights in the regulation[271] and especially that "The extent of the adverse impact caused by the AI system on the fundamental rights protected by the Charter is of particular relevance when classifying an AI system as high-risk", but the management of the corresponding risks is then seen as a largely technical and administrative matter to be ensured by the AI providers, for example by setting

menschlicher Kontrolle. Nur undeutlich und nicht konsistent abgesichert findet sich die Botschaft, dass die KI ein eng geführtes Werkzeug des Menschen bleiben soll. Ob diese Begrenzung im weltweiten Entwicklungs- und Verwertungswettbewerb und angesichts des demografischen Wandels durchgehalten werden kann, ist zweifelhaft" ("One finds – disguised as product regulation and imperfectly executed – requirements for the users of artificial intelligence, for the necessary but not easily achievable degree of human control. The message that AI should remain a strictly managed tool in human hands comes across only indistinctly and is not consistently reinforced. It is doubtful whether this limitation can be maintained in the global race for development and exploitation and in the face of demographic change.").

[269] See on the whole topic most recently also *De Stefano/Wouters*, AI and digital tools in workplace management and evaluation – An assessment of the EU's legal framework, May 2022, p. 63 with alternative courses of action.

[270] COM(2021) 206 final, p. 3.

[271] Cf. from the Explanatory Memorandum only p. 1 (CI "in accordance with the values, fundamental rights and principles of the Union"; "proposal is based on EU values and fundamental rights").

up suitable "risk management systems".[272] At least there is hope that the concept pursued in the draft regulation will be supplemented by civil liability rules in the foreseeable future.[273]

b) Legal basis

The Commission bases its draft "in particular" on the competence provision of Article 114 TFEU[274] and, insofar as the Regulation "contains specific provisions relating to the protection of individuals with regard to the processing of personal data,[275] also on Article 16 TFEU.[276] As regards Article 114 TFEU, which provides for the adoption of measures for the establishment and functioning of the internal market, the Commission points out that the "primary objective" of the Regulation is to "ensure the proper functioning of the internal market by setting harmonised rules in particular on the development, placing on the Union market and the use of products and services making use of AI technologies or provided as stand-alone AI systems". In the Commission's view, this means preventing fragmentation of the internal market and ensuring legal certainty for both providers and users of AI systems.[277]

That the legal basis named by the Commission is viable is often doubted in the literature.[278] Even more significant in the present context is that the primary link to Article 114 TFEU has given rise to the fear that the new regulation could set limits to the Member States' efforts to legally address the issue of AI (including the issue

[272] Cf. *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung*, How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 11; cf ibid, p. 10, with the criticism that according to the draft, fundamental rights ultimately represent only "balancing material", which is then set against opposing interests: "As presently drafted, the Regulation appears to treat fundamental rights as equivalent to mere interests. Each fundamental right engaged by this Proposal and the activities it enables is limited and made subject to a balancing process by the Proposal itself - at least to some degree - by virtue of (a) the inherent tension in this Proposal between the goals of promoting economic activity and innovation, with the protection of fundamental rights; and (b) the legal bases upon which it has been founded".

[273] *Ebers /Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (536) call for a "clarification to the effect that the AI Regulation does not affect claims for damages by persons who have been harmed by a breach of its regulatory requirements".

[274] Explanatory Memorandum of the Draft Regulation, p. 17.

[275] Explanatory Memorandum of the Draft Regulation, p. 17 f.

[276] See EDPB-EDPS, Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 18 June 2021, p. 2, which in this respect underlines the need for independent supervision of compliance with personal data processing requirements.

[277] Explanatory Memorandum of the Draft Regulation, p. 6.

[278] Cf only *Ebers /Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (529) with further references; critical also *Valta/Vasel*, ZRP 2021, 142 (143).

of labour law).[279] In this respect, it is argued that while the draft primarily targets high-risk AI systems, the scope of application is aimed at all AI systems. It is also argued that there is some reason to assume that the intended regulations (in the sense of complete harmonisation)[280] are to be understood as conclusive, which would then mean that Member State requirements for other than high-risk AI systems would be regarded as inadmissible.[281] Accordingly, it could not be ruled out that the regulation would lead to deregulation rather than "raising the regulatory bar".[282] As far as high-risk AI systems are concerned, however, Article 29(2) of the Draft Regulation (on the obligations of users of high-risk AI systems) states that the corresponding provisions of the Regulation "are without prejudice to other user obligations under Union or national law".[283] Beyond that, however, there is a lack of comparable provisions, with Recital 1 even explicitly emphasising that the "Regulation pursues a number of overriding reasons of public interest, such as a high level of protection of health, safety and fundamental rights, and it ensures the free movement of AI-based goods and services cross-border, thus preventing Member States from imposing restrictions on the development, marketing and use of AI systems, unless explicitly authorised by this Regulation". The inclusion of a provision in the Regulation comparable to Article 88 GDPR could remove doubts about the permissibility of Member State regulations.[284]

c) Choice of Instrument

Also with regard to the choice of a regulation as legal instrument, the Commission refers to the "need for uniform application of the new rules" and to the fact that the direct applicability of a Regulation under Article 288 TFEU reduces legal fragmentation and facilitates the development of an internal market for legitimate, secure and trustworthy AI systems.[285] In doing so, it additionally argues that the

[279] It should be noted that Art. 114 in para. 4 and 5 does contain derogation possibilities, inter alia, for the protection of the "working environment", but these are to be interpreted narrowly; cf. only *Tietje*, in: Grabitz/Hilf/Nessesheim, Das Recht der Europäischen Union 2021, Art. 114 TFEU, para. 156 et seq. In this context, measures taken by a Member State pursuant to Art. 114 (4) TFEU must comply with the principles of proportionality and non-discrimination; cf. ibid., marginal no. 177.

[280] In general, for example, *Tietje,* in: Grabitz/Hilf/Nessesheim, Das Recht der Europäischen Union 2021, Art. 114 TFEU marginal no. 38, according to which "a differentiation of different harmonisation methods only makes sense when using directives as harmonisation instruments" ("eine Differenzierung unterschiedlicher Harmonisierungsmethoden [ist] nur beim Einsatz von Richtlinien als Harmonisierungsinstrumentarium sinnvoll". As far as regulations are used, there is "regularly no more room for manoeuvre for the Member States" ("regelmäßig keine Handlungsfreiräume der Mitgliedstaaten mehr").

[281] Cf *Veale/Zuiderveen Borgesius*, CRi 2021, 97 (108 ff.).

[282] Cf *Veale/Zuiderveen Borgesius*, CRi 2021, 97 (112).

[283] See also *Veale/Zuiderveen Borgesius*, CRi 2021, 97 (110).

[284] *Kelly-Lyth*, The AI Act and Algorithmic Management, Comparative Labor Law & Policy Journal, Dispatch No. 39, p. 1 (9).

[285] Explanatory Memorandum of the Draft Regulation, p. 7.

provisions of the Regulation are "not overly prescriptive" and leave "room for different levels of Member State action for elements that do not undermine the objectives of the initiative, in particular the internal organisation of the market surveillance system and the uptake of measures to foster innovation".[286] In the present context, it is of decisive importance, as already mentioned, whether and to what extent the Regulation should have a blocking effect vis-à-vis Member State law, that is, whether the latter should be allowed to deviate from the substantive requirements contained therein.[287]

It remains to be seen whether the expectations placed by the Commission in the choice of the legal instrument of the Regulation will be fulfilled in practice. After all, the main burden of authority is borne by the authorities of the Member States, and the experience gained in implementing the GDPR suggests that the fear of inconsistent implementation is not unfounded.[288]

d) Relationship of the Regulation to the GDPR

The relationship of the regulation to the GDPR does not appear to be completely clarified, to put it mildly. The explanatory memorandum to the draft states that the latter remains "unaffected" by the Regulation.[289] However, in their Joint Opinion on the draft Regulation, the European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS), for example, strongly suggest clarifying that the GDPR applies to any processing of personal data that falls within the scope of the proposal.[290] Furthermore, a specific recommendation is to include

[286] Explanatory Memorandum of the Draft Regulation, p. 7.

[287] Thus explicitly *Ebers/Hoch/Rosenkranz/Ruschemeier/Steinrötter* , RDi 2021, 528 (529) pointing out that so far it is unclear whether the Member States are to have the power to prohibit certain AI systems or their uses beyond the prohibitions of Art. 5 of the Draft Regulation.

[288] See only *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung*, How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 46 f.; cf also EDPB-EDPS, Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 18 June 2021, p. 15, calling for national data protection authorities to be designated as competent authorities in this respect.

[289] Explanatory Memorandum of the Draft Regulation, p. 4.

[290] EDPB-EDPS, Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 18 June 2021, p. 8; cf on the whole also Ebers/Hoch/Rosseran/Ruschemeier/Steinrötter, RDi 2021, 528 (536); *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung,* How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 42: "Given the reliance of many AI systems on personal data, it is recommended to strengthen the ties between the Proposal and the GDPR more consistently, in order to ensure a more coherent and comprehensive data protection framework for AI systems".

in Chapter 2 of Title III of the Regulation the requirement for AI systems to ensure compliance with the GDPR and the EUDPR.[291]

e) Risk-based approach

In the Regulation the Commission follows a so-called risk-based regulatory approach. This means that the intensity of the regulation depends on the risks posed by an AI application. Three risk categories are distinguished: "unacceptable risk", "high risk" and "low or minimal risk". [292]

Regardless of how one feels about the Commission's risk-based approach[293], one would have liked the risks posed by AI to have been reflected more thoroughly in the draft. For example, the explanatory memorandum to the draft only states that "depending on the circumstances regarding its specific application and use, artificial intelligence may entail risks and cause harm to public interests and rights that are protected by Union law" and that such harm can be "material or immaterial". In contrast, the Joint Opinion of the European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS) states that "generating content, making predictions or taking decisions in an automated way, as AI systems do using machine learning techniques or logical and probabilistic inference rules, is not the same as humans performing these activities using creative or theoretical reasoning and bearing full responsibility for the consequences".[294] This statement is not especially profound either, but it seems much more "insightful" than the Commission's remarks.

[291] EDPB-EDPS, Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 18 June 2021, p. 10; see also *Burgess*, How GDPR Is Failing – The world-leading data law changed how companies work. But four years on, there's a lag on cleaning up Big Tech, May 23, 2022. https://www.wired.com/story/gdpr-2022/.

[292] Explanatory Memorandum of the Draft Regulation, p. 12.

[293] Extremely critical for example *Edwards*, Regulating AI in Europe: four problems and four solutions, March 2022, p. 11: "The alleged 'risk-based' nature of the Act is illusory and arbitrary. Impacts on groups and on society as a whole need to be considered, as well as risks to individuals and their rights, and risks should be considered throughout the AI lifecycle not just at market entry."

[294] EDPB-EDPS, Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 18 June 2021, p. 5.

aa) Unacceptable risk

The first category includes applications that are considered unacceptable because they violate Union values, such as fundamental rights.[295] Article 5 of the Draft Regulation contains a comprehensive prohibition. However, this raises concerns. One objection that has been expressed is that the provision is too closely oriented to current debates (e.g. about social scoring), is already recognisably too narrow[296] and is also not "open to the future". [297]

bb) High risk

As mentioned, the rules of the Regulation primarily target AI systems that pose a high risk to the health and safety or fundamental rights of natural persons. Such high-risk AI systems must meet certain mandatory requirements; a conformity assessment must also be carried out in advance.[298] The classification as a high-risk AI system "is based on the intended purpose of the AI system, in line with existing product safety legislation". This means that it depends not only on the function of this system, but also on its specific purpose and application modalities.[299] Chapter 1 of Title III defines the two main categories for high-risk AI systems: According to Article 6(1) of the Draft Regulation, systems are considered high-risk AI systems if they fulfil the conditions listed there. In addition, the AI systems listed in Annex III are also considered high-risk AI systems. As already mentioned, the second category (Article 6(2) in conjunction with Annex III) is of particular interest here. This category refers to stand-alone AI systems that pose a high risk to the health and safety or fundamental rights of natural persons (Article 7 of the Draft Regulation). Annex III of the Draft Regulation currently lists eight stand-alone AI systems for which it has already been shown or is foreseeable that

[295] The prohibitions apply "to practices that have a significant potential to manipulate persons through subliminal techniques beyond their consciousness or exploit vulnerabilities of specific vulnerable groups [...] in order to materially distort their behaviour in a manner that is likely to cause them or another person psychological or physical harm"; Explanatory Memorandum of the Draft Regulation, p. 12 f.

[296] Cf only *Ebers Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (530).

[297] In fact, a rule comparable to that in Art. 7 of the Draft Regulation is missing here; thus rightly critical *Ebers/Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (531); also criticical *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung*, How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 20 f.

[298] Recently, in anticipation of the entry into force of the Regulation, researchers have presented a procedure that providers can use in their conformity assessment; cf *Floridi/Holweg/Taddeo/Silva/Mökander/Yuni,* capAI - A Procedure for Conducting Conformity Assessment of AI Systems in Line with the EU Artificial Intelligence Act, 23 Mar 2022: https://ssrn.com/abstract=4064091.

[299] Explanatory Memorandum of the Draft Regulation, p. 13.

these risks will actually materialise. As already mentioned, these include AI systems from the areas of "employment, workers management and access to self-employment". According to Article 7(1) of the Draft Regulation, the Commission is empowered to adopt delegated acts to amend the list in Annex III in order to add high-risk AI systems that fulfil the two conditions listed there.[300] In this respect, it can be criticised that the qualification as a high-risk CI system appears uncertain, as it depends on the fulfilment of characteristics that require much fleshing out.[301] Quite apart from this, the approach chosen in Article 7(1) of the Draft Regulation of a supplementing of the list by the Commission may serve legal certainty, but appears to be relatively narrow, since adding to the list is only permissible within the framework of the areas named in Annex III (Article 7(1) Draft Regulation), and beyond that it is abundantly "technocratic" and therefore also not very "fundamental rights-friendly". There is also widespread concern that the approach chosen by the Commission could open up the possibility of circumvention for providers.[302] Finally, there is criticism of the Commission's prominent position, which is not even obliged to hold consultations.[303]

cc) Low risk

AI systems of the third category, those that pose only a "low or minimal risk", remain largely unregulated. In this respect, according to Article 52 of the Draft Regulation, the only obligations on "AI systems that are intended for interaction with natural persons" involve transparency.[304] This seems too narrow in two respects, since on the one hand, other systems can also pose risks to fundamental rights, and on the other hand, it is not entirely clear why the scope of obligations is narrowed from the outset to transparency obligations.

[300] That the AI systems are intended to be used in any of the areas listed in points 1 to 8 of Annex III and that they pose a risk of harm to the health and safety, or a risk of adverse impact on fundamental rights (Art. 7(1)a) and (b) Draft Regulation).

[301] Cf in particular Art. 7(1)(b) of the Draft Regulation: "the AI systems pose a risk of harm to the health and safety, or a risk of adverse impact on fundamental rights, that is, in respect of its severity and probability of occurrence, equivalent to or greater than the risk of harm or of adverse impact posed by the high-risk AI systems already referred to Annex III".

[302] See *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung*, How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 13.

[303] Crit. e.g. *Ebers/Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (532); also *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung,* How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 49 et seq.

[304] Moreover, according to Art. 69(1) of the Draft Regulation, there is an obligation on the Commission and the Member States to promote and facilitate the establishment of codes of conduct "aimed at ensuring that the requirements set out in Chapter 2 of Title III apply to AI systems that do not pose a high risk".

f) Standardisation and self-assessment of providers

High-risk AI systems are at the centre of the draft. Thus the Commission relies on the idea of co-regulation through standardisation. The basis for this is the so-called New Legislative Framework (NLF). Characteristic of the NLF is that only the central product safety requirements are regulated by the EU legislature itself, but their concretisation[305] is then left to the European standardisation organisations.[306] In the main, the AI Regulation is based on a self-assessment by the providers on the basis of the harmonised technical standards,[307] whereby, according to Article 40 of the Draft Regulation, there is a presumption of conformity if a high-risk AI system complies with harmonised standards, the references of which have been published in the Official Journal of the EU. According to Aticles 19 and 43(2) of the Draft Regulation, providers must ensure that high-risk AI systems are subject to a conformity assessment procedure before being placed on the market or put into service. Only in exceptional cases does the Regulation provide for an ex-ante conformity assessment by external third parties.[308] However, Article 63 of the Draft Regulation stipulates ex-post market surveillance by the competent authorities of the Member States, while Article 64(1) and (2) of the Draft Regulation grants the market surveillance authorities unrestricted access to all information, documents and data (including the source code, if applicable).

How effective the regulation of high-risk AI systems will be depends primarily on the harmonised standards to be developed.[309] Adherence to these standards is voluntary for providers. Relying on their own technical solutions, however, would mean concretising the (vague) requirements of the regulation at their own risk. In contrast, providers “can’t go wrong" if they adhere to the standards to be developed, as they can then invoke the presumption of conformity in Article 40 of

[305] For more details, see *Ebers*, RDi 2021, 588 (589), who illustrates this interplay using the example of Article 10(3) of the Draft Regulation; criticism of the extensive vagueness of the content of the requirements of Article 10 of the Draft Regulation can also be found in *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung,* How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 33 et seq.

[306] European Committee for Standardisation (CEN), European Committee for Electrotechnical Standardisation (CENELEC), European Telecommunications Standards Institute (ETSI).

[307] See also *Ebers*, RDi 2021, 588 (589) and reference to Recital 21 of Decision No. 768/2008/EC of the European Parliament and of the Council of 9 July 2008 on a common framework for the marketing of products and repealing Council Decision 93/465/EEC: "The manufacturer, having detailed knowledge of the design and production process, is best placed to carry out the complete conformity assessment procedure. Conformity assessment should therefore remain the obligation of the manufacturer alone."

[308] More closely *Ebers*, RDi 2021, 588 (590).

[309] Cf. however Art. 41(1), first sentence, of the Draft Regulation, according to which the Commission may establish "common specifications for the requirements" if there are no harmonised standards according to Art. 40 or if the Commission "is of the opinion that the relevant harmonised standards are insufficient or that certain concerns with regard to safety or fundamental rights need to be addressed"; cf. on this *Bomhard/Merkle*, RDi 2021, 276 (283).

the AI Regulation.[310] Quite apart from the fact that standardisation in the field of AI is particularly challenging for practical reasons,[311] the fundamental approach of the NLF has been met with concern for some time, on the one hand from the point of view of democratic legitimacy and on the other from the point of view of the lack of judicial control.[312] One of the critics therefore judges as follows: “Such a regulatory approach [NLF] meets with considerable concerns. The standardisation of AI systems is not only about purely technical issues. Rather, a number of legal as well as ethical decisions have to be made, which require a society-wide, democratic discourse that should be shaped by industry, civil society, consumer associations and other stakeholders”.[313]

The conformity assessment by way of self-assessment is also subject to criticism.[314] It is not only the considerable leeway that arises under the regulation[315] and ultimately amounts to outsourcing the decision on which risks are acceptable to the AI provider that is disturbing.[316] This leaves virtually no room for an ex ante review by external third parties[317] and there is also no provision for the participation of potentially affected parties. As far as the first aspect is concerned, reference should be made to Article 64(5) of the Draft Regulation, which only opens the way to carrying out technical tests of the high-risk AI system if "the documentation referred to in paragraph 3 is insufficient to ascertain whether a breach of obligations under Union law intended to protect fundamental rights has occurred". On the other

[310] Also on this point *Ebers*, RDi 2021, 588 (591) with reference to *Veale/Zuiderveen Borgesius*, CRi 2021, 97 (105): "Consequently, standardisation is arguably where the real rule-making in the Draft AI Act will occur".

[311] In more detail *Ebers*, RDi 2021, 588 (593); cf. also *Philipp Roos/Caspar Alexander Weitz*, MMR 2021, 844 (851), calling "security requirements of the draft AI Reg partly too vague or difficult to implement in practice".

[312] *Ebers*, RDi 2021, 588 (593 ff.) with further references.

[313] *Ebers*, RDi 2021, 588 (596); crit. also *Ebers/Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (532); cf. also *Veale/Zuiderveen Borgesius*, CRi 2021, 97 (112): "The high-risk regime looks impressive at first glance. But scratching the surface finds arcane electrical standardisation bodies with no fundamental rights experience expected to write the real rules, which providers will quietly self-assess against".

[314] See only *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung*, How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 37: "The enforcement architecture relies heavily on (self-) conformity assessments. This leaves too much discretion to AI providers in assessing risks to fundamental rights without meaningful independent oversight, leaving many safety-critical and fundamental-rights critical AI applications without any ex ante review or systematic ex post review".

[315] Cf. only Art. 9 Draft Regulation, which refers to the "reasonably foreseeable misapplication" (para. 2(a)), demands "due consideration" (para. 3) and requires an assessment "as justifiable" (para. 4); criticised, for example, by *Ebers/Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (533).

[316] Thus *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung*, How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 30.

[317] Crit. also *Ebers /Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (533); *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung,* How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 39 f.; cf. also EDPB-EDPS, Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 18 June 2021, p. 39. also EDPB-EDPS, Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 18 June 2021, p. 12, calling for the establishment of a general ex-ante third-party conformity assessment for high-risk AI systems.

hand, as far as the latter aspect is concerned, reference should be made in particular to Article 9(4) of the Draft Regulation, according to which residual risks only have to be communicated to users, whereas according to Article 35(9) of the GDPR, as part of the data protection impact assessment "where appropriate, the controller shall seek the views of data subjects or their representatives on the intended processing, without prejudice to the protection of commercial or public interests or the security of processing operations".[318]

In general, however, it seems alarming that the Regulation focuses almost exclusively on the providers of AI systems. Thus, a risk assessment is to be carried out by the provider alone, although in most cases the users and not the providers will be the responsible for the processing. Also, the provider will often not be able to assess all uses of the AI system, so that the (initial) risk assessment to be carried out by the provider will almost necessarily be more general than that of a later user of the system.[319] According to Article 28(1)(c) of the Draft Regulation, users can become providers "if they make a significant change to the high-risk AI system". However, it is doubtful whether the legislature has thereby sufficiently taken into account the fact that AI is recognised and regulated less as a single product or service than as a process that goes through a life cycle of development, adaptation and use.[320]

g) Rights of the data subjects and legal protection

It is also unsatisfactory that the Regulation almost completely ignores the (potentially) affected parties. The transparency obligation under Article 13(1) of the Regulation thus only applies to the user,[321] without taking the data subject into account.[322] This is criticised not least by the European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS), who in their joint

[318] Crit. *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung*, How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 30.

[319] See EDPB-EDPS, Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 18 June 2021, p. 9.

[320] See *Circiumaru*, Three proposals to strengthen the EU Artificial Intelligence Act - Recommendations to improve the regulation of AI - in Europe and worldwide: https://www.adalovelaceinstitute.org/blog/three-proposals-strengthen-eu-artificial-intelligence-act/: "[...] an AI system used to manage labour and review performance may seem equitable in the abstract, but could become a vehicle for discrimination when deployed in the workplace".

[321] Art. 13 para. 1 sentence 1 Draft Regulation: "High-risk AI systems shall be designed and developed in such a way that their operation is sufficiently transparent to enable users to interpret and use the results of the system appropriately".

[322] For example, see *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung*, How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, p. 35.

opinion on the draft speak of a "blind spot" in the Commission's proposal.[323] This is in line with the fact that although the Regulation sets out a number of objective obligations (of providers and users) and provides for fines and other sanctions for infringements (cf. for example Art. 65(5) of the Regulation), it does not grant any claims for damages to those affected[324] and does not grant them any formal (participation) rights, for example within the framework of a complaints procedure.[325] This is all the more disconcerting, as the regulation precisely invokes the risks to fundamental rights emanating from AI.[326]

[323] EDPB-EDPS, Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 18 June 2021, p. 8 f.: "Whether they are end-users, simply data subjects or other persons concerned by the AI system, the absence of any reference in the text to the individual affected by the AI system appears as a blind spot in the Proposal. Indeed, the obligations imposed on actors vis-a-vis the affected persons should emanate more concretely from the protection of the individual and her or his rights. Thus, the EDPB and the EDPS urge the legislators to explicitly address in the Proposal the rights and remedies available to individuals subject to AI systems"; critical also *Kelly-Lyth*, The AI Act and Algorithmic Management, Comparative Labor Law & Policy Journal, Dispatch No. 39, p. 1 (8).

[324] Cf. for example *Ebers /Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (536) with the statement that "the weakly designed official enforcement [is] thus [...] not compensated by private enforcement".

[325] See only *Smuha/Ahmed-Rengers/Harkens/Li/MacLaren/Pisellif/Yeung*, How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, 2021, pp. 44 f., 51, who point out the differences to the GDPR that exist thereafter; cf. also *Veale/Zuiderveen Borgesius*, CRi 2021, 97 (111).

[326] Crit. also *Bomhard/Merkle*, RDi 2021, 276 (283) with the additional note that the AI Regulation also does not know a "consent concept" and that, in contrast to data protection, "which is based on a fundamental right to informational self-determination, [...] an individual "claim to trustworthy AI" cannot be derived from fundamental rights and freedoms without further ado".

F. Germany

Looking at Germany, there are three AI-related initiatives that deserve closer attention: the AI strategy of the German government, the work of the so-called Data Ethics Commission and the Enquete Commission set up by the German Bundestag to assess the social responsibility and economic, social and ecological potential of AI.[327] In addition, concrete legislative efforts have been made in the meantime, particularly in the form of the so-called *Betriebsrätemodernisierungsgesetz* (Workers Councils Modernisation Act), which came into force on 18 June 2021. However, this will be discussed in more detail later.[328]

I. AI Strategy of the Federal Government

In 2018, the Federal Government presented its "Artificial Intelligence (AI) Strategy". This strategy pursues a variety of goals. Among other things, it aims at a "responsible and public welfare-oriented development and use of AI". The federal government also expressly wants to "secure the possibilities of co-determination in companies in the introduction and use of AI" and "promote experimental spaces in companies for AI applications in the world of work".[329] Some of these ideas have meanwhile found expression in the just-mentioned Workers Council Moderniastion Act on the Modernisation of the Workplace[330] which will be presented in more detail below. After an interim report in 2019[331] the Federal Government published an "update" of its strategy in December 2020.[332]

[327] https://www.bundestag.de/ausschuesse/weitere_gremien/enquete_ki.

[328] Cf. G. VIII. 4.

[329] Artificial Intelligence Strategy of the Federal Government, November 2018, p. 7.

[330] Cf. G. VIII. 4.

[331] Interim Report One Year AI Strategy.

[332] Strategy for Artificial Intelligence of the Federal Government - Update 2020 of 8.12.2020, BT-Drucks. 19/25095, p. 2; cf. also ibid., p. 3, according to which the Federal Government is reacting to new developments with the update of the AI strategy and supplementing it with further measures. Current developments explicitly mentioned are "COVID-19 pandemic and sustainability issues, especially environmental and climate protection as well as European and international networking".

Within the framework of the AI strategy, a "Focus Group on Artificial Intelligence in the World of Work" was formed. This group advises the Federal Ministry of Labour and Social Affairs (BMAS) and is intended specifically to support the implementation of the AI strategy within the purview of the BMAS by identifying applications of AI in business practice and considering their impact on the world of work. The group consists of personalities from business, trade unions, associations and companies.[333] In addition, the so-called AI Observatory was established in 2020.[334] This pursues the goal of "analysing the effects on work and society associated with the application of AI and developing recommendations for action and measures to shape them". A further aim is to establish "European and international structures on the topic of AI in work and society".[335] The "ExamAI - AI Testing & Auditing" project of the Gesellschaft für Informatik (GI), which is funded within the framework of the AI Observatory, is investigating which control and testing procedures can be used to ensure safety, transparency, equal treatment and data protection when humans and machines work together.[336]

II. Data Ethics Committee

In the coalition agreement of the 19th legislative period of the German Bundestag between the CDU, the CSU and the SPD, which was negotiated after the 2017 Bundestag elections and concluded on 7 February 2018, the parties agreed to set up a data ethics commission and at the same time formulated their mandate to propose a "development framework for data policy, the handling of algorithms, artificial intelligence and digital innovations". The commission began its work in September 2018 and presented its report in October 2019.[337]

In it, the Commission recommended a "risk-adapted" approach to the regulation of algorithmic systems. This should be "based on the principle that an increasing potential for harm goes hand in hand with increasing requirements and depth of intervention of the regulatory instruments". In the context of risk assessment "the

[333] Strategy, p. 9.
[334] Cf. https://www.ki-observatorium.de/.
[335] Strategy, p. 10; cf. also: KI-Observatorium - Denkfabrik: Digitale Arbeitsgesellschaft: denkfabrik-bmas.de.
[336] Cf. https://testing-ai.gi.de/.
[337] Expert opinion of the Data Ethics Commission; cf. on this, for example, *Raue/von Ungern-Sternberg,* ZRP 2020, 49; also *Kelber*, ZD 2020, 73.

entire socio-technical system is important, that is, all components of an algorithmic application including all human actors, from the development phase (e.g. with regard to the training data used) to the implementation in an application environment and the phase of evaluation and correction". In this context, the Data Ethics Committee considered it "sensible to distinguish between five levels of criticality in a first step with regard to the damage potential of algorithmic systems", whereby the respective damage potential should be taken into account.[338]

In its final report, the Data Ethics Commission also commented on issues of worker protection and co-determination. It stated that "the sometimes far-reaching recording of employees' movement and performance data in modern working environments and the creation of biometric profiles necessary for certain forms of collaboration pose considerable risks to employees' informational self-determination and general personal rights".[339] In this respect, it recommended a further development of employee data protection in cooperation with the social partners. Here the concerns of persons "in unusual forms of employment" should also be taken into account. Collective agreements and company agreements should continue to play an important role in the area of employee data protection.[340] When structuring the co-determination rights of the representatives of interest groups over the processing of personal data in the company, "the existing asymmetry of knowledge between the employer and the employee side about the mode of operation and details of the processing procedures must be adequately taken into account". Therefore, "models must be found that enable the representatives to have recourse to external expertise beyond the existing mechanisms, while paying attention to the appropriate involvement of the company data protection officer, but also to the protection of business secrets". In view of the constant development of data processing systems in the company (software updates, self-learning elements, etc), "there should be a further development from selective consent to permanent monitoring of processes by the representatives of the respective interest groups".[341]

[338] Expert Opinion, p. 183.
[339] Expert Opinion, p. 112.
[340] Expert Opinion, p. 112.
[341] Expert Opinion, p. 113.

III. Enquete Commission

In this context, the work of the Enquete Commission "Artificial Intelligence - Social Responsibility and Economic, Social and Ecological Potentials", which was appointed by the German Bundestag in June 2018, is also worth mentioning. This commission consisted of 19 members of the German Bundestag and 19 experts.[342] According to the decision to set up the commission, its task was to "develop recommendations for action, including for the legislature, on how to foster the opportunities AI offers for people's lives, for the development of our prosperity and for society as a whole and how to limit the risks".[343] This was not least a matter of clarifying "whether and in what form national, European and international rules are needed", whereby "the framework conditions for AI should be defined on the basis of European values".[344] Among the concrete tasks were: to present the "opportunities and challenges of AI for the individual, society, the state, the economy and the world of work" and to research the "effects on equality and gender justice" as well as the "changes in the world of work through AI". A presentation of the "effects of technological change on the social market economy, collective bargaining and co-determination" was also explicitly addressed.[345]

The Commission presented its final report in October 2020.[346] It contains a large number of findings and recommendations that do not specifically concern working life but are certainly relevant to it. Suffice it to mention the section of the report dedicated to questions of discrimination,[347] or that dealing with "AI and law", where, among other things, questions of liability law are problematised.[348]

In addition, the final report also contains a separate section on the topic of "Artificial Intelligence and Work".[349] In this section, "examples of AI applications in use or being tested in companies" are presented: assistance and service robots, knowledge and assistance systems, process optimisation through predictive analysis, AI-based chatbots and intelligent speech analysis.[350] In addition, the

[342] On the establishment of the Commission, cf BT-Drucks. 19/2978 of 26 Jun 2018, p. 4.
[343] BT-Drucks. 19/2978, p. 1.
[344] BT-Drucks. 19/2978, p. 1.
[345] BT-Drucks. 19/2978, p. 2 f.
[346] Report of the Enquete Commission on Artificial Intelligence - Social Responsibility and Economic, Social and Environmental Potentials, BT-Drucks. 19/23700 of 28 Oct 2020.
[347] BT-Drucks. 19/237oo, p. 57 ff.
[348] BT-Drucks. 19/237oo, p. 67 ff.
[349] For the area of "work, education, research", as for other areas, a separate project group was formed; cf. BT-Drucks. 19/23700 of 28 Oct 2020 (under V.), p. 289 ff.
[350] BT-Drucks. 19/23700, p. 300 ff.

report contains recommendations for action.[351] These include, in particular, the demand that work be made more human-oriented, that the distribution of roles in human-machine interaction be clearly defined and that co-determination be "modernised". Especially with regard to the latter point, the report is quite specific: employees and their representatives should, among other things, “be able to participate just as effectively in the definition of the objectives and configuration of AI systems as in the evaluation, operation and further development of the socio-technical conditions of use [...];[352] due to the increasing importance of personnel planning and development as well as the qualification of employees, be given a right of co-determination and initiative in questions of further training; be able to use effective co-determination so that all rights of personality defined in the constitution are protected; be able to base their actions on a comprehensible technology assessment, quality criteria, certifications, audits and the work of the Federal Government's Observatory; have an influence on work density and the amount of work resulting from the machine-human interface; have easy access to further training and counselling offers in order to develop their own AI competence”.[353] The modernisation of co-determination must also take into account that "in addition to the employees in the company, external service providers are increasingly participating in value creation". In addition, "gaps in co-determination must be closed where AI systems with transnational accountability and transnational corporations are involved". At the same time, "the principles and contents of traditional works agreements based on Section 87(1), No. 6 of the Works Council Constitution Act (BetrVG) should be further developed or rethought". It then goes on to say that it is "a matter of strengthening process orientation and making it more agile, and of basing effects analysis and evaluations on standards and scientific findings". The Commission also suggests that "employer and works council conclude a principle-based framework agreement and application-specific individual agreements" in order to " speed up the approval process.[354]

[351] BT-Drucks. 19/23700, p. 328 ff.
[352] Whether this should apply generally to the use of data or only to the use of personal data was debated in the project group; BT-Drucks. 19/23700, p. 330.
[353] BT-Drucks. 19/23700, p. 321.
[354] BT-Drucks. 19/23700, p. 321.

The Commission also looks at other issues in detail. These include the issue of the use of automated decision-making systems and AI in human resource management. Again, the report contains relatively concrete recommendations. For example, it should be made clear by law "that people analytics processes may only be used if there is a company agreement or if employees have given their individual consent".[355] Before the introduction of ADM systems[356] and AI-supported systems, "company impact assessments regarding the potential for damage, personality rights and the effects on working conditions are indispensable".[357]

[355] BT-Drucks. 19/23700, p. 330.
[356] Algorithmic decision making.
[357] BT-Drucks. 19/23700, p. 331.

G. Problem areas under labour law

I. AI and the concept of the employee under German law

One question is clearly of fundamental significance: Does one contractual partner's use of AI have an influence on whether the other partner is to be qualified as an employee? It does not seem implausible that this question could be answered in the affirmative. After all, there is some evidence to suggest that the use of AI exacerbates the asymmetry of information between the parties involved and generally shifts the balance in favour of one side. However, nothing other than the "real imparity" existing between the parties to the employment contract and the resulting need for social protection of the employee forms the starting point for the emergence and development of labour law.[358]

1. Possible starting points

If one looks at the English-language literature on the subject, one indeed encounters authors keen to consider the possibilities opened up by the use of AI in the question of whether a certain contractual relationship is to be qualified as an employment relationship. Since the use of AI has dramatically concentrated the control exercised by the employer, but the qualification of contractual relationships is essentially based on precisely this, namely on control and/or dependence, there are arguments in favour of affirming the status of employee.[359] Nevertheless, caution is called for when using the term *control* as the basis for a definition, as this test, which is particularly widespread in the United Kingdom, is not, or at least not exclusively, about "control" in the sense of monitoring or supervision.[360] Rather, in answering the question of whether a person exercises control in a relevant way, in the United Kingdom, too, aspects such as "being subject to instruction" or

[358] Cf. only MünchArbR/Fischinger, 4th ed. 2018, § 3 marginal no. 29.

[359] See in particular *Prassl*, Comparative Labour Law & Policy Journal 2019, 123; most recently also *De Stefano/Durri/Stylogiannis/Wouters,* Platform work and the employment relationship, ILO Working Paper 27, March 2021, p. 34: "control through technology".

[360] *Prassl/Jones*, in: Waas/Heerma van Voss (eds.), Restatement of Labour Law in Europe, vol. 1, The Concept of Employee, 2017, p. 755 list as factors to be considered in this respect: "the extent of actual control over the substance of work done, the method of control, transfer of control, the extent of integrationinto the employing organisation, the powers of appointment, suspension and dismissal, and the existence and form of payment of wages, salaries and other benefits".

"integration" must be taken into account. The essential aspects of the term are thus: *work instructions*, *work control* and *integration*.[361] And in other legal systems, too, it seems to be the case that although aspects such as supervision and control are also taken into account in the qualification of a contractual relationship, these hardly ever stand on their own.[362] Even if this were not the case, the gain in knowledge for German law would remain manageable, as the concept of employee in foreign legal systems cannot be transferred one-to-one to the concept that exists in German law. With regard to the legal situation in Germany, it must also be taken into account that attempts made some time ago to bring a (pure) "informational dependency" into play in determining the status of employee were not successful.[363]

However, there is no reason to stop the considerations at this point. The fact that one should not be misled into jumping to conclusions about the status of an employee by a possibly more pronounced control due to the use of AI does not mean that the increasing availability and the ever more intensive use of AI cannot be brought into line with the characteristics that Section 611a of the German Civil Code (BGB) mentions for the existence of an employment contract or for the status of an employee: "being bound by instructions", "external determination and "personal dependence". It is therefore necessary to take a closer look at these characteristics.[364]

2. The features of Section 611a BGB

a) Bound by instructions

It should be made clear at the outset, however, that personal dependency is "merely a generic legal term which itself has no material content". It only becomes

[361] See *Prassl/Jones*, in: Waas/Heerma van Voss (eds.), Restatement of Labour Law in Europe, vol. 1, The Concept of Employee, 2017, p. 754.

[362] See for example *Hießl*, Case law on the classification of platform workers: Cross-European comparative analysis and tentative conclusions, p. 48 ff. with a discussion of the prerequisite "direction and control": https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3839603.

[363] *Linnenkohl/Kilz/Rauschenberg/Reh*, ArbuR 1991, 203, who at the time relied on the increasing integration of data processing in the production process and referred primarily to ISDN and DSL; see *Koch,* Selbstständigkeit in der virtualisierten Arbeitswelt, Diss. Kassel, 2010, p. 37.

[364] The approach of *Prassl*, The Concept of the Employer, 2015, on the other hand, is fundamentally different, starting - quite generally - with the *concept of the employer* instead of the concept of the employee, whereby his considerations are subject to the objection that it would have to be justified to what extent the employer position arises from certain employer functions, which the author names in detail.

practicable "by its degree of personal dependence [...], which is measured according to the criteria of being bound by instructions or external control".[365]

The first question that must be answered in this context is whether and to what extent the use of AI has an influence on the obligation by or subjection to instructions of the person at whom the use of the AI is aimed. However, certain doubts immediately arise as to whether one can really get anywhere with this question. For one can only either affirm or deny that there is a duty to follow instructions, so that an "intensification" that one might associate with the use of AI would not make any difference to begin with. It is true that there are gradations of the obligation to follow instructions in case law and literature. However, these generally refer to the type or the objects of instructions; for example, there is agreement that a lack of technical instructions can be compensated for by instructions in other respects (regarding location or time).

However, case law and literature repeatedly speak of the "degree"[366] or "extent" to which one can be subject to instructions,[367] but also of the characteristic of being obligated to follow instructions.[368] Following this – contrary to first impressions – there would certainly be indications that the use of AI should already be taken into account when the characteristic of being bound by instructions is given. This is because AI systems linked to the monitoring of employees and their results can react immediately to observed "misconduct" and, if so programmed, even automatically invite an employee to a staff appraisal interview. Furthermore, even if, for legal reasons, automatic machines are not able to make "independent" decisions,[369] they can at least prepare and thus accelerate these decisions. Under

[365] *Preis*, NZA 2018, 817 (819).

[366] ErfK/Preis, 22nd ed. 2022, § 611a BGB marginal no. 40: "Je stärker die Weisungsbindung, umso eher ist ein Arbeitsverhältnis anzunehmen“ (The stronger the obligation to follow instructions, the more likely it is to be assumed that there is an employment relationship); also *Willemsen/Mehrens*, NZA 2019, 1473: "Das nach der herkömmlichen Dogmatik bedeutsamste und mittlerweile im Gesetzestext (§ 1 Abs. 1 S. 2 AÜG) verankerte Kriterium der Abgrenzung von Dienst- bzw. Werkverträgen und Arbeitnehmerüberlassung ist der Grad der Weisungsgebundenheit" ("According to conventional doctrine, the most important criterion for differentiating between service contracts or contracts for work and labour leasing, which is now anchored in the wording of the law (Sec. 1(1) second sentence AÜG), is the degree to which the employee is bound by instructions").

[367] Cf. only BAG, NZA 1995, 622: "An employment relationship can also exist if the employee participates in the design of the program, but is subject to extensive instructions as to content, i.e. he or she has only a small degree of creative freedom, initiative and independence". The "low degree" of creative freedom then necessarily corresponds to a "high degree" of being bound by instructions.

[368] *Preis*, NZA 2018, 817 (820): "Freilich ist mit dem Kriterium [Fremdbestimmung] dann nichts gewonnen, am Ende entscheidet die Qualität der Weisungsgebundenheit". ("Admittedly, nothing is gained with the criterion [external determination]; in the end it is the quality of the obligation to follow instructions that decides").

[369] For more information on Art. 22(1) GDPR, see G. V. 6. In the USA, however, there seem to be cases where AI applications are used on a wider scale. In any case, there are reports that at Amazon AI applications not only monitor the productivity of workers, but also automatically give warnings or terminate workers; cf. *Lecher*, How Amazon automatically tracks and fires warehouse workers for 'productivity', Documents show how the company tracks and terminates workers, The Verge, April 25, 2019: https://www.theverge.com.

these circumstances, the employee is at any rate likely to be under far greater "pressure to follow instructions" from the employer than would be the case without the use of AI.

b) Heteronomy

Ultimately, however, the question of whether such "pressure to follow instructions" can be taken into account within the framework of the obligation to receive instruction stated in Section 611a(1), first sentence of the German Civil Code (Bürgerliches Gesetzbuch, BGB) could be moot if it were possible to relate the use of AI categorically to the characteristic of heteronomy also mentioned in that provision. In this respect, it should first be recalled that it is precisely this characteristic that has attracted particular attention in the recent discussion on the concept of employee. Some acknowledge for example that "work in the service of another and work subject to instructions from third parties go hand in hand", but at the same time point out that the characteristic of heteronomy is "recognisably broader than that of being bound by instructions". Working under external control, they say, is "also possible below the threshold of being bound by instructions".[370] Remarkably, some of the hopes associated with the characteristic of heteronomy are addressed specifically to the digital world of work.[371] In its decision on the legal status of crowdworkers, the Federal Labour Court (Bundesarbeitsgericht - BAG) also pointed out that the concept of external control encompasses more than the right to issue instructions.[372]

aa) Heteronomy outside of subjection to instructions?

However, as plausible as it may be to assume that the characteristic of being subject to external control goes further than that of being subject to instructions, it

[370] *Preis*, NZA 2018, 817 (824).

[371] *Preis*, NZA 2018, 817 (824): "Even someone who works on the other side of the world can be externally determined and thus personally dependent. The same applies vice versa to a person who works independently in the business of an employer. Consequently, it is not necessary to be bound by instructions in order to establish an employment relationship. Digital work is thus easier to grasp"; similarly *Bayreuther*, RdA 2020, 241 (248): "The constituent element of the performance of 'externally determined work' contained in Section 611a(1), first sentence of the German Civil Code (BGB) deserves far more esteem than is currently accorded to it in case law and literature. This feature would certainly have the potential to compensate for the fact that in the modern world of work the obligation to give instructions is evaporating".

[372] Cf. BAG, NZA 2021, 552 (para. 31) "The concepts of being bound by instructions and external control are closely connected and partly overlap. As a rule, an activity that is bound by instructions is at the same time externally determined. The obligation to follow instructions is the narrower criterion that characterises the core of the type of contract, which is defined in more detail by Sec. 611a (1) sentences 2 to 4 BGB; cf. also *Riesenhuber*, ZfA 2021, 1 (5.).

is unclear at the current state of the discussion how exactly to arrive at the conclusion that the former exists (sufficiently) outside the latter. This also applies to "digital work", which is primarily of interest here, where neither the supplementary reference to the European concept of the employee[373] nor the case law of the ECJ,[374] nor the additional reference to the aspect of "subordination"[375] or "being dependent on employment in a continuing obligation" which one sometimes encounters, provide any real guidance.[376]

As concerns the European concept of the employee, a closer analysis shows that it is largely unproductive; clear standards that could help in filling the criterion of heteronomy cannot be found in the case law of the CJEU. However, the aspects of "subordination" and "dependence" occasionally mentioned in the literature are also of only limited use in the present context. With regard to the former, this is true if only because "subordination" and "heteronomy" are obviously terms that can largely be used synonymously, whereby the former term appears to be even more in need of elaboration than the latter.[377] The aspect of "being dependent" on the other hand does not even have a place in the context of self-determination or heteronomy, but is rather related to the question of economic dependence or lack of autonomy. Accordingly, provides no answer to the question of when one can speak of such a strong degree "external determination" that it is justified to assert a person's status as an employee and to apply labour law to that person.[378]

bb) "Anticipation" of instructions through detailed contractual provisions

Consequently, the task remains, first, to determine the area in which there can be talk of "external control" without the person in question being bound by instructions (to a legally significant extent). The first such case that comes to mind is one in

[373] See *Risak/Dullinger*, The concept of 'worker' in EU law - Status quo and potential for change, ETUI Report 140, 2018.

[374] *Preis*, NZA 2018, 817 (824 f.) with the assessment that "on the one hand, Union law [...] can act as a guideline for a more flexible conceptualisation, on the other hand [...] the Union law concept of employee also noticeably challenges national law. *Bayreuther*, RdA 2020, 241 (245), on the other hand, is sober in his assessment: "Nevertheless, when viewed in the light of day, the relevant decisions hardly represent more than the concept of dependency under section 611a (1) BGB".

[375] Explicitly *Preis*, NZA 2018, 817 (824): "New impetus through the subordination relationship as an expression of externally determined work?"; the existence of a "subordination relationship" is also the focus of ECJ v. 22.4.2020 - C-692/19, NZA 2021, 1246 (para. 37).

[376] *Preis*, NZA 2018, 817 (824): "[...] both external determination and subordination occur through being dependent on employment in a continuing obligation".

[377] Ultimately, this is also the case with *Preis,* NZA 2018, 817 (824) itself: "However, the formula is obvious that subordination means or results in external determination and external determination means or results in subordination".

[378] The same objection is raised by *Bayreuther,* RdA 2020, 241 (248), who wants to get closer to the characteristic of external determination, among other things, by saying that the service provider has "no serious chance of winning".

which the contract itself gives the obligor such concrete instructions that that person has no leeway for self-determined action. The external determination is then not expressed in the fact that a person must follow the instructions of another person, but in the fact that the former is already obliged to act in a (very) specific way according to the clauses of the contract.[379] That the two situations are quite close together is shown by the fact that the employee who disregards instructions that are lawful under the contract ultimately violates the contractual obligation which concretises the instruction. However, despite all similarity, there remains a significant difference: this is that the employee, when concluding the contract, cannot know which instructions of the employer she will be subject to, whereas when detailed provisions are made in the contract the behaviour expected of her is already predetermined, which is why she can – at least theoretically – adjust much better to the concrete constraints that await her. One can possibly even go one step further: If the employee is already faced with concrete behavioural requirements, the "degree of voluntariness" with which she "waives" her right to self-determination is much greater than if this is not the case[380] – so that in the end it might be doubtful whether one can really speak of external determination in this case. Ultimately, the problem arises here of whether detailed stipulations of duties (making the exercise of rights to issue instructions superfluous) rule out the status of employee or – conversely – possibly even imply it.[381] The problem cannot be solved here, but is ultimately decided by whether the very fact that one person is subject to the authority of another to issue instructions prompts a special need for protection under labour law, as well as by whether one is prepared to consider permanent restrictions on self-determination based on detailed contractual conditions acceptable: the employee has, after all, consented to them "eyes wide open" by concluding the contract. However, the problem does not need to be solved here because the affirmation of personal dependency in the present context would not occur due to a specific *contract design*, but would result as a

[379] In this respect, one should think in particular of the cases of simpler activities where there is a lower need from the outset (in particular for technical instructions). The decision of the BAG, NZA 2021, 552, also dealt with this type of arrangement.

[380] Cf. in this respect also the considerations of the Co-determination Commission, BT-Drucks. VI/334, p. 61: "The economic compulsion to conclude an employment contract continues in the necessity to consent to the planning responsibility of the enterprise, its concretisation by the right to issue instructions and thus the existence and exercise of powers to issue instructions. Thus, in the Commission's view, the power to issue instructions cannot be justified solely by the employee's contractual consent. It is not a result of mutual agreement between the contracting parties, but exists independently of this".

[381] Cf. on this ErfK/Preis, 21st ed. 2021, § 611a BGB marginal no. 33 (with further references): "The assumption that the *binding drafting of the contract with* a specification of the areas relevant to instructions in the contract should speak against the assumption of an employment relationship is incorrect [...]. It is difficult to avoid labour law by skilfully drafting a contract that specifies the right to issue instructions in detail [...]. On the contrary, such a contractual arrangement can justify the need for protection [...]. If the content of the activity is precisely prescribed by contractual provisions, this speaks in favour of the status of employee [...]"; cf. on the whole also BAG, NZA 2021, 552, which refers not least to the fact that "the individual work steps of the activities to be performed […] were precisely specified by the job descriptions on the online platform [...]".

consequence of a specific *implementation* of the contract, namely, one characterised by the use of AI.[382]

cc) Inclusion

Considering against this background in what other ways an "external determination beyond being bound by instructions" could be justified, one could see it resulting from the insertion of a person into an external work organisation, thus under the aspect of integration. In this respect, however, one encounters the problem – which is quite controversial – of whether and to what extent this aspect can still be taken into account at all under Section 611a BGB – a question that represents an important facet of the general problem that it is uncertain to what extent it is still possible to rely on case law for the understanding of Section 611a BGB which the legislature did not explicitly include in the provision. While some want to continue to rely on the aspect of integration (and thereby attach importance to the feature of external determination),[383] others assume that recourse to this feature developed by case law is no longer permissible since the entry into force of Section 611a BGB.[384] The former merits agreement. From a practical point of view, the conclusion of an employment contract is accompanied by "subjection to external management and organisational authority",[385] which takes account of the fact that the employee is typically assigned no more than a specific function in the process structured by the employer based on the division of labour.[386] However, there is nothing to suggest that a "subjection" to this should be treated differently than a subjection to the employer’s authority to issue instructions.[387] It must therefore be noted that external control also occurs as a result of integration into the company.[388] In this respect, it is also true that integration lives on in the characteristic of heteronomy.[389]

[382] However, detailed specifications on the one hand and control on the other hand can "come together"; cf. in this respect indeed BAG, NZA 2021, 552 (and para. 43) with a reference also to the fact that the plaintiff "no longer had any significant scope for decision-making in the execution of the orders taken over, i.e. in the legal relationship already established".

[383] See only *Wank*, AuR 2017, 140 (150 f).

[384] Thus *Preis,* NZA 2018, 817 (820) with reference to the intention of the historical legislature, but quite critical in substance.

[385] Thus Mitbestimmungskommission, BT-Drucks. VI/334, p. 56.

[386] Cf. in this respect also BT-Drucks. VI/334, p. 59.

[387] Similarly, *Schubert*, RdA 2020, 248 (251): "The actual constraints of an organisation make specifications without these having to be expressed in the form of concrete directives"; cf. also *Schwarze,* RdA 2020, 38 (45), according to which "digital forms of work [...] will to an increased extent replace legal by factual availability of labour".

[388] Also BAG, NZA 2021, 552 (para. 31): "It [the external determination] shows itself in particular in the integration of the employee into the employer's work organisation".

[389] Thus also ErfKomm/Preis, 22nd ed. 2022, § 611a BGB marginal no. 41.

c) The aspect of "direction"

It is questionable whether external determination can also be assumed in other cases, thus even if a person is neither bound by instructions nor subject to the "management and organisational authority" of another[390] so that the employer "does not determine but directs" the employee's behaviour, as it were.

However, one point should be clarified at the start. Here we are obviously looking at more subtle and (at least potentially) weaker means of an employer to exert influence than when employees are subject to instruction. Thus, it seems clear from the outset that the only meaningful question is whether these means limit the other party's right of self-determination to *a legally significant extent*. Of course, it must be considered that being subject to instruction does not completely eradicate the employee's right of self-determination.

aa) AI-mediated possibilities of direction

In concrete terms, the question is whether there are possibilities to "direct" the behaviour of other persons that do not place them in a significantly different position than being subject to a right of instruction. It seems obvious that the availability and use of AI play a role in this context, and here the aspect of "direction" introduced here gains importance: anyone who is subjected to (or at least has to reckon with) comprehensive, detailed monitoring by a highly adaptive and "intelligent" system has every reason to adapt their behaviour from the beginning so as not to disappoint their counterpart's expectations if possible. Such "anticipatory obedience" is not to be treated any differently from "obedience" demonstrated by a person following the instructions of another.

bb) In particular: "AI Nudging"

Closely related to this is another set of questions. It must be taken into account that AI applications can not only be used for nudging, that is intentionally guiding people by deliberately triggering unconscious behavioural changes, but are in fact

[390] Nor – if one answers this question in the affirmative – does it result from detailed contractual specifications.

increasingly being used for the very purpose of inducing people to take certain decisions or to behave in a certain way.[391] In addition, AI applications are increasingly able to set very concrete incentives tailored to the individual (micro-targeting). It is also worth recalling the findings of the Committee of Ministers of the Council of Europe that draws attention to the growing ability of AI systems "not only to predict decisions, but also to influence emotions and thoughts and to change an expected course of action[392] and therefore attests to AI's "potential for highly personalised manipulation".[393] Once one realises this, one can no longer categorically deny the equality of "AI nudging" and "being bound by instructions".

In this context it seems advisable to pause for a moment and take a brief look at an issue that, while it may not appear to be directly comparable, will nevertheless be shown shortly to provide important insights. This issue is the relevance of nudging to fundamental rights, which has recently been the subject of lively discussion, particularly in connection with a legal regulation of organ donation.[394] As is well known, nudging is about making use of systematic human decision-making weaknesses so as to establish circumstances that steer people's behaviour in a certain direction without a priori excluding any of their choices.[395] How one judges the use of nudging by a state's legislature under the fundamental rights aspect depends greatly on one's conception of humanity. This cannot be discussed in detail here. However, two things should be beyond dispute, namely that the constitution "does not formulate any substantive-material guidelines for a well-understood use of freedom by the individual", and that "the individual is fundamentally not in the service of the supposedly "greater good", but conversely the state or constitutional order is in the service of the individual".[396] Against this background, however, there is good reason to acknowledge – contrary to some authors who emphasise the "autonomy-friendliness" of nudging,[397] and at least with regard to certain manifestations of state nudging – that not only does

[391] Illustrative in this context is the increasing use of so-called *dark patterns* on the internet; see *Sara Morrison*, Dark patterns, the tricks websites use to make you say yes, explained- How design can manipulate and coerce you into doing what websites want, April 1, 2021, www.vox.com. This is the design of user interfaces that are intended to induce users to make unintended and potentially disadvantageous decisions. *Dark patterns are* used in particular on shopping websites, where visitors are encouraged to make more purchases or disclose more information than they otherwise would; see *Mathur/Acar/Friedman/Lucherini/Mayer/Marshini/Narayanan*, Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites, Proceedings of the ACM on Human-Computer Interaction, 2019, p. 81.
[392] Recommendation of 13 February 2019 (at 8).
[393] Recommendation of 8 April 2020 (under A.4.).
[394] See only *Kirchhof*, ZRP 2015, 136.
[395] See only *Honer*, DÖV 2019, 940 (942); also *Ivankovics*, JuWissBlog No. 57/2018 v. 31.5.2018: https://www.juwiss.de/57-2018.
[396] Thus *Honer*, DÖV 2019, 940 (944).
[397] See, for example, *Sunstein*, The Ethics of Nudging, Yale Journal on Regulation 32 (2015), 413 (415), who makes two arguments: Firstly, that nudges and a certain "choice architecture" are unavoidable anyway, and secondly, that nudges and many forms of decision architecture are justifiable and even necessary for ethical reasons.

encroachment on fundamental rights exist but also – and this is important when it comes to potential justifications of the same – that it tends to be particularly intensive.[398] However, just as the citizens' right to self-determination is to be taken seriously by the state, it should also be respected by third parties. And even if the latter is not (necessarily) about imposing prohibitions on nudging, there is every reason to see it for what it is, even in the relationship between private individuals: a (potentially) serious encroachment on the right of self-determination of the other party.

However, if one wants to gain a clearer impression of how AI is used in the present context, then one should also take a closer look at the literature on the topic outside the field of law. One recent study by researchers from different disciplines on how modern AI works looks, for example, at so-called "algorithmic recommending", by which the authors mean the use of algorithms that are supposed to "prompt the targeted worker to make decisions preferred by the choice architect".[399] They point out that such recommendations influence workers' decisions by a process of derivation based on patterns found in the data, which can eliminate the need for instructions.[400] Companies also often use "algorithmic restricting" as the use of algorithms "to display only certain information and allow specific behaviors while preventing others".[401] In addition, there is the widespread use of "algorithmic rating" in the form of "ongoing aggregation of quantitative and qualitative feedback about worker performance from both internal and external sources".[402] Overall, the authors conclude that "[w]orker activities can be more constrained under algorithmic control than under previous regimes of rational control because algorithmic control can be more comprehensive in terms of how it directs, evaluates, and disciplines workers". They also point out that algorithms " can also

[398] *Honer*, DÖV 2019, 940 (947): "Insbesondere dort, wo die Steuerung unbemerkt erfolgt und kein eigener Reflexionsprozess ansetzt, ist der *Nudge* [...] eben nicht darauf ausgerichtet, eine bewusste eigene Entscheidung des Bürgers herbeizuführen. Seine normativ zu unterstellende Entscheidungskompetenz soll hier nicht zur Entfaltung gebracht werden. Die Entfernung zur Manipulation ist nicht weit. Das steht in klarem Widerspruch zum aufgezeigten, an der Menschenwürde orientierten Menschenbild des Grundgesetzes. Auch das intensiviert den Grundrechtseingriff." ("In particular in those situations where the control takes place unnoticed and no own reflection process begins, the nudge [...] is not aimed at bringing about a conscious decision of the citizen. His decision-making competence, which is to be assumed normatively, is not to be brought to fruition here. The distance to manipulation is not far. This is in clear contradiction to the Basic Law's conception of the human being, which is oriented towards human dignity. This also intensifies the encroachment on fundamental rights".)

[399] See *Kellogg/Vantine/Christin*, Algorithms at Work: The New Contested Terrain of Control, Academy of Management Annals 2020, 366 (372).

[400] See *Kellogg/Vantine/Christin*, Algorithms at Work: The New Contested Terrain of Control, Academy of Management Annals 2020, 366 (372).

[401] See *Kellogg/Vantine/Christin*, Algorithms at Work: The New Contested Terrain of Control, Academy of Management Annals 2020, 366 (375).

[402] See *Kellogg/Vantine/Christin*, Algorithms at Work: The New Contested Terrain of Control, Academy of Management Annals 2020, 366 (378).

provide rewards and penalties in real time".[403] Some authors speak of "hypernudging" as an "algorithmic real-time personalisation and reconfiguration of electoral architectures based on large amounts of personal data".[404] That at least "a "soft" incentive system should be considered as a minus to contractual obligation", as is claimed in the German literature,[405] therefore appears to be anything but a foregone conclusion.

Particularly in the area of the platform economy, there are studies that describe the manipulative use of AI in very concrete terms. This applies, for example, to AI that is used by ride service platforms to induce drivers to preferentially drive to certain areas via certain pricing mechanisms, and which thereby take advantage of the existing information asymmetry between platform and driver.[406] There are reports that Uber has "[e]mploy[ed] hundreds of social scientists and data scientists [and] has experimented with video game techniques, graphics and noncash rewards of little value that can prod drivers into working longer and harder — and sometimes at hours and locations that are less lucrative for them".[407] Last but not least, it is pointed out that platforms sit, in a sense, at the intersection of consumers and service providers, and therefore "have a unique capacity to monitor and nudge all participants".[408] The advantages of nudges are obvious: “[N]udges are usually not very intrusive, easily scalable, and employees are not forced to make extensive changes to their working habits. Of course, for most companies, it will be difficult to easily measure the effectiveness of nudges and new default rules. However, herein lies the great opportunity of digitalisation, big data, and an evidence-based approach to management: through continuous collection and analysis of data, companies will soon be able to assess quickly which nudges tend to work for which

[403] See *Kellogg/Vantine/Christin*, Algorithms at Work: The New Contested Terrain of Control, Academy of Management Annals 2020, 366 (386, 381). See also Pignot, Who is pulling the strings in the platform economy? Accounting for the dark and unexpected sides of algorithmic control, 2021, 20, according to which the persuasive *performance of* algorithms potentially goes deeper than that of conventional control mechanisms.

[404] Also *Lanzing*, „Strongly Recommended“Revisiting Decisional Privacy to Judge Hypernudging in SelfTracking Technologies, Philosophy & Technology 2019, 549 (553); vgl. dazu auch Mendelsohn, MMR 2021, 857 (859). Cf. *Mendelsohn*, MMR 2021, 857 (859): "algorithmischen Echtzeit-Personalisierung und Rekonfiguration von Wahlarchitekturen auf der Grundlage großer Mengen persönlicher Daten".

[405] Thus *Thüsing/Hütter-Brungs*, NZA-RR 2021, 231 (234): " "ein 'weiches' Anreizsystem als Minus zur vertraglichen Verpflichtung".

[406] See *Rosenblatt/Stark*, International Journal of Communication 2016), 3758. However, it is recently reported that Uber is testing a new algorithm in some cities in the U.S. that discloses destination and pay; see *Bellon*, Uber revamps driver pay algorithm in large U.S. pilot to attract drivers, Feb 26, 2022: https://www.reuters.com.

[407] *Scheiber*, How Uber Uses Psychological Tricks to Push Its Drivers' Buttons: https://www.nytimes.com. For example, drivers who want to opt out of the system are alerted that they are about to reach a certain earnings target. They are also alerted to the next driving opportunity even before their current journey is over.

[408] *Calo/Rosenblat*, The Taking Economy: Uber, Information, and Power, Columbia Law Review 2017, 1623 (1624).

knowledge worker, and which ones don't—ultimately, leading to more personalised nudges and default rules individually tailored to each knowledge worker".[409]

Looking at the aspect of "external control" through nudging, it also becomes clear that the various aspects put forward in the literature such as "being dependent on employment in a continuing obligation" do have a legitimate core.[410] Admittedly, this only refers to economic dependence and thus "being dependent" can at best constitute a "similarity to an employee".[411] However, the present context has to do with a completely different phenomenon, namely that people as a rule are that much more receptive to nudges from their contractual partner the more reason they have to give in to the incentives set by that partner. Accordingly, it is one thing to base the status of employee solely on the existence of economic dependence (which is excluded *de lege lata*), but quite another to not completely disregard this dependence when testing for heteronomy, but instead to include it in one's consideration. Let it be noted in passing that the aspect of "being dependent" should not be disregarded in other respects either. When assessing the significance of ratings, for instance, it is quite significant how tangible their material consequences are for the person concerned.[412]

Against this background, the recent decision of the BAG on the legal status of a crowdworker seems downright emblematic. In this case, the court deduced that an employee was externally directed (and thus had the status of employee) from the fact that, on the one hand, the employee's behaviour was "directed" by the other party (using the incentive function of an evaluation system) based on demand,[413] and on the other hand, the employee was only able to "exercise his activity in an economically meaningful way" by regularly accepting orders over a long period of time.[414]

[409] *Ebert/Freibichler*, Nudge management: applying behavioural science to increase knowledge worker productivity, Journal of Organization Design 2017, 6:4 (5).
[410] *Preis*, NZA 2018, 817 (824).
[411] So also LAG Munich, NZA 2020, 316 (and para. 129).
[412] See most recently the decision of the UK Supreme Court, Judgment *Uber BV and others (Appellants) v. Aslam and others (Respondents)*, [2021] UKSC 5 on the qualification of Uber drivers as workers, where the court considered both the rating system in place and the fact that a driver whose percentage acceptance rate falls below a level set by Uber London (or whose cancellation rate exceeds a set level) "will (receive) an escalating series of alerts which, if performance does not improve, will result in the driver being automatically logged off the Uber app and barred from re-registering for ten minutes", ibid (at para. 97).
[413] Cf. BAG, NZA 2021, 552 (para. 50); crit. *Heckelmann*, NZA 2022, 73 (74); Sittard/Pant, jm 2021, 416. In contrast, *Schmidt,* NZA 2021, 1232 (1235) speaks of an "algorithm-based behavioural control" and "external determination in a modern guise".
[414] Cf. BAG, NZA 2021, 552 (at para. 48 f.); crit. *Häferer/Koops*, NJW 2021, 1787 (1789 f.). Criticism of the decision also *Wisskirchen/Haupt*, RdA 2021, 355 (359), who, however, consider legislative action to be necessary.

d) Summary

In summary, three things can be stated about external control: first, it is indeed broader than the characteristic of being subject to instructions; second, it encompasses more than mere integration;[415] and third and above all, the availability and use of AI should be taken into account when considering external control (and thus when qualifying a contractual relationship as an employment contract). Whether this can lead to results that sufficiently take into account the need for protection of platform employees (or whether additional intervention by the legislature is required in this respect)[416] is not to be decided here. However, the aim here was only to show that "control by AI" already carries weight in determining employee status under current law.

3. AI and platform employment

As stated, the aspect of "control by AI" plays a central role, especially in the context of platform employment. It is therefore not surprising that the way AI functions, as sketched out above, has also had a strong influence on the Proposal for a Directive on improving working conditions in platform work.[417] For example, in the communication underlying the proposal, the Commission explicitly states that "algorithmic management can conceal subordination behind a claim of human supervision" in that "the control exercised through algorithms deprives [a purportedly self-employed person] of the autonomy enjoyed by a genuine self-employed person."[418] The explanatory memorandum of the proposal also states that algorithmic management "conceals the existence of subordination and control by the digital labour platform on the persons performing the work". Accordingly, Article 6 of the Directive imposes on platforms far-reaching informational

[415] Cf. therefore ErfKomm/Preis, 21st ed. 2021, § 611a BGB marginal no. 41.

[416] In favour of the latter, for example *Kocher*, ZEuP 2021, 606 (632) with the assessment that the "conventional categories, criteria and indications [...] do not (can) capture the specifics of digital indirect control on labour platforms well and that an "appropriate regulation of this field [...] must take better account of the character of the platforms as market organisers" (Die "herkömmlichen Kategorien, Kriterien und Indizien [können] die Spezifika der digitalen indirekten Steuerung auf Arbeitsplattformen nicht gut erfassen"; eine "angemessene Regulierung dieses Feldes muss den Charakter der Plattformen als Marktorganisatoren besser berücksichtigen".

[417] Commenting in depth on this proposal *Krause*, NZA 2022, 521.

[418] See Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Better working conditions for a stronger social Europe: harnessing the full benefits of digitalisation for the future of work, 9.12.2021, COM(2021) 761 final, p. 3. For a review and analysis of the (limited) case law of European courts on the significance of *algorithmic management,* see *Hießl*, Case law on algorithmic management at the workplace: Cross-European comparative analysis and tentative conclusions: https://ssrn.com/abstract=3982735.

obligations vis-à-vis workers, while Article 9 lays down information and consultation rights to which primarily the representatives of the platform workers and secondarily the platform workers themselves are entitled. Though this particular proposal addresses only platform work, the Commission explicitly recognises that while "algorithmic management (…) is clearly inherent to digital labour platforms“, it „is used in a growing number ofways in the wider labour market“.[419]

Above all, however, this aspect comes into play in connection with the determination of worker status. Thus, according to Article 4(2)(c) and (d) of the proposed Directive, the circumstances that give rise to a presumption of employee status include: "supervising the performance of work or verifying the quality of the results of the work including by electronic means" as well as "effectively restricting the freedom, including through sanctions, to organise one’s work, in particular the discretion to choose one’s working hours or periods of absence, to accept or to refuse tasks or to use subcontractors or substitutes". The regulation appears to be rather vague, avoiding as it does the question of where exactly the threshold lies between "supervision of the performance of work or verification of the quality of the work results" and sufficient "control" and when the freedom of work organisation, including through sanctions, is "effectively restricted".[420] Quite apart from this, however, there is no denying the fact that the proposed Directive considers "algorithmic management" to have significance for determining a worker's status as an employee.

II. Possible legal capacity of AI

A question of obvious fundamental importance is whether a robot or an AI application should be endowed with its own legal personality. The discussion on this was triggered by the European Parliament's 2017 resolution on the question of civil liability of robots, where the Parliament had indeed considered creating a specific legal status for robots in the long term "so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause",

[419] Proposal for a Directive of the European Parliament and of the Council on improving working conditions in platform work, 9.12.2021, COM(2021) 762 final, p. 2.

[420] Similarly crit. *Racabi*, What Can U.S. Labor Take from the Proposed E.U. Directive of Regulations of Platform Workers?, https://onlabor.org: "Both the first and the second steps of the classification route are murky as to substance and procedure, which are left for E.U. states and judicial venues to develop and actualize, and for platform employers to exploit and leverage". See also *Waas*, ZRP 2022; similarly crit. *Krause*, NZA 2022, 521 (528); *Junker*, EuZA 2022, 141.

while at the same time considering "applying electronic personality" to cases "where robots make autonomous decisions or otherwise interact with third parties independently".[421] In the labour law literature on the subject, the discussion on the "legal capacity of AI" reverberates in contributions that invoke the "employer status of algorithms" in their titles.[422] Some authors may be motivated by the desire to cleverly "market" their products. But as the resolution of the European Parliament shows, the idea of a "robo boss"[423] is by no means pure science fiction.

However, the ensuing discussion promptly revealed that the granting of legal capacity considered by Parliament has hardly any supporters. This was already made clear in an open letter written by several experts on AI and robotics, leading figures in business and legal, medical and ethical experts, who rejected the idea of legal capacity for AI under all relevant aspects: an analogy with natural persons, an analogy with legal persons and the use of the trust model.[424] In particular, the analogy to the recognition of legal persons, which has been used on various occasions in favour of granting legal capacity, does not in fact hold water. For while legal persons are given "capacity to act" by natural persons, this is precisely not the case with robots.[425] Above all, however, machines are not (yet) capable of autonomous decision-making, which could justify putting them on an equal footing with natural persons in terms of liability law.[426] Also, the recognition of legal persons is intended to enable individuals to pursue objectives they cannot pursue over a long term or only through a division of labour.[427] Nothing comparable applies to AI.[428] However, caution is also required with regard to the analogy to the legal person because not only does the term aim to consolidate very different organisational forms "under one roof", but also the legal person in some respects

[421] European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL), paragraph 59.

[422] See for example *Aloisi/De Stefano*, Introducing the Algorithmic Boss, April 20, 2021: https://www.ie.edu/insights/articles/introducing-the-algorithmic-boss/.

[423] Cf. *HöpfnerDaum*, ZfA 2021, 467.

[424] http://www.robotics-openletter.eu/.

[425] Cf. the Open Letter of 29.06.2020 (under 2b)): "The legal status for a robot can't derive from the Legal Entity model, since it implies the existence of human persons behind the legal person to represent and direct it. And this is not the case for a robot": http://www.robotics-openletter.eu/.

[426] Cf. *Bertolini*, Artificial Intelligence and Civil Liability - Study requested by the JURI Committee, 2020, p. 36: "In particular, machines are things, products and artefacts of human intellect, and there are no ontological grounds to justify their equation to humans, so long as they do not display such a form of strong autonomy that amounts to freedom of self-determination in the outcomes the system pursues and in the ways it chooses to accomplish them. Currently there is no machine that would be able to display such a level of autonomy, and there is no reason to desire the development of such a system that being more intelligent and capable than any human life form, and being also independent, could pursue its own intended ends. Technological development does not justify acknowledging such a level of autonomy on the side of any AI application existing or being developed".

[427] https://www.staatslexikon-online.de/Lexikon/Juristische_Person.

[428] In conclusion, *Haagen,* Verantwortung für Künstliche Intelligenz - Ethische Aspekte und zivilrechtliche Anforderungen bei der Herstellung von KI-Systemen, 2021, p. 184; also *Banteka*, Artificially Intelligent Persons, Houston Law Review 2020: https://ssrn.com/abstract=3552269.

enjoys a more advantageous position than natural persons (which is by no means unobjectionable). [429] Whether it is advisable to apply a blanket solution to AI that at the same time grants privileged status is anything but settled.[430]

This leaves the (albeit rather pragmatic) aim of granting legal capacity in order to close possible gaps in liability by granting a kind of electronic legal capacity. It should not, however, be considered as a foregone conclusion that these gaps exist, since both the development and the use of AI systems involve people who can as a rule be addressed. In the current situation, then, it is more a matter of "clearing the way" to imposing liability on these persons – for instance by establishing rules of presumption or easing the burden of proof - but not of replacing it with liability for machines.[431] In all of this, it must also be considered that it can only make sense to grant legal capacity to machines if they are also allocated recoverable assets that can be accessed if necessary. This would ultimately lead to a limitation of liability for "damage by machine", though, and this can hardly be the intended outcome. However, it would still be possible to create a duty to take out liability insurance with a certain minimum coverage. But even then there would still be a limitation of liability in every case, namely, up to the insured sum. Considering, furthermore, that the risk of a claim cannot have a deterrent effect on a robot and the imposition of liability thus fails to have a behaviour-controlling effect, the idea of subjective liability fails completely to offer an attractive argument.[432]

The European Parliament itself has also departed from its earlier position in a recent resolution on civil liability, stating that "all physical or virtual activities, devices or processes that are driven by AI-systems may technically be the direct or indirect cause of harm or damage, yet are nearly always the result of someone building, deploying or interfering with the systems". Accordingly, it has also stated that "it is not necessary to give legal personality to AI systems".[433] Furthermore,

[429] https://www.staatslexikon-online.de/Lexikon/Juristische_Person.

[430] *Negri*, Robot as Legal Person: Electronic Personhood in Robotics and Artificial Intelligence, frontiers in Robots and AI, Hypothesis and Theory: 10.3389/frobt.2021.789327.

[431] See Expert Group on Liability and New Technologies - New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 38: "Harm caused by even fully autonomous technologies is generally reducible to risks attributable to natural persons or existing categories of legal persons, and where this is not the case, new laws directed at individuals are a better response than creating a new category of legal person".

[432] See also *Wagner*, VersR 2020, 717 (739).

[433] European Parliament Resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for the use of artificial intelligence, 2020/2014(INL) at 7, where it further states that "opacity, connectivity and autonomy of AI-systems could make it in practice very difficult or even impossible to trace back specific harmful actions of AI-systems to specific human input or to decisions in the design", but "one is nevertheless able to circumvent this obstacle by making the different persons in the whole value chain who create, maintain or control the risk associated with the AI-system liable".

the European Parliament even considers that "Any required changes in the existing legal framework should start with the clarification that AI-systems have neither legal personality nor human conscience, and that their sole task is to serve humanity".[434] However, this argument does not quite seem to serve its purpose: the fact that AI systems do not have a "human conscience" is reason enough not to grant them decision-making powers over humans, but is not enough to stand in the way of granting them legal capacity. And the fact that AI systems are intended to serve humanity is likewise not an argument against granting legal capacity if, for example, it should turn out that compensation for damage that has occurred would otherwise be endangered or even (practically) impossible.

However, the discussion on legal capacity does not seem likely to fall silent, if only because technical development is continuing and the "autonomy capacity" of AI systems will in all likelihood increase. There is also no getting around the fact that the proponents of giving AI systems legal capacity assess the chances of injured parties to always be able to find a respective injuring party considerably less favourably than, for example, the European Parliament,[435] but consider access to a legally capable machine to be relatively straightforward.[436] Moreover, flexible solutions could be developed, possibly also differentiating between individual areas of law.[437] Nevertheless, with the current state of the art, the granting of legal capacity seems neither necessary nor sensible.[438]

[434] Annex (6). The European Parliament took the same position in its resolution of 20 October 2020 on intellectual property rights in the development of AI technologies, but justified this (under 13.) with the protection of human creators: "[...] the autonomisation of the creative process of generating content of an artistic nature can raise issues relating to the ownership of IPRs covering that content [...] in this connection, [...] it would not be appropriate to seek to impart legal personality to AI technologies and points out the negative impact of such a possibility on incentives for human creators"; also instructive on the problem *Chesterman*, Artificial Intelligence and the Limits of Legal Personality, in: International and Comparative Law Quarterly 2020, 819 (834 et seq.).

[435] However, there are also sceptical voices in this respect; cf. only *Papakonstantinou/de Hert*: Refusing to award legal personality to AI: Why the European Parliament got it wrong - European Law Blog: "Exactly because AI will infiltrate all of human activity, indistinguishable from any technology and embedded in all of our daily decision-making systems, it will be impossible to "trace back specific harmful actions of A" to a particular "someone". Any AI setup will most likely involve a number of (cross-border) complex agreements between many developers, deployers and users before it reaches an end-user. Identifying the "someone" liable within this international scheme will be extremely difficult for such end-users without the (expensive) assistance of legal and technical experts. On the contrary, end-users would be better served through a one-on-one relationship, whereby the AI effect that affects them is visibly caused by a specific entity; only by granting legal personality to AI may warrant that this will be an identifiable entity, rather than a string of opaque multinational organisations hiding behind complex licensing and development agreements".

[436] See again *Papakonstantinou/de Hert*: Refusing to award legal personality to AI: Why the European Parliament got it wrong - European Law Blog: "Legal personality to AI will mean that each individual affected by it will have a specific legal entity facing him or her locally, in the same manner as is the case with legal persons today".

[437] See also *Papakonstantinou/de Hert*: "Legal personality will mean that each field of law (civil law, tax law, employment law, penal law, competition law) will be allowed with the freedom to assess the legal issues posed by AI within its own boundaries and under its own rules and principles".

[438] *Chesterman*, Artificial Intelligence and the Limits of Legal Personality, in: International and Comparative Law Quarterly 2020, 819 (843): "At least for the foreseeable future, the better solution is to rely on existing categories, with responsibility for wrongdoing tied to users, owners, or manufacturers rather than the AI systems themselves".

III. AI and exercise of the right to issue instructions

A characteristic feature of the employment contract is the right to issue instructions, to which the legislature indeed assigns a prominent role in the definition of the employment contract in Section 611a(1), first sentence of the German Civil Code (BGB) in the form of its mirror image of "being bound by instructions". As is well known, the right to issue instructions is regulated in Section 106 of the Commercial Code (GewO). Pursuant to Section 106, first sentence, of the GewO, the employer may "determine the content, place and time of the work performance in more detail at his reasonable discretion, insofar as these working conditions are not stipulated by the employment contract, provisions of a works agreement, an applicable collective agreement or statutory provisions". Pursuant to Section 106, second sentence, this also applies "with regard to the order and conduct of the employees in the enterprise".

In the present context, in view of the existence of autonomous AI systems, one must ask whether instructions by these systems themselves are permissible or fail due to the requirement that the direction of work can only be further determined "in accordance with reasonable discretion". It is true that Article 22(1) of the GDPR sets limits on automated decisions;[439] this will be discussed in more detail later.[440] However, this must be separated from the completely different question of whether Section 106, first sentence GewO leaves any room at all for instructions by AI systems vis-à-vis employees. It will not be argued in detail here that so-called "autonomous declarations of intent", that is, declarations by AI systems that are no longer necessarily based on previously clearly defined conditions, are strictly attributable to the employer.[441] To assume otherwise would be to ignore the fact that the employer quite deliberately outsources its decision-making power to AIs in such cases. The fact that the declarations made by the system are neither predictable nor traceable, or only to a limited extent, only affects the so-called business intention, which is not a necessary component of a declaration of intent.[442]

[439] Cf. on this also e.g. *Däubler*, Digitalisierung und Arbeitsrecht, 7th ed. 2020, p. 299 with a qualification of such instructions as mere "recommendations".
[440] Cf. G. V. 6.
[441] Cf. *Paulus/Matzke*, ZfPW 2018, 431 (443).
[442] Cf. on this *Höpfner/Daum*, ZfA 2021, 467 (475 f.).

1. Deficits in human decision-making

In answering the question of whether "AI instructions" are permissible, it seems helpful to first take a step back and consider whether AI might actually have advantages over human decision-making.[443] In fact, there is little reason to approach the issue with total self-confidence. Numerous studies - for example, on the possible use of AI in the judiciary - show that people often make mistakes and that rationality deficits and influences outside of the law can also be observed in judicial decisions. Thus, humans do not always prove capable of "compensating for unconscious weaknesses and errors entirely through conscious reflection".[444] This is demonstrated quite strikingly by the Implicit Association Test (IAT), a measurement method used in social psychology.[445] It shows that a majority of test persons associate positive terms with pictures of light-skinned people more often than with pictures of dark-skinned people.[446] The effects of priming or framing are also not uncommon.[447] With the former, a certain previous experience leads to the activation of special associations in the memory. In the case of so-called media priming, for example, politicians are judged preferentially according to criteria that are foregrounded in general media coverage.[448] The latter (and probably more significant) effect concerns the fact that different formulations of a message with the same content influence the recipient's behaviour in different ways. In other words: Information with the same content is processed differently depending on the form of presentation. This effect can be particularly exascerbated by the fact that people often cannot completely block out information they are familiar with.[449] Another such effect called the anchor effect has recently become a focus of discussion. This term from cognitive psychology describes the systematic distortion of numerical judgements in the direction of a numerical value (arbitrarily

[443] However, it should not be overlooked in this context that there are always people "behind" algorithms; cf. *Zekos*, Political, Economic and Legal Effects of Artificial Intelligence - Governance, Digital Economy and Society, 2022, p. 483 f.: "Algorithms are mathematical models of the real world and scientists construct algorithms to take in data and find correlations or make predictions. Thus, humans energetically design algorithms in a number of manners by choosing an algorithm's objectives, determining what the input will be, picking whether to use proxies, etc. and, once the algorithm is functioning, decide whether and how to confirm inpractice that it is generating accurate results".

[444] Thus *Nink*, Justiz und Algorithmen, 2021, p. 47 f.

[445] See, for example, *Kang/Bennett/Carbado/Casey/Dasgupta/Faigman/Godsil/ Greenwald/Levinson/Mnookin,* Implicit Bias in the Courtroom, UCLA Law Review, 2012, UCLA School of Law Research Paper: https://ssrn.com/abstract=2026540; *Rachlinski/Johnson/Wistrich/Guthrie, Chris,* Does Unconscious Racial Bias Affect Trial Judges? Notre Dame Law Review, 2009, Vanderbilt Public Law Research Paper No. 09-11: https://ssrn.com/abstract=1374497; *Kang*, What Judges Can Do About Implicit Bias, Court Review: https://ssrn.com/abstract=4033906.

[446] *Nink*, Justiz und Algorithmen, 2021, p. 48 with further references.

[447] *Nink*, Justiz und Algorithmen, 2021 p. 50 ff.

[448] Wikipedia: Media priming.

[449] See also *Nink*, Justiz und Algorithmen, 2021, p. 52.

given as a starting point), the so-called anchor.[450] This effect is used, for example, in bait-and-switch offers. In contrast, so-called hindsight bias occurs when earlier predictions about an event are systematically misjudged after the outcome of the event is known.[451] Confirmation bias, the last example to be mentioned here, describes the phenomenon that people are more likely to perceive information and evaluate it as correct the better it fits their own expectations and appears suitable to support their own point of view.[452] The so-called echo chambers in social media, which are themselves based on corresponding algorithms,[453] are an illustrative example. In this respect, it is also true that human decision-making is prone to error in many respects and, in particular, has a not insignificant "discrimination potential".[454] At this point, we are not even talking about the fact that decisions are never made in a vacuum, that is, a judge or, more generally, a decision-maker may be exposed to public pressure of some kind, which then has an impact on the content of his or her decision.[455]

2. Inadmissibility of "machine decisions"

If discretionary decisions cannot be left to machines, then it is because the decision of a machine can never be based on the exercise of discretion in the legal sense: "Individual case justice" cannot be forced into an automation system, both conceptually and by its very nature,[456] since this is necessarily based on schematisation and therefore the relevant circumstances can only be anticipated to a limited extent.[457] Another factor is that decisions are always based on facts. However, in addition to "hard" facts, there are also "soft" facts that cannot or not easily be quantified, and therefore cannot be imported into an automatic decision-making system.[458] For decisions, the time of the decision regularly plays an

[450] *Nink*, Justiz und Algorithmen, 2021, p. 53. These and other defects must also be taken into account when it comes to detecting errors in algorithmic decisions; cf. *Rhue,* Affectively Mistaken? How Human Augmentation and Information Transparency Offset Algorithmic Failures in Emotion Recognition AI, November 22, 2019: https://ssrn.com/abstract=3492129.

[451] *Nink*, Justiz und Algorithmen, 2021, p. 61 ff.

[452] *Nink*, Justiz und Algorithmen, 2021, p. 63 ff.

[453] See for example *Lambrecht/Sen/Tucker/Wiertz,* Algorithmic Recommendations and Earned Media: Investigating Product Echo Chambers on YouTube, 27 Oct 2021: https://ssrn.com/abstract=3951425.

[454] Cf. on this *Nink*, Justiz und Algorithmen, 2021, p. 76 ff.

[455] Cf. on this *Nink*, Justiz und Algorithmen, 2021, p. 66 ff; fundamental *Esser*, Vorverständnis und Methodenwahl in der Rechtsfindung: Rationalitätsgarantien der richterlichen Entscheidungspraxis, 1970; cf. on the whole topic also *Möllers*, Juristische Methodenlehre, 3rd ed. 2020, p. 24 ff.

[456] *Nink*, Justiz und Algorithmen, 2021, p. 196: "It remains indispensable for judicial decisions that the decision-maker can understand and evaluate the individual case and all its aspects".

[457] Cf. again *Nink*, Justiz und Algorithmen, 2021, p. 198, who at the same time draws attention to the danger "that algorithmic forecasts and decisions reduce them [the individual affected by the decision] to belonging to certain groups".

[458] Similarly *Nink*, Justiz und Algorithmen, 2021, p. 179 f.

essential role. This also sets limits to the "pre-programming" of decisions.[459] In this context, it is important to realise that AI is necessarily "backward-looking" and that there is therefore always a danger that the past will simply be "written down" in AI decisions. Also, according to case law, the exercise of the right to issue instructions pursuant to Section 106, first sentence GewO requires "a weighing of the respective interests according to constitutional and legal value decisions, the general principles of proportionality and appropriateness as well as custom and reasonableness", whereby "all circumstances of the individual case (must) be included" in the weighing.[460] However, no one will claim that AI systems can make such "value decisions".[461] Closely related to this is the fact that the employer is obliged under Section 106, first sentence GewO to weigh up, that is, to "evaluate legal positions from the perspective of priority" with the aim of achieving a balance between conflicting interests and concerns.[462] AI systems are not able to do this either, at least not at present.[463]

A parallel to administrative discretion[464] may clarify the foregoing: Discretionary decisions by an administration are intended to enable decisions to be made that adhere to the facts with the aim of fairness in individual cases. The granting of discretion is based precisely on the fact that the legislature cannot assess the interests and concerns of the parties involved a priori and, moreover, cannot take into account the particularities of the individual case.[465] Nor can this be fed into the process of "decision-making" of a machine: one cannot speak of "discretion", much less of "reasonable discretion".[466]

However, there are those who cite a difference between "employer's discretion" and administrative discretion and would derive from it that decisions without (sufficient) consideration also meet the requirements of Section 106, first sentence

[459] See also *Nink*, Justiz und Algorithmen, 2021, p. 194.

[460] BAG, NZA-RR 2018, 568 (and para. 39).

[461] Cf. also *Rollberg*, Algorithmen in der Justiz - Rechtsfragen zum Einsatz von Legal Tech im Zivilprozess, 2020, pp. 69 ff, 128 ff.

[462] Thus (on consideration in company law) *Freund,* Die Abwägung im Gesellschaftsrecht, NZG 2020, 1328 (1328).

[463] Which is why they can fail at the simplest tasks; cf. only *Pavlus*, The Easy Questions That Stump Computers – What happens when you stack logs in a fireplace and drop a match? Some of the smartest machines have no idea, 2 May 2020. https://www.theatlantic.com/technology/archive/2020/05/computers-common-sense/611050/; vgl. auch *Choi*, The Curious Case of Commonsense Intelligence, Daedalus 2022, 139. https://doi.org/10.1162/DAED_a_01906. *Hutson*, Can Computers Learn Common Sense? A.I. researchers are making progress on a longterm goal: giving their programs the kind of knowledge we take for granted, 5 Apr 2022. https://www.newyorker.com/tech/annals-of-technology/can-computers-learn-common-sense.

[464] On AI in administrative practice, most recently *Tischbirek*, Zeitschrift für Digitalisierung und Recht (ZfDR) 2021, 307.

[465] Similarly, *Höpfner/Daum,* ZfA 2021, 467 (477), also emphasise that there are "some similarities" between the equitable discretion within the meaning of section 106 sentence 1 GewO and the administrative discretion.

[466] It is controversial whether "equity" is a uniform standard; cf. only *Völzmann-Stickelbrock*, in: *Herberger/Martinek/Rüßmann/Weth/Würdinger*, jurisPK-BGB Vol. 2, 2020, § 315 marginal no. 16 et seq. However, this should not be relevant in the present context.

GewO and are thus generally permissible.[467] This, they claim, results from the fact that the administrative court's review of discretionary decisions of the administration - for reasons of separation of powers[468] - is limited to the process of weighing (non-use or misuse of discretion), whereas determining the fairness of an instruction issued by the employer only depends on "whether the result, i.e. the content of the instruction, meets the legal requirements". Since the principle of separation of powers does not apply in the relationship between employer and employee, it is argued, "a limitation of judicial review to the process of weighing [...] is not appropriate". However, whether the employer has carried out a comprehensive weighing of interests or "the instruction merely happened to be in accordance with equity" is irrelevant for the lawfulness of the instruction, in this view.[469] With regard to the latter, it is argued, the wording of Section 315(3), first sentence BGB, according to which the provision is only binding on the other party "if it is equitable", already indicates a mere review of the result.[470]

However, this view cannot be followed. First of all, it should be noted that if judicial review were limited to the result of the balancing process, possible impairments of the employee's interests and concerns would be deliberately ignored only because the result might "happen to be fair". In other words, an instruction would be valid because it is not inequitable, even though a different instruction might have better served the interests of the parties involved. If one realises this, it immediately becomes clear that it is misleading to speak of "a restriction of judicial review to the weighing process [...] being inappropriate" in connection with Section 106, first sentence GewO. In reality, it is not a question of whether judicial review is limited to the weighing process, but whether judicial review is limited to the result of the weighing process. In this respect, however, all factors argue in favour of understanding as the "(more detailed) provisions (of the performance)" referred to in Section 106, first sentence GewO and Section 315 BGB only those that are not only attributable to people as declarations of intent - which is not problematic - but are also the *responsibility* of people. One may still be content to accept certain impairments of workers' interests if the consideration is deficient, but at least "the result" is right. However, this cannot justify the use of automatic decision-making systems, if only because the corresponding deficits are already inherent in them

[467] Cf. *Höpfner/Daum*, ZfA 2021, 467 (480), who only recommend "equipping instruction-issuing systems with a remonstration function and instructing employees to make use of this should an AI instruction be inequitable in their view"; as already *Göpfert/Brune*, NZA-Beil. 2018, 87 (90).
[468] For more details, see *Höpfner/Daum*, ZfA 2021, 467 (479).
[469] Thus *Höpfner/Daum*, ZfA 2021, 467 (478).
[470] Cf. *Höpfner/Daum*, ZfA 2021, 467 (478 f.).

and are therefore "structurally determined". To put it another way: While it would be one thing to hold back on the control of human decisions to a certain extent, if necessary, it is quite another to do so even if the decision is not the responsibility of humans in the first place. To decide otherwise would indeed be to open the door to chance. But there is all the less reason to do so, as it hardly seems justifiable for the employer's interest in using AI to prevail over the employee's interest in the "best possible" decision by the employer.

It should only be hinted at here that the personal connection of the employment relationship, but above all the protection of human dignity according to Article 1(1) of the Basic Law, also speak in favour of excluding automatic decisions and reserving them for humans, who, unlike AIs, are capable of "empathy and the assessment of the social consequences of their decisions".[471] Incidentally, demands by British trade union lawyers run along the same lines, aiming to establish a legal right to personal (analogue) participation as concerns decisions of considerable importance to the employee. One of the justifications for this reads: “Machines and technology are not human, and we cannot have a personal relationship with them in the same way that we can and do with other humans. [...] They can only be an aid to human interaction if the employment relationship is to remain personal and built on mutual trust and confidence. Employees are entitled to more than just a "relationship" with a machine”.[472]

Also only mentioned here in passing is the need to counter the danger that too much openness to the possibility of "machine decisions" will lead to a suppression of human judgements based on constant learning in and adaptation to complex socio-technological environments, which in the long run would be paid for with a weakening of human judgement.[473] It is important to keep in mind the fundamental difference between human judgements and "machine judgements": Human judgement is about what is "appropriate, right, good, fair or just to do in an

[471] Cf. *Nink*, Justiz und Algorithmen, p. 463: " Eine vollständig automatisierte Rechtsprechung, die den Einzelnen nur mehr als Input und Output einer formalisierten Zahlenlogik und damit als Objekt, aber nicht mehr als autonomes Individuum behandelt, ist auch mit Art. 1 Abs. 1 GG nicht in Einklang zu bringen." ("A completely automated administration of justice that treats the individual only as input and output of a formalised numerical logic and thus as an object, but no longer as an autonomous individual, cannot be reconciled with Article 1(1) of the Basic Law"); cf. also *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 96 ff. on "technikimmanenten Erkenntnisgrenzen", which the author includes, for example, social and emotional intelligence as well as common sense.

[472] Thus *Allen/Master*, Technology Managing People - the legal implications, 2021, p. 107.

[473] Cf. *Moser/den Hond/Lindebaum*, Morality in the Age of Artificially Intelligent Algorithms, 7 Apr 2021: https://doi.org/10.5465/amle.2020.0287: "[...] we offer the strong thesis that we are at risk, now, that these algorithms change, perhaps irreversibly so, our morality in fundamental ways by suppressing judgment in decision-making"; cf. also *Moser/den Hond/Lindebaum*, What Humans Lose When We Let AI Decide - Why you should start worrying about artificial intelligence now, MIT Sloan, Feb 07, 2022: https://sloanreview.mit.edu/article/what-humans-lose-when-we-let-ai-decide/.

ambiguous, troubled, problematic or puzzling situation, having explored and considered the various characteristics of that situation and having (creatively) developed and (carefully) evaluated multiple options in their respective potential to ‘better’ that situation. Judgment, therefore, requires imagination, reflection, empathy, and valuation. In judgment, it is acknowledged that data are value-laden, and that the identification of which values are relevant for decision-making is an inherent part of the process (…). Moral considerations thus inescapably come into play when developing judgment because they cannot be excluded or separated from the very situation that demands judgment.“[474] In contrast, "’reckoning’ is the processing of data through calculation and formal rationality. It relies on data as correct representations of reality (‘facts’), and values can only find their place in reckoning as stable ex ante givens, indeed a form of ‘data’. Driven by predefined rules and goals, reckoning is insensitive to context and time. [...]. In this view, the world is understood in terms of logical and ‘objective’ relationships that are fully and unambiguously defined. [...] Data and information are seen as unproblematic representations of the world, rather than – from a pragmatist viewpoint – as discriminatively selected, assembled and created with the purpose of “affording signs or evidence to define and locate a problem, and thus give a clew [sic] to its resolution".[475] Unsurprisingly, AI research is increasingly calling for collaboration with social scientists to lift the gaze of AI engineers beyond the realm of mere metrics.[476]

Due to the inability of machines to exercise discretion[477] and due to the strong personal connection of the employment relationship, discretionary decisions of machines are ruled out according to Section 106, first sentence GewO, so that they cannot effectively give instructions to people, in the view of labour law.[478]

[474] *Moser/den Hond/Lindebaum*, Morality in the Age of Artificially Intelligent Algorithms, p. 9.

[475] *Moser/den Hond/Lindebaum*, Morality in the Age of Artificially Intelligent Algorithms, p. 10, with reference to *Dewey*, The quest for certainty, 1929, p. 178.

[476] Cf. only *Bartolo/Thomas*, Qualitative humanities research is crucial to AI. https://www.fast.ai/2022/06/01/qualitative/.

[477] Cf. in this respect also *Alkhatib/Bernstein*, Street-Level Algorithms: A Theory at the Gaps Between Policy and Decisions, CHI 2019 Paper. https://doi.org/10.1145/3290605.3300760 referring to differences between human and algorithmic decision-makers: “Street-level bureaucrats are capable of making sense of new situations and then construct rationales that fill in the gaps. […] Street-level algorithms, by contrast, can be reflexive only after a decision is made, and often only when a decision has been made incorrectly. Even reinforcement learning systems, which require tight loops of feedback, receive feedback only after they take an action. Often, these algorithms only ever receive feedback after a wrong decision is made, as a corrective measure. Sometimes, algorithmic systems don’t receive corrective input at all. Algorithmic systems don’t make in-the-moment considerations about the decision boundary that has been formed by training data or explicit policies encoded into the program. Instead, the decision boundaries are effectively established beforehand, and street-level algrithms classify their test data without consideration of each case they encounter, and how it might influence the system to reconsider its decision boundary.”

[478] Likewise *Klebe,* Soziales Recht 2019, 128 (134). In view of this, it is at most conceivable to stratify the degree of human "responsibility" with regard to the potentially impaired employee interests. However, this will not be discussed further here. *Knitter*, Digitale Weisungen – Arbeitgeberentscheidungen aufgrund algorithmischer Berechnung, 2022, p. 194 considers capable of automatisation "uniform equity judgments taken within a narrow margin".

IV. Anti-discrimination law

One of the most discussed issues in connection with AI concerns the problem that AI systems can have discriminatory effects. At first glance it would seem there is something to be said for letting machines rather than humans decide in some cases.[479] After all, unlike the latter, the former are free of emotions, for example.[480] And indeed, the claim of such systems is precisely to replace fallible human judgements with neutral decision-making, or to put it in the words of the chief scientist of a US provider of workforce analytics programmes: "Let's put everything in and let the data speak for itself".[481] This is in line with the fact that in the USA, for example, automated systems are widely used and considered a central HR tool.[482] However, certain doubts have arisen about such promises, even setting aside the fact that there are inevitably people behind AI programmes.[483] One reason for such doubts is that with AI everything depends on the quality of the data with which the algorithm has been trained.[484] Whether AI can fulfil the expectations placed in it would therefore seem, to put it mildly, not to be a foregone conclusion. And indeed, it seems as if the initial euphoria about the potential of AI is giving way to increasing disillusionment.[485] For example, a recent article on the use of AI in the US healthcare system finds that "the 'promise' of AI is misleading. Without a comprehensive [...] framework that addresses biases in AI, patients that have historically not benefited from the healthcare industry will continue to face discrimination – engrained systemic biases will only become solidified, automated ones."

[479] Some time ago, the journalist Malcolm Gladwell coined the term *hiring nihilism* to sum up his conviction that all hiring decisions made by people are ultimately arbitrary; see *Sullivan*, Interviews Don't Work so Why Not be a Hiring Nihilist? Because it's all a lottery anyway, 28 Oct 2020.

[480] After all, recent AI applications are reported to counter unconscious bias by examining written evaluations of employees by their supervisors to determine whether they are more performance-based or more personality-based: https://www.textiq.com.

[481] Quoted in *Kim*, Data-Driven Discrimination at Work, William & Mary Law Review 2017, 857 (871).

[482] See *Ajunwa*, An Auditing Imperative for Automated Hiring Systems, Harvard Journal of Law & Technology 2021, 1 (12).

[483] See only *Ajunwa*, The Paradox of Automation as Anti-Bias Intervention, Cardozo Law Review 2020, 1671 (1704 et seq.).

[484] See also, for example, *Nink*, Justiz und Algorithmen, 2021, p. 167 f.: "Auch Systeme maschinellen Lernens sind nur so diskriminierungsfrei wie die Daten und Beispielsfälle, mit denen sie trainiert und gefüttert wurden".

[485] So *Takshi*, Unexpected Inequality: Disparate-Impact From Artificial Intelligence in Healthcare Decisions, Journal of Law and Health, 2021, 215 (251).

In the following, it will be shown that new risks of discrimination can also arise through partially or fully automated decision-making systems.[486] In order to better understand why this is the case, it is first necessary to take a closer look at how AI applications work and where the "gateways" for possible errors are.[487]

1. The way AI works

AI applications based on data mining exist to find correlations in existing data sets.[488] For example, one can "feed" a computer with emails that are marked as "spam" or "non-spam". These then make up what is called the training data. The computer determines which features of email messages correlate with their classification as spam. The set of correlations found is often called a "model" or "predictive model". AI thus discerns patterns and reveals regularities on which subsequent decision-making can be based. The "model" is the sum of the accumulated set of discovered relationships that can be used to automate the process of classification, estimate the value of unobserved variables, or predict future outcomes. To return to the example: Familiarise the algorithms with examples of spam that contain certain terms or phrases ("You won") and it will learn which related content also points to the features or outcomes of interest, that are being searched for, called the target variable. In contrast, so-called "class labels" transfer all possible values of the target variable into mutually exclusive categories. The task of the programmer is now to translate a problem into formal terms so that they can be analysed by computers. Herein, it is often said, lies the "art" of data mining.[489] Project goals and requirements must be transformed into a "data mining problem definition". Here, the definition of the target variables and the class labels determines which results are achieved. This does not pose any major difficulties for the question of "spam" or "non-spam". However, if a program is to determine a person's "creditworthiness" or to filter out "good workers", the task becomes much more complex: "Good" must be defined in a way that corresponds

[486] Thus *Orwat*, Diskriminierungsrisiken durch Verwendung von Algorithmen, 2020, p. 22: "durch teil oder vollautomatisierte Entscheidungssysteme [können sich] auch neue Risiken der Diskriminierung ergeben". See also *Parviainen*, Can algorithmic recruitment systems lawfully utilise automated decision-making in the EU?, ELLJ 2022, 225.

[487] See in particular *Barocas/Selbst*, California Law Review 2016, 671 (679), on whose work the following brief outline is primarily based, but see also, for example, *Orwat*, Diskriminierungsrisiken durch Verwendung von Algorithmen, p. 76 et seq, *Zuiderveen Borgesius*, Discrimination, artificial intelligence, and algorithmic decision-making, Council of Europe, 2018; *Sullivan*, Employing AI, Seton Hall Public Law Research Paper 2018; *Kim*, Data-Driven Discrimination at Work, William & Mary Law Review 2017, 857 (883 et seq.).

[488] There are reports of newer approaches, however, that are supposed to reduce discrimination risks with the help of causal inferences; cf. on this only *Nink*, Justiz und Algorithmen, 2021, p. 170 with further references.

[489] See *Barocas/Self*, California Law Review 2016, 671 (678).

to measurable results (e.g. productivity). But there is also the question of whether workers should only be classified as "good" or "bad" or whether a comprehensive ranking should be established. Employers may have criteria in mind that show whether an employee is "good". But these will hardly ever be exhaustive. There is also a danger that target variables and class labels are chosen in such a way that the application simply locks in certain assessments.[490] In other words, if the definitions are already "flawed", the application of AI will almost inevitably lead to discriminatory effects. [491]

Further difficulties arise, as already mentioned, when computers are "trained" with "problematic" training data. It is not even necessary to consider that data may reflect biases on the part of the programmers;[492] in this case one often speaks of machine bias.[493] Quite independent of this, there is the danger that previous biases will be solidified and possibly even reinforced - due to so-called feedback loops.[494] On this, the literature sometimes cites the (real) example of a program in which applicants for medical studies were to be sorted on the basis of previous admission decisions, but it then turned out that ethnic minorities and women were systematically disadvantaged in the process.[495] The automated process thus perpetuated existing prejudices - without the users' input or knowledge.[496] It is also conceivable that AI in a sense prolongs existing biases. For example, if a program makes hiring recommendations based on interest in certain types of candidates, it will ultimately produce results that not only reflect, but even perpetuate, any

[490] See *Barocas/Selbst*, Big Data's Disparate Impact, California Law Review 2016, 671 (679): "These may seem like eminently reasonable things for employers to want to predict, but they are, by necessity, only part of an array of possible definitions of 'good'. An employer may instead attempt to define the target variable in a more holistic way-by, for example, relying on the grades that prior employees have received in annual reviews, which are supposed to reflect inherit the formalizations involved in preexisting assessment mechanisms, which in the case of human-graded performance reviews, may be far less consistent".

[491] For more details, see *Barocas/Self*, Big Data's Disparate Impact, California Law Review 2016, 671 (679).

[492] See *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 87: "The nimbus of objectivity of mathematical formula language that surrounds the automation of the process or its delegation to a technical assistance system therefore often conceals the fact that the decision patterns are subject to subjective bias". Accordingly, the lack of diversity in development teams is increasingly considered a problem; cf. only Stanford University Human-Centered Artificial Intelligence, Artificial Intelligence Index Report 2021, p. 137: "The AI workforce remains predominantly male and lacking in diversity in both academia and the industry, despite many years highlighting the disadvantages and risks this engenders. The lack of diversity in race and ethnicity, gender identity, and sexual orientation not only risks creating an uneven distribution of power in the workforce, but also, equally important, reinforces existing inequalities generated by AI systems, reduces the scope of individuals and organisations for whom these systems work, and contributes to unjust outcomes".

[493] Foundational *Friedman/Nissenbaum*, Bias in Computer Systems, ACM Transactions on Information Systems, 1996, 330; cf. on this also *Nink*, Justiz und Algorithmen, 2021, p. 170 ("Algorithmen als Projektion der Werteinstellungen ihrer Schöpfer"), who also points out that especially later optimisation processes offer "gateways for subjective elements" and that "programmers and algorithm designers [...] usually have no legal or ethical training"; ibid. p. 174.

[494] See for example *Kim*, Data-Driven Discrimination at Work, William & Mary Law Review 2017, 857 (882).

[495] See *Barocas/Self*, Big Data's Disparate Impact, California Law Review 2016, 671 (682).

[496] AI researchers are working on applications to detect this bias; one example is Facebook's Fairness Flow program; sceptical, however, *Wiggers,* AI experts warn Facebook's anti-bias tool is 'completely insufficient', 31 Mar 2021: https://venturebeat.com.

existing bias on the part of employers.[497] Also, training data fed into the AI application may itself be "flawed" or biased. The reasons for a possible "flaw" are manifold. For example, it may be that fewer and/or less timely records exist for certain groups of people than for others from the outset. But even a data set with individual records of consistently high quality can suffer from statistical distortions that result in different groups not being represented in correct proportions.[498] For example, members of a certain group may have limited access to the internet and therefore find it more difficult to indicate their interest in and qualifications for a job advertised on the internet than members of other groups.[499] Distortions may also occur when members of certain groups are overrepresented in a data set. An example mentioned in the literature is that the behaviour of members of certain groups is monitored particularly closely by their superiors and their offences are therefore disproportionately found in the corresponding data records.[500] At the same time, all this raises the fundamental problem that AI is "past-oriented" or "backward-looking" because it necessarily works with historical data. [501]

Another problem arises from the selection of attributes and the weight given to them (feature selection). An example may illustrate the problem: In hiring decisions, great importance is attached to the reputation of the college or university that an applicant has attended. However, if members of a certain group attend these institutions far less frequently or, for whatever reason, graduate from them less frequently, they are systematically disadvantaged. Another example: If only the degree from a particular university is taken into account as such, and if, for example, the degree grade and duration of study are ignored, the result is that ultimately the best-qualified candidate is not selected. This example shows that in feature selection it is crucial to take the context into account and find the right balance between features and the size of the data set.[502]

[497] *Barocas/Self*, Big Data's Disparate Impact, California Law Review 2016, 671 (683).

[498] *Barocas/Self*, Big Data's Disparate Impact, California Law Review 2016, 671 (684).

[499] *Barocas/Self*, Big Data's Disparate Impact, California Law Review 2016, 671 (685).

[500] *Barocas/Self*, Big Data's Disparate Impact, California Law Review 2016, 671 (687).

[501] Cf. *Nink*, Justiz und Algorithmen, 2021, p. 171: "Für den Betroffenen wächst es sich zum schwer durchschaubaren Problem aus, wenn unreflektiert ältere Daten in eine Entscheidung einfließen, die nunmehr aus dem ursprünglichen Kontext gerissen oder eine zwischenzeitliche Veränderung und Entwicklung des Betroffenen nicht mehr abdeckt." ("For the person concerned, it becomes a problem that is difficult to understand if older data is unreflectedly incorporated into a decision that is now torn out of its original context or no longer covers a change and development of the person concerned in the meantime"); cf. also *Mayson*, Bias In, Bias Out, YALE L.J. 2019, 2218 (2224): "All prediction functions like a mirror. Algorithmic prediction produces a precise reflection of digital data. Subjective prediction produces a cloudy reflection of anecdotal data. But the nature of the analysis is the same. To predict the future under status quo conditions is simply to project history forward"; see also *Chander*, The Racist Algorithm?, Michigan Law Review 2017, 1023 (1034).

[502] Lernen wie Maschinen: Was ist algorithmische Voreingenommenheit (Algorithmic Bias)?: https://www.lernen-wie-maschinen.ai/ki-pedia/was-ist-algorithmische-voreingenommenheit-algorithmic-bias/. The example mentioned can also be found here.

While the error-proneness of AI in the cases mentioned so far is ultimately based to a large extent on "human error", errors can also arise from the way AI itself works. It is important to realise that AI applications, insofar as they use data mining, analyse data with statistical techniques to uncover patterns. There may be a causal relationship between correlating data, but the relationship may also be purely coincidental. In any case, the AI application is only "interested" in the correlation as such. As long as the discovered relationships can be considered robust, the data model will use them to classify or predict future cases.[503] Although data mining cannot explain the relationship, a model will predict that applicants who exhibit a certain trait will be better workers, and recommend their selection to the employer.[504] For example, there is a report of software that based its recommendation on the fact that particularly qualified applicants had conspicuously visited a certain Japanese manga site. The fact that this is highly problematic, however, is almost palpable when one considers that a non-Japanese person is highly unlikely to visit such a site to begin with.[505] The discriminatory effect is based on the fact that the program is linked to a characteristic that is also a "proxy" for a certain group membership. This group membership is coded in other data.[506] In this respect, it is only partially reassuring when companies advertise their AI products with the fact that they constantly change the variables that are considered important. One might see this as proof of the adaptability and flexibility of an application. But it is at least as likely that the predictions made by algorithms are of limited value because they often only capture temporary, random correlations. If they depicted causal relationships, they would be more stable.[507]

A paper published some time ago on the susceptibility of AI and Big Data to discrimination contains a "taxonomy" of biases – including those already outlined here, but also others - that can creep in during the various stages of making an AI system, from data creation and problem formulation to data preparation and

[503] Critical of this is *Smith*, High-tech redlining: AI is quietly upgrading institutional racism: How an outlawed form of institutionalised discrimination is being quietly upgraded for the 21st century: https://www.fastcompany.com: "Is our faith in computers so blind that we are willing to trust algorithms to reject job applications and loan applications, set insurance rates, determine the length of prison sentences, and put people in internment camps? Favoring some individuals and mistreating others because they happen to have irrelevant characteristics selected by a mindless computer program isn't progress: it's a high-tech return to a previous era of unconscionable discrimination".

[504] See also *Kim*, Data-Driven Discrimination at Work, William & Mary Law Review 2017, 857 (875).

[505] See *Smith*, High-tech redlining: AI is quietly upgrading institutional racism: How an outlawed form of institutionalised discrimination is being quietly upgraded for the 21st century: https://www.fastcompany.com.

[506] Cf. *Barocas/Self*, Big Data's Disparate Impact, California Law Review 2016, 671 (691 f.); for a German perspective see *Buchholtz/Scheffel-Kain*, NVwZ 2022, 612.

[507] See *Smith*, High-tech redlining: AI is quietly upgrading institutional racism: How an outlawed form of institutionalised discrimination is being quietly upgraded for the 21st century: https://www.fastcompany.com: "the algorithm captures transitory coincidental correlations that are of little value. If these were causal relationships, they would not come and go. They would persist and be useful".

analysis. The paper is all the more impressive because it was written by and for practitioners. The authors are not even concerned with the design of fair algorithms. Rather, it is intended to help ML developers avoid errors in the various project phases and, in particular, to raise their awareness of the error-proneness of the systems.[508] The authors distinguish between different forms of bias:[509] Sampling bias is characterised by the fact that the selected data is not representative of reality, such as when a face recognition algorithm is "fed" with more photos of light-skinned people than of dark-skinned people. Measurement bias occurs, for example, when photographers only provide views of objects from certain perspectives when creating image and video data sets. Label bias is based on inconsistencies in the assignment of labels that occur when different people assign different labels to the same object type.[510] Bias can also occur in problem formulation. Depending on how a problem is formulated and how the information is presented, the results obtained may be different and possibly biased (framing effect bias). For example, a program designed to estimate or predict the future creditworthiness of customers will have business specifications that determine how "creditworthiness" is defined for the purposes of the program. However, bias can also occur in the algorithm or during data analysis. Sample selection bias occurs when the samples chosen are not representative of the population being analysed. Confounding bias in the AI model occurs when the algorithm learns the wrong relationships by not taking into account all the information in the data or when it "overlooks" the relevant relationships between features and target outputs. A design-related bias occurs when biases arise as a result of algorithmic limitations or other system limitations such as lack of computer power. Finally, possible biases related to the evaluation and validation of an AI model's performance need to be considered. These range from the various forms of human evaluation bias – which is based on the fact that humans are at work and can, for example, make confirmation errors – to sample treatment bias, in which the test sets selected for the evaluation of an algorithm are biased, to various forms of validation and test dataset bias, which is ultimately characterised by the fact that errors in data generation can reappear in the model evaluation phase.[511]

[508] *Srinivasan/Chander*, Biases in AI Systems - A Survey for Practitioners, acmqueue 2021, 47 (48); cf. most recently also *Mehrabi/Morstatter/Saxena/Lerman/Galstyan*, A Survey on Bias and Fairness in Machine Learning, 25 Jan 2022: https://arxiv.org/pdf/1908.09635.pdf.; cf. also European Parliament, AIDA Working Paper on AI and Bias, November 2021.

[509] *Srinivasan/Chander*, Biases in AI Systems - A Survey for Practitioners, acmqueue 2021, 47 (48 et seq.) with further references.

[510] The so-called negative set bias will be excluded here because it is related to the image recognition that is less relevant here.

[511] The competent US regulatory authority has recently presented practical guidelines for eliminating bias in AI systems, favouring a socio-technical approach; cf. National Institute of Standards and Technology (NIST), Towards a Standard for Identifying and Managing Bias in Artificial Intelligence, NIST Special Publication 1270, March 2022.

Summarising the considerations at this point, it emerges that while human decision-making is certainly not perfect, AI applications, for their part, are, to put it cautiously, far more prone to error than might be assumed at first glance.[512] At the same time, it often seems anything but easy to eradicate existing distortions.[513]

In this context, it is also worth recalling the above-mentioned Recommendation of the Council of Ministers of the Council of Europe of 8 April 2020, which stated that "most algorithmic systems are based on statistical models in which errors form an inevitable part".[514] The fact that some of the shortcomings have the effect of inviting job applicants (for example) to devise circumvention strategies (by avoiding certain terms in the application letter or, conversely, specifically including them),[515] does not make matters any better. As one expert put her reservations: "The fact remains that there are myriad of ways that automated hiring could systematically replicate biases that have calcified from organizational practice".[516] In this context, reservations about AI and Big Data exist even if one does not assume bad faith on the part of the developers.[517] After all, there is every reason to "take a close look" at AI applications, also from the point of view of anti-discrimination law. Even if one disregards the fact that deliberate discrimination in AI applications is relatively easy to conceal (masking),[518] there is always a risk of discriminatory effects when using AI.

2. Discrimination problems using the example of "AI recruiting"

At the centre of the discussion on "discriminatory AI" is the use of AI applications in recruitment procedures.[519] Already today, such applications are used to a

[512] For a comprehensive analysis of existing programs, see *Raghavan/Barocas/Kleinberg/Levy*, Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices, 2020.

[513] Cf. only *Quach*, AI models still racist, even with more balanced training, 1 May 2022. https://www.theregister.com/2022/05/01/ai_models_racist/?tpcc=nleyeonai; see also *Jingwei Li/Danilo Bzdok/Jianzhong Chen* et al., Cross-ethnicity/race generalization failure of behavioral prediction from resting-state functional connectivity, Science Advances 2022, 144. DOI: 10.1126/sciadv.abj18.

[514] Recommendation CM/Rec(2020)1 of the Committee of Ministers to Member States on the human rights impacts of algorithmic systems (Adopted by the Committee of Ministers on 8 April 2020 at the 1373rd meeting of the Ministers' Deputies (under A5.).

[515] See *Ajunwa*, An Auditing Imperative for Automated Hiring Systems, Harvard Journal of Law & Technology 2021, 1 (21 f.). Cf. *Hern*, Amazon's Alexa could turn dead loved ones' voices into digital assistant – Technology promises ability to 'make the memories last' by mimicking the voice of anyone it hears, 23 June 2022. https://www.theguardian.com/technology/2022/jun/23/amazon-alexa-could-turn-dead-loved-ones-digital-assistant?tpcc=nleyeonai.

[516] So *Ajunwa*, An Auditing Imperative for Automated Hiring Systems, Harvard Journal of Law & Technology 2021, 1 (16).

[517] See only *Johnson*, Automating the Risk of Bias, The George Washington Law Review 2019, 1214 (1221): "Even when well-intentioned developers aspire to create ADM platforms that are more inclusive, bias may creep in and compromise the outcomes".

[518] See also *Barocas/Self*, Big Data's Disparate Impact, California Law Review 2016, 671 (693).

[519] Cf. also *Söbbing*, InTer 2018, 64.

considerable extent in HR work and many more companies are dealing with the question of using AI.[520] The risk of discrimination is palpable. It is therefore not surprising that the academic debate on discriminatory AI deals almost exclusively with the use of AI vis-à-vis job applicants.

There is a lot of discussion in the literature about automated personality tests and telephone interviews or job interviews conducted by chatbots, which, for example on the basis of an AI-controlled speech analysis, are supposed to provide information about the psychological state of a candidate. The fact that the use of such instruments threatens to discriminate, whether on the basis of a disability or an ethnic origin etc, needs no explanation.[521] The risk of indirect discrimination within the meaning of Section 3(2) of the General Act on Equal Treatment (AGG) must be taken particularly seriously, because although on the surface a neutral characteristic is used, it is fulfilled by members of a certain group of people more frequently than by others and thus the applicant is ultimately disadvantaged on grounds mentioned in Section 1 AGG. Thus, a certain emotional state of a person is in itself a neutral characteristic. However, if an AI application focuses on this, it constitutes indirect discrimination if it is shown that, for example, people with disabilities exhibit this state of mind significantly more often than others. [522]

Applications that "pre-screen" applicants deserve at least as much scrutiny as automated personality tests and the like, even if the system does not go so far as to "recommend" the hiring of certain persons to the potential employer. In this respect, there are widespread reservations in the literature about the restriction of the AGG to the characteristics listed in Section 1 AGG: If, for example, an algorithm sorts out applicants from a certain district,[523] this may constitute (indirect) discrimination on grounds of social origin. However, this is not prohibited under the AGG - in contrast in particular to ILO Convention No. 111 on Discrimination in Employment and Occupation of 1958.[524] Against this background, it is no longer surprising when, in view of the analytical capabilities of AI systems, demands are made to search through the grounds of discrimination of the AGG and, if necessary, to create new grounds of discrimination, or at least to align them with

[520] Künstliche Intelligenz in der Personalarbeit Netzwerk Weiterbildung Interessenvertretung Information www.bpm.de Evaluation of the survey 30 Apr 2019; available at: https://www.bpm.de/sites/default/files/20190429_auswertung_bpm-pressemitteilung_final_0.pdf; cf. also *Freyler*, NZA 2020, 284 (285).

[521] Both examples in *Dzida/Groh*, NJW 2018, 1917 (1919).

[522] Cf. again *Dzida/Groh*, NJW 2018, 1917 (1919).

[523] Ex. again after *Dzida/Groh*, NJW 2018, 1917 (1919).

[524] Art. 1 No. 1a of the Convention.

the evaluations of the GDPR.[525] Indeed, the use of AI can reinforce inequalities, even if no protected groups of persons are affected.[526]

In the following, however, the focus will not be on legal policy considerations and demands, but rather on illuminating the problems that the use of AI entails under current law, thus *de lege lata.*

a) Existence of "treatment" within the meaning of Section 3 AGG

Initially one encounters the problem that Section 3 AGG requires an action or, more precisely, "treatment". According to some authors, there are doubts as to whether this is the case when AI is used, as "algorithm-based discrimination" does not constitute treatment within the meaning of Section 3 AGG.[527] However, this cannot be followed, as the required treatment is simply to be seen in the decision that the HR manager makes on the basis of the AI and with which he or she adopts its "weighing".[528] The "action or omission that emanates from human behaviour" required by Section 3 AGG[529] is thus certainly given.[530]

b) Subjective facts

Further concerns expressed in the literature focus on the subjective facts. In fact, in many cases not even the employer itself may have knowledge of the characteristics that were decisive for the result achieved by the algorithm; this may even be the case if the employer itself developed the algorithm, but the algorithm then "developed itself further". However, there is widespread consensus that

[525] See only *Wachter/Mittelstadt/Russell*, Why Fairness cannot be automated Bridging the Gap between EU Non-Discrimination Law and AI, 1 (11 f.): "Groups which do not map to a legally protected characteristics may suffer levels of disparity which would otherwise be considered discriminatory if applied to a protected group. These new patterns of disparity may force legislators and society to re-consider whether the scope of non-discrimination remains broad enough to capture significant disparity as caused not only by humans and organisations, but machines as well".

[526] See only *Zuiderveen Borgesius*, Strengthening legal protection against discrimination by algorithms and artificial intelligence, The International Journal of Human Rights 2020, 1572. See also Wachter, The Theory of Artificial Immutability: Protecting Algorithmic Groups under AntiDiscrimination Law (February 15, 2022). Tulane Law Review, Forthcoming. https://ssrn.com/abstract=4099100.

[527] Thus *Steege*, MMR 2019, 715 (718).

[528] In the result, also *Dzida/Groh*, NJW 2018, 1917 (1919). The fact that a decision is made directly by the AI system itself is unlikely to be the case in practice, but it hardly poses a problem because it is not obvious why the employer should not be attributed in such a case.

[529] Thus *BeckOKArbR/Roloff*, 64th ed., § 3 AGG marginal No. 2. *Sesing/Tschech*, MMR 2022, 24 (26) rightly point out that the protection against discrimination under the AGG is "technology-neutral" and thus also applies to "discriminatory AI".

[530] Contrary to *Steege*, MMR 2019, 715 (718), the "output of the AI" only becomes effective if the potential employer relies on it.

knowledge of the "frowned-upon" feature on the part of the employer is a prerequisite for discrimination within the meaning of section 3 AGG.[531] However, there is every reason to attribute the knowledge of the machine to the employer - if one wants to see the machine as an independent "bearer of knowledge" at all - as one's own knowledge in accordance with Section 166 of the German Civil Code (BGB), if the employer makes use of the machine in the selection of applicants.[532] This cannot be elaborated on here. However, it should be pointed out that for the dogmatic justification of the imputation of knowledge both the aspect of the so-called "knowledge responsibility" of the principal[533] and the circumstance that a "splitting of knowledge" cannot lead to unjustified privileges are important.[534] However, the "machine knowledge" can be attributed under one point of view as well as the other.[535] Moreover, if one did not want to follow this, it would be tantamount to an unjustifiable "carte blanche" under discrimination law for employers who use AI in recruitment procedures.[536]

c) Causality

In the literature, one occasionally encounters the assessment that "due to the size of the data set, it is to be expected that the decision is not based on only one characteristic".[537] This will be true in many cases. After all, the "attraction" of using AI and Big Data lies precisely in their ability to recognise complex patterns in a vast amount of data, so that it should be almost the rule if the result achieved by the AI system is based on more than one feature. In this respect, an "analogy to the cases of a bundle of motives" is used in the literature,[538] whereby, according to the case law, it is indeed sufficient for the affirmation of causality if the impermissible feature has only influenced the decision.[539]

[531] Cf. *Lewinski/de Barros Fritz*, NZA 2018, 620 (622).

[532] Accurately *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 354, who bases an attribution of knowledge on § 166 BGB analogue and furthermore refers to BAG, NZA 2009, 79 (and para. 39), where the court - with regard to the severely disabled status of an applicant - decisively focused on the possibility of acquiring knowledge (in the "sphere of influence" of the employer); cf. ibidem, marginal no. 35: "Every employer must organise the handling of its personnel matters in such a way that it can fulfil its legal obligations to promote severely disabled applicants"; for an attribution to the employer, most recently also *Sesing/Tschech*, MMR 2022, 24 (26).

[533] Cf. only MünchKomm/Schubert, 9th ed. 2021, § 166 BGB marginal no. 60 with further references.

[534] Cf. also MünchKomm/Schubert, 9th ed. 2021, § 166 BGB marginal no. 61 with further references.

[535] In favour of unrestricted knowledge attribution under the aspect of qualifying AI as a "knowledge agent" most recently *Kuntz*, ZfPW 2022, 177.

[536] See also *Starker*, in: *Hoeren/Sieber/Holznagel* (eds.), Handbuch Multimedia-Recht, 57th ed. 2022, Part 15.6 Big Data und Arbeit, para. 49. Another question is whether and to what extent one assumes that there is an obligation to carry out so-called *algorithmic audits* (cf. also on this point the author, op. *cit.,* para. 50) and attaches discriminatory effects to their absence.

[537] *Freyler*, NZA 2020, 284 (287); similarly *Lewinski/de Barros Fritz*, NZA 2018, 620 (622).

[538] Thus *Lewinski/de Barros Fritz*, NZA 2018, 620 (622).

[539] Cf. only BeckOK/Roloff, § 3 AGG marginal no. 16 with further references.

d) Indirect discrimination

In the literature, problems are also caused in particular by the question of under which conditions the "decision" by an AI could be seen as indirect discrimination within the meaning of Section 3(2) AGG and, if this were to be affirmed, could it be justified according to the conditions also mentioned in Section 3(2) AGG.[540] The difficulties are caused by the fact, already mentioned above, that the aim of AI is to uncover correlations and classify applicants on this basis. If it turns out that members of a group protected under the AGG are overrepresented in the group formed by the AI, the danger of inadmissible indirect discrimination is clear.[541] This applies all the more since, as already stated above, according to the case law, it is sufficient for the existence of indirect discrimination if the proscribed characteristic has only influenced the decision.[542] If one also considers that the (supposed) added value of the use of AI and Big Data consists precisely in detecting correlations that remain hidden for humans due to the volume of the underlying data, then it becomes clear that the susceptibility of AI to indirect discrimination is not only *a problem,* but may even be *the central problem* that arises in the present context.[543] It is perhaps best illustrated by the example of the manga site mentioned at the beginning of this section: anyone who favours visitors to such a site in job applications is inevitably discriminating on the basis of ethnic origin (and possibly also gender).[544]

It should be clear that in such cases there will often be indirect discrimination. The question that therefore arises is whether and under what conditions this is justified. In this respect, Section 3(2) AGG requires that an "apparently neutral provision,

[540] See also *Sesing/Tschech,* AGG und KI-VO-Entwurf beim Einsatz von Künstlicher Intelligenz, MMR 2022, 24 (26) on the distinction between indirect and refined discrimination.

[541] Thus *Lewinski/de Barros Fritz*, NZA 2018, 620 (622); similarly *Freyler*, NZA 2020, 284 (288).

[542] Cf. on the latter only *Schrader*/*Schubert*, in: Däubler/Beck (eds.), Allgemeines Gleichbehandlungsgesetz mit Entgelttransparenzgesetz, Berliner LADG, 5th ed., 2021, § 3 AGG marginal no. 83 with further references.

[543] Cf. in this respect also *Straker/Niehoff*, ABIDA-Fokusgruppe - Diskriminierung durch Algorithmen und KI im eRecruiting, ZD-Aktuell 2018, 06252; cf. also Council of Europe, Discrimination, artificial intelligence, and algorithmic decision-making, 2018, p. 10 et seq.; *Xenidis*/*Senden*, EU non-discrimination law in the era of artificial intelligence: Mapping the challenges of algorithmic discrimination, in *Bernitz* et al. (eds.), General Principles of EU law and the EU Digital Order, 2020, 151, 20 ff. SSRN: https://ssrn.com/abstract=3529524. *Adams-Prassl*/*Binns*/*Kelly-Lyth*, Directly Discriminatory Algorithms, Modern Law Review 2022 argue that direct discrimination is also significant.

[544] See also *Kim*, Data-Driven Discrimination at Work, William & Mary Law Review 2017, 857 (865) f.: "The nature of algorithmic decision-making raises particular concern when employers rely on these models to make personnel decisions. Data mining techniques used to build the algorithms seek to uncover any statistical relationship between variables present in the data, regardless of whether the reasons for the relationship are understood. As a result, if employers rely on these models, they may deny employees opportunities based on unexplained correlations and make decisions that turn on factors with no clear causal connection to effective job performance".

criterion or practice [must be] objectively justified by a legitimate aim and that the means [...] used to achieve that aim are appropriate and necessary".[545] Thus, for example, in the case of an employer who is interested in the fact that an applicant's vocational qualification does not date back too far, in simple terms, the question is whether he can refer to a justified interest or to a "real need of the company"[546] which, moreover, prevails over the interests of the disadvantaged persons. In a case decided by the BAG, which was in fact based on exactly this factual situation, it was therefore necessary to examine, among other things, whether the requirement formulated by the employer was "necessary and appropriate for the best possible completion of the work".[547]

The fact that all this does not "fit" the use of AI and Big Data hardly needs explanation. It is probably even less decisive that in the one case the employer formulates a requirement autonomously, while in the other a machine "derives" its validity from data analysis. It is much more significant that it is not the same thing whether a requirement exists because, from the employer's point of view, it "corresponds to a real need of the company"[548] (which then has to be verified in court), or whether it exists because statistical evidence makes it appear valid without, to put it casually, the machine or even the employer "overthinking it".[549] To return to the "manga example": To say the least, it seems far-fetched to assume that visiting a certain website can provide information about labour productivity or professional success. Accordingly, this case gives every reason to "question" the correlation. However, this is a matter fundamentally different from the examination indicated under Section 3(2) AGG, which asks whether there is a "real need" for a particular requirement and, moreover, whether there is not in fact a reason to fear an excessive impairment of the legitimate interests of the persons protected under the AGG. No one will claim that, for example, an applicant's frequent visit to a manga site corresponds to a "real need".

[545] On the fact that this already excludes the existence of indirect discrimination in the affirmative, for example BeckOK/Roloff, § 3 AGG marginal no. 18.

[546] Cf. only NZA 2017, 715 (and para. 38) and reference to ECJ, AP EC Art. 138 No. 2.

[547] BAG, NZA 2017, 715 (and para. 44).

[548] Cf. also on this NZA 2017, 715 (and marginal no. 38).

[549] Cf. only *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 353: "Indirect discrimination is often not foreseen by the manufacturers and operators of a software application themselves and may not even be recognised later".

The problem can be further illustrated if one also takes a look at US law and in particular at the legal institution of disparate impact discrimination,[550] which ultimately also underlies the protection against indirect discrimination in EU law.[551] In this respect, too, it is argued that the standards resulting therefrom for a possible justification of unequal treatment cannot be transferred to AI and Big Data, where it is solely a matter of "min[ing] the available data, looking for statistical correlations that connect seemingly unrelated variables, such as patterns of social media behavior, with workplace performance".[552] Accordingly, the employer is in any case required to do more than merely point out the existence of such a correlation.[553]

If we return to German law, we have to answer the question of what consequences it has for the justification of indirect discrimination by AI if Section 3(2) AGG does not open up the possibility for it (which is indeed the case according to the foregoing). This question is relatively easy to answer. In view of the lack of justification, *de lege lata* there is no other option than to assume the existence of inadmissible indirect discrimination in such cases. A different result could only be reached if one assumed that the AGG is incomplete because the use of AI is either a special form of discrimination not covered by the applicable law[554] that requires a specific possibility of justification, or in any case the existence of a possibility of justification must be required which also covers "indirect discrimination by AI". However, there is no reason why the existence of indirect discrimination should be denied only because the corresponding "decision" was made by an AI application. And also from the point of view of the lack of a justification tailored to the specific circumstances of AI, a legal loophole seems difficult to justify, though it is also true

[550] A special case is so-called *proxy discrimination*, where the benefit of an apparently neutral practice for the discriminator results at least in part from the fact that it produces a disparate impact; see *Prince/Schwarcz*, Proxy Discrimination in the Age of Artificial Intelligence and Big Data, Iowa Law Review 2020, 1257. Platforms' algorithms for the flexible determination of fares are examined from the perspective of disparate impact *Pandey/Caliskan*, Disparate Impact of Artificial Intelligence Bias in Ridehailing Economy's Price Discrimination Algorithms: https://dl.acm.org/doi/pdf/10.1145/3461702.3462561.

[551] Cf. only *Thüsing*, NZA 2000, 570 (570).

[552] See *Kim*, Data-Driven Discrimination at Work, William & Mary Law Review 2017, 857 (866): "Classification bias may seem amenable to challenge under disparate impact doctrine, which targets facially neutral employment practices that have disparate effects on racial minorities or other protected classes. However, a mechanical application of existing disparate impact doctrine will fail to meet the particular risks that workforce analytics pose. That doctrine evolved to address employer use of tests purporting to measure workers' abilities, and therefore focused on the validity of those measures and their relevance to a particular job. In contrast, data mining models do not rest on psychological or any other theories of human behaviour. [...]. As a result, they pose a different set of risks-risks that existing doctrine does not address well".

[553] See again *Kim*, Data-Driven Discrimination at Work, William & Mary Law Review 2017, 857 (916 et seq.): "Under disparate impact doctrine, an employer may defend against a prima facie showing of disparate impact by demonstrating that the challenged practice is "job related [...]. and consistent with business necessity." The exact meaning of this phrase is ambiguous, and the standard has proven difficult to apply consistently in practice. When applied to data analytics, however, it is difficult to make sense of the standard at all. When an algorithm relies on seemingly arbitrary characteristics or behaviours interacting in some complex way to predict job performance, the claim that it is "job related" often reduces to the fact that there is an observed statistical correlation. If a statistical correlation were sufficient to satisfy the defence of job-relatedness, the standard would be a tautology rather than a meaningful legal test. In order to protect against discriminatory harms, something more must be required to justify the use of an algorithm that produces biased outcomes".

[554] See *Kim*, Data-Driven Discrimination at Work, William & Mary Law Review 2017, 857 (925) for US discrimination law.

that in the current law no standards exist by which to close them. Accordingly, it must suffice that those who use AI and Big Data run the risk of indirectly discriminating against protected persons, which cannot be justified under current law. Admittedly, this is of little help to the disadvantaged party, as will be shown in a moment,[555] in view of the existing rules on the distribution of evidence. However, the fact that the user of AI "turns a blind eye" to the risk of indirect discrimination appears to be a point of view that could justify the easing of the burden of proof in favour of the potentially injured party.[556]

e) Fault

As far as the legal consequences of discrimination are concerned, the requirement of fault under Section 15(1) AGG for claiming damages is particularly difficult, though it is presumed, as is well known.[557] Liability of the potential employer then depends on whether Section 278 BGB applies to "machines" (or to algorithms) by analogy.[558] This general question will not be explored in depth here. However, it should not be overlooked that trying to answer it raises serious difficulties. These arise less from the fact that, as things stand now, AI applications are for the most part denied their own legal capacity,[559] since this would not prevent one from assuming a partial legal capacity precisely for the purposes of Section 278 BGB.[560] What it is more important is that Section 278 BGB requires fault on the part of the vicarious agent, but it is difficult to accuse AIs of subjective fault.[561]

[555] Cf. GG. IV. 2f).

[556] See also *GrimmelmannWestreich*, Incomprehensible Discrimination, California Law Review Online 2017, 164 (176): "A test that turned only on the employer's knowledge of how its model functions would discourage employers from looking too closely at models that superficially seemed to work. Where a model has a disparate impact, our test in effect requires an employer to explain why its model is not just a mathematically sophisticated proxy for a protected characteristic".

[557] Cf. only ErfKomm/Schlachter, 22nd ed. 2022, § 15 AGG marginal no. 6.

[558] Affirmatively *von Lewinski/de Barros Fritz*, NZA 2018, 620 (623); in the result also *Dzida/Groh*, NJW 2018, 1917 (1920), according to which it should not make a difference whether the employer uses natural persons or software programs; in agreement ErfK/Schlachter, 21st ed. 2021, § 15 AGG marginal no. 9.

[559] Cf. only *Freyer*, NZA 2020, 284 (286 f.).

[560] See only *Lampe*, in: Hoeren/Sieber/Holznagel (eds.), Handbuch Multimedia-Recht, September 2021, Part 29.2 KI im Zivilrecht para. 41; also *Riehm*, in: Kaulartz/Braegelmann (eds.), Rechtshandbuch Artificial Intelligence and Machine Learning, 2020, p. 235 et seq.

[561] Cf. also insofar *Lampe*, in: Hoeren/Sieber/Holznagel (eds.), Handbuch Multimedia-Recht, September 2021, part 29.2 KI im Zivilrecht marginal no. 14.

f) Burden of proof

There are particularly intense discussions in the literature regarding the burden of proof. As one widespread opinion has it, in the face of the often opaque workings of AI applications and the complexity of algorithms, applicants regularly run into difficulties, despite the applicability of Section 22 AGG, demonstrating even circumstantial evidence of discrimination.[562] One thing that is often pointed out in this context is that algorithms commonly take a multitude of characteristics into consideration, but under these circumstances the "decision-making processes" can no longer be traced at all. AI is thus not only "susceptible to discrimination", but also makes it difficult to uncover errors in retrospect. In general, it is complained that the full "transparency risk" is imposed on the protected subjects, without there being any way to change this under current law.[563] In all of this, it must also be taken into account that although the protected subject can in the framework of claiming indirect discrimination rely on statistics that demonstrate a regular but also significantly greater adverse impact,[564] case law is relatively reluctant to find statistical data robust enough to trigger a reversal of the burden of proof.[565] The "personalisation logic" of AI also makes it almost impossible to provide statistical evidence that can be based on a sufficiently large comparison group.[566] It is little consolation for potential victims of discrimination that the informational duties of Article 13 f of the GDPR and the right to information under Article 15 of the GDPR come to their aid:[567] they only cover the "basic logic" of the processing and does not necessarily include a comparison with others, which is a prerequisite for establishing unequal treatment.[568] It must also be taken into account that in many cases not even the employer itself will be able to make the decision comprehensible to a rejected job applicant.[569]

[562] See only *von Lewinski/de Barros Fritz*, NZA 2018, 620 (622); also *Orwat*, Diskriminierungsrisiken durch Verwendung von Algorithmen, p. 108.

[563] For example, Michael *Grünberger*, Reformbedarf im AGG: Beweislastverteilung beim Einsatz von KI, ZRP 2021, 231 (234), who - rightly - considers this incompatible with Art. 21 Charter of Fundamental Rights and proposes the development of a "two-step model" of the burden of proof.

[564] Cf. only ErfKomm/Schlachter, 22nd ed. 2022, section 22 AGG marginal no. 8.

[565] Cf. only BAG, NZA 2011, 93; cf. also *Grünberger* ZRP 2021, 231 (233).

[566] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 361.

[567] See also *Sesing/Tschech*, AGG und KI-VO-Entwurf beim Einsatz von Künstlicher Intelligenz, MMR 2022, 24 (27) with further references.

[568] Cf. *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 360.

[569] Cf. only *Hinz*, in: Kaulartz/Braegelmann (eds.), Rechtshandbuch Artificial Intelligence and Machine Learning, 2020, p. 556: "In the use of unsupervised learning systems [...], in which the AI systems independently form different categories and contexts and adapt/change its learning objectives itself and the resulting clustering can no longer be traced by the employer or programmer in the sense of 'reverse engineering'. programmer can no longer be traced in the sense of "reverse engineering", it is hardly conceivable here that the employer, in the event of assertion of information requests, should be in a position to explain to a rejected applicant/employee for what motives he was rejected".

Against this background, it is understandable that there are many demands in the literature for a modification of Section 22 AGG.[570] It has been suggested, for example, that Section 22 AGG be amended such that so-called black box algorithms, which are uncoupled from knowledge of the internal functioning and implementation of the system, should in future be sufficient as an indication of discrimination; it would then be up to the party using the AI to refute this, for example by disclosing the technical and organisational measures implemented to avoid discrimination.[571]

3. Fundamental deficits of the current anti-discrimination law

After what has just been said, an adjustment of the concept of "indirect discrimination" seems urgently required. Likewise, the demands for facilitation of evidence for potentially aggrieved parties seem plausible.[572] While considering how far changes to the AGG should go and what they should look like in concrete terms, however, one must not neglect the fundamental question of whether the current anti-discrimination law is still structurally capable of guaranteeing sufficient protection against discrimination emanating from AI systems.

a) Identifiability of acts of discrimination

In particular, there are doubts as to whether individual legal protection directed at claims for damages and compensation is effective enough with regard to the use of AI. The concerns that exist in this context are due not only to the difficulties of proof outlined above, which a victim of discrimination will regularly face. The problems go deeper. For example, the literature has rightly pointed out that discrimination will often be difficult to detect under the conditions of AI use: “Humans discriminate due to negative attitudes (e.g. stereotypes, prejudice) and unintentional biases (e.g. organisational practices or internalised stereotypes) which can act as a signal to victims that discrimination has occurred. Equivalent

[570] Cf. also, for example, the Third Equality Report of the Federal Government, BT-Drs. 19/30750 of 10.06.2021, p. 138, according to which "platform operators [should] bear the burden of proof that they do not violate the provisions of the AGG on protection against discrimination when using algorithmic systems".

[571] Cf. *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 361 f.

[572] On this cf. also e.g. *De Stefano*, *Valerio*/*Wouters*, *Mathias*: AI and digital tools in workplace management and evaluation – An assessment of the EU’s legal framework, May 2022, p. 66 f. with further reform proposals.

mechanisms and agency do not exist in algorithmic systems. … Compared to traditional forms of discrimination, automated discrimination is more abstract and unintuitive, subtle, and intangible”.[573] Against this background, there is good reason to doubt whether the "traditional legal remedies and procedures for detection, investigation, prevention, and correction of discrimination which have predominantly relied upon intuition" are still fit for purpose.[574]

b) Collective legal protection

The AGG is to a large extent directed to individual legal protection, which provides victims of discrimination with rights, in particular to damages and compensation. In addition, Section 23 AGG opens up the possibility of support by anti-discrimination associations and creates an institution, the Federal Anti-Discrimination Agency, to support persons in asserting their rights. However, this does not change the fact that a person protected by anti-discrimination law generally has to face the user of the algorithm alone, clearly in an inferior position to both the user and the machine, and also under considerable pressure to assert their rights in a timely manner.[575] There are indeed ways and means to strengthen the position of the protected party if they try to enforce their rights by taking legal action. However, this alone does little to change the considerable knowledge asymmetry[576] and most importantly the fact that the potentially injured party is procedurally relegated to the role of the aggressor. It appears downright overwhelming for the discriminated party to take on an algorithm in a court battle.[577]

[573] Thus *Wachter/Mittelstadt/Russell*, Why Fairness cannot be automated Bridging the Gap between EU Non-Discrimination Law and AI, 1 (2), p. 10, p. 67.

[574] See *Wachter/Mittelstadt/Russell*, Why Fairness cannot be Automated Bridging the Gap between EU Non-Discrimination Law and AI, 1 (2).

[575] In this respect, cf. in particular also the two-month period of Sec. 21(5), first sentence AGG.

[576] From a US perspective, *Kim*, Data-Driven Discrimination at Work, William & Mary Law Review 2017, 857 (921): "The claimants would have to trace how the data miners collected the data, determine what populations were sampled, and audit the records for errors. Conducting these types of checks for a dataset created by aggregating multiple, unrelated data sources containing hundreds of thousands of bits of information would be a daunting task for even the best-resourced plaintiffs. In addition, the algorithm's creators are likely to claim that both the training data and the algorithm itself are proprietary information. Thus, if the law required complainants to prove the source of bias, they would face insurmountable obstacles".

[577] Cf. *Orwat*, Diskriminierungsrisiken durch Verwendung von Algorithmen, 2020, p. 137 f.: "Both the right to informational self-determination and anti-discrimination place burdens of responsibility on the individual concerned to identify and take action against unlawful data processing and unjustified unequal treatment. However, questions arise as to whether these basic legal concepts are still appropriate at all, given an increasing amount of data and algorithm-based as well as automated decision-making processes and their specific characteristics. This is because such burdens of responsibility require very high professional, cognitive and temporal prerequisites on the part of the individuals concerned in order to (a) perceive the many situations with data processing and differentiations at all, b) process the information resulting from the information duties under data protection law if necessary [...] as well as to enforce rights of access, correction or deletion and, above all, to (c) assess for themselves the individual consequences resulting from data processing and diverse (potential) differentiation decisions and to recognise the risk of possible discrimination for themselves".

In view of this situation, it is advisable to supplement individual legal protection with collective legal protection.[578] There are indeed two reasons for this – apart from the weakness of the former: On the one hand, this would seem to offset the knowledge asymmetry that characterises the problem as a whole at least to some extent, since the collective, at least potentially, "knows more" than the individual (and is also far more likely to be able to shoulder the costs of legal proceedings). On the other hand, and above all, discrimination when using algorithms is precisely not an individual "outlier", but is inherent in the underlying technology, which is why a "bundling" of interests seems the obvious choice from the outset. A right of action by associations could (at least partially) remedy this.[579]

c) The idea of prevention

Quite independently of this, however, the question arises whether the issue of "discriminatory algorithms" can be addressed with legal remedies that primarily aim to grant the affected party claims for damages and/or compensation. Incidentally, the "backward-looking, liability-focused model of legal regulation" is also subject to criticism in the USA (and in Europe, as well).[580] Instead, more efforts should be made to counter discrimination preventively.[581] Accordingly, many call for comprehensive operator obligations, compliance with which would have to be monitored, most likely by state authorities.[582] What is remarkable in all of this, however, is the scepticism that exists in many places towards an approach that relies solely on transparency and explainability of AI. Reliability, security and fairness of AI could, according to a widespread assessment, ultimately only be achieved through measures such as algorithm impact assessments, auditing and

[578] In this sense, e.g. also *Grünberger,* ZRP 2021, 231 (235) with further references: " Es ist daher dringend an der Zeit, über ein intelligentes Design kollektiver Rechtsschutzinstrumente nachzudenken und zu überlegen, wie man private und public enforcement auch im Nichtdiskriminierungsrecht sinnvoll kombiniert" ("It is therefore high time to think about an intelligent design of collective legal protection instruments and to consider how to combine private and public enforcement in a meaningful way in non-discrimination law as well".

[579] See also *Orwat*, Diskriminierungsrisiken durch Verwendung von Algorithmen, p. 135: " Der Ansatz des lediglich punktuellen Vorgehens im Einzelfall erscheint mit Blick auf die gegebenenfalls systematische Schlechterbehandlung von vielen Betroffenen durch algorithmenbasierte Differenzierungen nicht sachgerecht." ("The approach of merely proceeding selectively in individual cases does not appear appropriate in view of the possibly systematic worse treatment of many affected persons through algorithm-based differentiations").

[580] See *Kim*, Data-Driven Discrimination at Work, William & Mary Law Review 2017, 857 (867 f.).

[581] In favour of linking to Section 12 AGG, most recently *Sesing/Tschech*, AGG und KI-VO-Entwurf beim Einsatz von Künstlicher Intelligenz, MMR 2022, 24 (26); but cf. also, for example, the Third Equality Report of the Federal Government, BT-Drs. 19/30750 of 10.06.2021, p. 168 f. with the demand for the specification of preventive organisational duties.

[582] In this respect, the proposals range from the introduction of official controls of results to detect potential discrimination, if necessary using so-called control algorithms, to the establishment of official rights of information and inspection to control the processing mechanisms; cf. only *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 342 and 365 et seq.

certification.[583] At least it is encouraging that there are apparently increasing attempts to design algorithms in the sense of "built-in fairness".[584] In fact, in view of the above, it seems urgent, not to say inevitable, to take appropriate technical precautions against indirect discrimination.[585] From a German perspective, the so-called Hambach Declaration of the Data Protection Conference (DSK), the body of independent German data protection supervisory authorities of the Federation and the Länder, is also worth mentioning in this respect. This declaration contains the demand that "before AI systems are implemented the risks to the rights and freedoms of individuals shall be assessed with the aim inter alia of reliably excluding covert discrimination through countermeasures". Furthermore, "appropriate risk monitoring must also be carried out during the use of AI systems".[586]

[583] For example, *Castelluccia/Le Métayer*, Understanding algorithmic decision-making: Opportunities and challenges, 2019, p. 78; see *also Koene/Clifton/Webb/Patel/Machad/LaViolette/Richardson/Reisman*, A governance framework for algorithmic accountability and transparency, 2019.

[584] See only *Zehlike/Hacker/Wiedemann*, Matching code and law: achieving algorithmic fairness with optimal transport, in: Data Mining and Knowledge Discovery, 2020, p. 163; cf. on the whole also *Barocas/Hardt/Narayanan*, Fairness and Machine Learning Limitations and Opportunities, 2021: https://fairmlbook.org/pdf/fairmlbook.pdf.

[585] Cf. *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 353 ff. This in particular by making algorithm-based decision-making "blind" to specific factors susceptible to discrimination; ibid., p. 357.

[586] Declaration, p. 3 f. This was concretised in the position paper of the Conference of Independent Data Protection Authorities of the Federation and the Länder on recommended technical and organisational measures in the development and operation of AI systems of 6 November 2019; cf. on the whole also *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 360, who wants to impose an obligation on the operator to take technical precautions against indirect discrimination; similarly Wachter//Mittelstadt/Russell, Bias Preservation in Machine Learning. 360, who wants to impose an obligation on the operator to take technical precautions against indirect discrimination; similarly *Wachter//Mittelstadt/Russell,* Bias Preservation in Machine Learning: The Legality of Fairness Metrics Under EU Non-Discrimination Law, https://ssrn.com/abstract=3792772.

V. Data protection

It hardly needs explaining that the use of AI also has far-reaching implications for the area of data protection. This will be discussed in more detail below.

As far as employee data protection is concerned, the General Data Protection Regulation (GDPR) on the one hand and Section 26 of the Federal Data Protection Act (BDSG) on the other hand must be observed: The GDPR does not contain any specific provisions on employee data protection, but it does contain the essential data protection law evaluations for this area as well.[587] At the same time, Article 88 of the GDPR allows Member States to adopt "more specific rules to ensure the protection of the rights and freedoms in respect of the processing of employees' personal data in the employment context".[588] Article 88(2) of the GDPR specifies that national provisions also cover "monitoring systems at the work place", which should be understood as automated monitoring.[589] Article 88 GDPR takes into account the fact that special regulatory problems arise in the field of protection of employee data, on the one hand due to the structural inferiority of employees and on the other hand due to the special interest of employers in being able to monitor the performance of work.[590] As is well known, the German legislature made use of the option to create specific regulations with Section 26 of the BDSG. A specific employee data protection law does not exist.[591] It is contested whether Article 88 GDPR really only allows more specificity in the rules,[592] as the wording suggests, or also "real" deviations ("downwards" or "upwards").[593] In any case, however, the

[587] Cf. only *Seifert*, in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, 1st ed. 2019, Art. 88 marginal No. 1.

[588] See only *Gola*, in: Gola/Heckmann, Bundesdatenschutzgesetz, 13th ed. 2019, § 26 BDSG marginal no. 1 f.

[589] Thus *Seifert*, in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, 1st ed. 2019, Art. 88 marginal no. 43 with further references.

[590] For more details on the specific regulatory issues, see *Seifert*, in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 88 marginal no. 11 et seq.

[591] This could change however; cf. *Löber*, ZD-Aktuell 2022, 01120 with further references; most recently also *Franzen*, EuZA 2022, 261.

[592] See also Recital 155 in this respect: "Member State law or collective agreements, including 'works agreements', may provide for specific rules concerning the processing of personal data relating to employees in the employment context".

[593] Cf. *Hanloser*, in: Forgó/Helfrich/Schneider, Betrieblicher Datenschutz, 3rd ed. 2019, Chapter 1 marginal no. 16; cf. also *Wybitul,* NZA 2017, 413 (413), differentiating *Seifert*, in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 88 marginal no. 22 f., who believes that the purpose of Art. 88(1) DSGVO speaks in favour of the permissibility of national deviations "upwards"; cf. on the whole also most recently the order for reference of the VG Wiesbaden, ZD 2021, 393 (on the compatibility of Section 26(1)(1) BDSG corresponding to Section 23(1)(1) HDSiG with Art. 88 DSGVO). See also *Schild*, ZD-Aktuell 2022, 01178.

GDPR applies with regard to the processing of personal data unless a processing purpose has been conclusively regulated by Section 26 BDSG.[594]

The following look at the GDPR and BDSG will show that AI and, above all, Big Data pose considerable problems for the current law and, in part, also create considerable pressure for reform.

1. Basic terms

Article 1 of the GDPR specifies the subject matter and objectives of the Regulation. A distinction must be made between two equally important objectives: the protection of the fundamental rights of natural persons with regard to the processing of personal data (paragraphs 1 and 2) on the one hand, and, on the other, the free movement of data within the EU (paragraphs 1 and 3). [595]

Article 4 GDPR contains definitions of the main terms. These definitions alone are put to a serious test by AI and Big Data.

a) Personal data

The linchpin of the GDPR is the protection of "personal data". Accordingly, it is not surprising that the list of definitions in Article 4 GDPR starts with this term.

According to Article 4 No. 1 GDPR, personal data are "any information relating to an identified or identifiable natural person [...]; an identifiable natural person is one who can be identified, directly or indirectly". Overall, the definition of "personal data" was deliberately kept extraordinarily open and thus flexible.[596] Equally deliberately, the Union legislature accepted the resulting legal uncertainty.[597]

[594] Cf. only *Gräber/Nolden*, in: Paal/Pauly, DS-GVO BDSG. 3rd ed. 2021, § 26 BDSG, marginal no. 10; *Malorny*, RdA 2022, 170.

[595] See also, for example, *Spindler/Dalby,* in: Spindler/Schuster, Recht der elektronischen Medien, 4th ed. 2019, Art. 1 DSG-VO marginal no. 1.

[596] According to the case law of the ECJ, the term must also be interpreted broadly; cf. on this only *Karg,* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 1 marginal no. 3 with further references. (and footnote 10).

[597] See only *Tosoni/Bygrave*, in: *Kuner/Bygrave/Docksey/Drechsler*, The EU General Data Protection Regulation - A Commentary, 2020, Art. 6 note 7 with further references.

If information is not assigned to a person or cannot be assigned to a person, the the GDPR does not apply. Anonymous data are not protected by the GDPR.[598] However, this is where the problems begin in connection with AI, considered together with Big Data, the added value of which consists precisely in the fact that data can be statistically correlated with each other, where this was previously not possible or feasible for reasons of time or cost.[599] Indeed, Big Data analytics and AI regularly draw non-intuitive and unverifiable conclusions and make predictions, about, for example, people's behaviour or certain inclinations. The GDPR undoubtedly applies to Big Data analytics based exclusively on personal data. However, such analytics can also use exclusively non-personal data.[600] If, for example, analyses on the behaviour of certain groups are then applied to persons belonging to groups, the GDPR might not be taken into account, even though the risk to these persons is obvious. Against this background, it is understandable that some authors claim that in reality the data being processed here are also personal (derived) data – which, however, is “not at the beginning but at the end” of the data processing, as is the case with the classic personality profile.[601]

According to Article 4 No. 1, personal data means any information relating to an identified or identifiable natural person. A reference to a person therefore exists if a person is directly identified by the information.[602] However, it also exists if a person is identifiable through the addition of further information or intermediate steps. An "identifiable" person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person. To answer the question of what is meant by "identifiability", one must also refer to

[598] Cf. only *Karg*, in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 1 marginal no. 19.

[599] See only *Roßnagel/Geminn/Jandt/Richtert*, Datenschutzrecht 2016 "Smart" genug für die Zukunft? Ubiquitous Computing und Big Data als Herausforderungen des Datenschutzrechts, 2016, p. 21 f.

[600] For more details, see *Roßnagel/Geminn/Jandt/Richtert*, Datenschutzrecht 2016 "Smart" genug für die Zukunft? Ubiquitous Computing und Big Data als Herausforderungen des Datenschutzrechts, 2016, p. 29 ff. with examples.

[601] Thus *Roßnagel/Geminn/Jandt/Richtert*, Datenschutzrecht 2016 "Smart" genug für die Zukunft? Ubiquitous Computing und Big Data als Herausforderungen des Datenschutzrechts, 2016, p. 26: "„Es wird, um es bildlich auszudrücken, keine Akte über eine bestimmte Person geführt, sondern es gibt eine Vielzahl dynamischer anonymer Akten, die in einem Augenblick auf eine bestimmte Person konkretisiert werden können" ("To put it figuratively, there is no file kept on a specific person, but there are a multitude of dynamic anonymous files that can be concretised to a specific person in an instant"). In conclusion, likewise *Zuiderveen Borgesius*, Singling out people without knowing their names - Behavioural targeting, pseudonymous data, and the new Data Protection Regulation, Computer Law & Security Review 2016, 256; also *Wachter/Mittelstadt*, A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI (October 5, 2018). Columbia Business Law Review 2019: https://ssrn.com/abstract=3248829, who call for the establishment of a "right to reasonable *inferences*". (Problems of a different kind arise when, as is often the case with AI and Big Data*, there* are mixed data sets, i.e. those containing personal and non-personal data; see *Tosoni/Bygrave,* in: Kuner/Bygrave/Docksey, The EU General Data Protection Regulation - A Commentary, 2020, Art. 4 note 6.

[602] Cf. only *Karg*, in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 1 para. 46, 54 et seq.; on the requirements for sufficient "identification" ibid., para. 48 et seq.

Recital 26. This states: “To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments”. This is significant mainly because the legislature has thus decided in principle in favour of the so-called objective or absolute theory (and thus against the so-called subjective or relative theory). According to the former theory, identifiability is already given if either the responsible body or any third party is able to connect the information to a person. Under the subjective or relative theory, in contrast, only those means are to be taken into account that are actually available to the respective responsible body in the concrete individual case in order to establish the personal reference.[603]

It is clear from what has just been said that the question of identifiability is burdened with considerable uncertainties from the outset. Data processing in the context of AI does not make it any easier to answer this question. Illustrative of this is, for example, the necessity arising from Recital 26 to take into account the means which are “generally likely to be used to identify the natural person directly or indirectly". This is a dynamic test which, in addition to objective factors (such as time and costs), must also take into account the technology available at the time of the processing. In other words, whether identifiability is given depends largely on the state of the art at the time of the legal assessment of the facts. However, in view of the increasing capabilities of AI to assign information to individuals, this means that a processing of data that is still anonymous today may very well be a processing of personal data at a later point in time. Accordingly, data controllers are obliged to conduct a continuous review and risk analysis to ensure that originally anonymous data can continue to be considered as such.[604]

However, the fact that AI noticeably increases the possibility of linking (initially) anonymous data with concrete persons should be beyond question.[605] Here, the

[603] Cf. *Karg*, in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 1, para. 58 et seq., who believes that "case law and the GDPR" have "now arguably answered the question in favour of the relative theory, albeit with strong limitations and adoption of some elements of the absolute theory".
[604] Cf. *Karg*, in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 1, para. 63.
[605] Cf. also *Holthausen*, RdA 2021, 19 (25) with the conclusion that "anonymity-preserving data mining in the context of Big Data [...] thus (remains) a challenge for data protection as well as data security and a task for research".

possibility of identification is based on statistical correlations between unidentified data and personal data concerning the same person. To put it another way, a data element that is anonymous at first glance is placed in the context of further data through the application of AI, which then enables a personal attribution.[606] In the meantime, the constantly expanding linking possibilities have even led to calls for a restrictive interpretation of Article 4 No. 1 of the GDPR, with the argument that an excessive application of the GDPR must be counteracted. It is argued that the technical possibilities now allow the linking of almost any data with a person,[607] although one could add that the use of AI systems can not only noticeably reduce the "costs of identification", but also the "time required" for this. If one followed this, however, there would be a risk of curtailing the scope of protection of Article 8 of the Charter of Fundamental Rights.[608]

Conversely, however, it cannot be overlooked that the GDPR could degenerate into a "law of everything" in view of the extremely open definition of "personal" data.[609] In this context, some gloomy forecasts state that the GDPR’s "system of legal protection based on such an all-encompassing notion and high intensity of positive compliance obligations is not going to be sustainable in the long run".[610]

According to Recital 26, the Regulation does not apply to the processing of personal data "which have been rendered anonymous in such a way that the data subject is not or no longer identifiable". Anonymisation procedures (as well as pseudonymisation procedures) are among the methods that can contribute to the implementation of data protection requirements through technology design.[611] Effective anonymisation prevents "all parties from singling out an individual in a dataset, from linking two records within a dataset (or between two separate

[606] For more details, see *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 36 ff.

[607] Cf. in particular *Forgó/Krügel*, MMR 2010, 17 (using the example of geodata).

[608] Thus *Karg*, in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht. 1st ed. 2019, Art. 4 No. 1 DSGVO marginal no. 65, who additionally opines that "the expansion of the scope of application of the DSGVO [...] is not caused by an extensive interpretation of the concept of personal data, but by the constantly increasing analytical capabilities of information and communication technology and the associated gain in knowledge about the personality".

[609] See *Purtova*, The Law of Everything. Broad Concept of Personal Data and Future of EU Data Protection Law, Law, Innovation and Technology 2018, 40.

[610] Thus *Purtova*, The Law of Everything. Broad Concept of Personal Data and Future of EU Data Protection Law, Law, Innovation and Technology 2018, 33. The author calls for abandoning the concept of 'personal' data as the cornerstone of data protection altogether and instead providing remedies for 'information-related harm' in the broadest sense.

[611] Thus *Hansen* in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 5 marginal no. 50.

datasets) and from inferring any information in such dataset".[612] However, it is unclear how specific the reference to a natural person must be and to what extent a sufficient reference to a person is also given if general statistical statements allow certain conclusions to be drawn about the training data used.[613] Technically, it is probably true anyway that completely anonymous data do not exist.[614] In particular, the literature calls for the establishment of criteria that can be used to verify beyond doubt whether data are personal or anonymous. In the absence of such verifiability, there would be "no guarantee that a dataset anonymised according to the state of the art is actually anonymous".[615]

In any case, however, there are considerable anonymity risks, especially in machine learning.[616] For example, there are findings that certain ML techniques can unexpectedly clearly "remember" the data used to train the model and that this "memory" may be so strong that a faithful image of the training data can be reconstructed.[617] In the meantime, the European legislature has also explicitly recognised that in the future, the possibility of converting anonymised data into personal data must be increasingly expected.[618] There are proposals in the literature on how "de-anonymisation" by AI could be prevented or at least sanctioned more effectively.[619]

[612] Cf. Article 29 Working Party, WP 216, p. 9. The Working Party (Article 29 Data Protection Working Party) was an independent advisory body to the European Commission on data protection issues, established on the basis of Article 29 of Directive 95/46/EC (Data Protection Directive) of 24 October 1995. The statements of the group - now replaced by the European Data Protection Board (cf. Art. 68 GDPR) - still carry weight in the interpretation of the GDPR. General on anonymisation methods *Karg*, in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 5, para. 50 et seq.

[613] Cf. *Winter/Battis/Halvani*, ZD 2019, 489 (89 f.) with reference to Opinion 05/2014 of the Art. 29 Group, which was based on a very broad definition of inference; cf. also *Meents*, in: Kaulartz/Braegelmann (eds.), Artificial Intelligence and Machine Learning, 2020, p. 465 marginal no. 1; cf. also *Gierschmann*, ZD 2021, 482 (482): "Science for determining and assessing anonymity still under development". "Privacy-friendly" methods for training AI models are described by *Puschky*, ZD-Aktuell 2022, 00019.

[614] See *Kolain/Grafenauer/Ebers*, Anonymity Assessment - A Universal Tool for Measuring Anonymity of Data Sets Under the GDPR with a Special Focus on Smart Robotics , November 24, 2021, Rutgers University Computer & Technology Law Journal 2022: https://ssrn.com/abstract=3971139, 29.

[615] *Winter/Battis/Halvani*, ZD 2019, 489 (490).

[616] Cf. in this respect also *Thieltges,* ZfP 2020, 3 (13 ff.) with the additional reference to the fact that "especially in the mixing of private and professional contexts [keyword: "*bring you won device*"] the personal reference is immanent".

[617] The trained network reacted noticeably differently to information that had already been used for training than to previously unseen test data; for more details, see *Winter/Battis/Halvani*, ZD 2019, 489 (492).

[618] However, this is not in the GDPR, but in Regulation (EU) 2018/1807 of the European Parliament and of the Council of 4 November 2018 establishing a framework for the free movement of non-personal data in the European Union, OJ L 303/59, which not only explicitly recognises that "the growing Internet of Things, artificial intelligence and machine learning [...] are significant sources of non-personal data". It also states that when "technological developments make it possible to transform anonymised data back into personal data, these data must be treated as personal data".

[619] Cf. *Roßnagel/Geminn*, ZD 2021, 487 with a consideration of Japanese law.

b) Pseudonomysation

Article 4 No. 5 GDPR contains a definition of "pseudonymisation". This means “the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person”. In this respect, Recital 26 clarifies that "[p]ersonal data which have undergone pseudonymisation […] should be considered to be information on an identifiable natural person".[620] With regard to the question of identifiability, Recital 26 states that, as stated above, "all objective factors*,* such as the cost of identification and the time required for identification" should be taken into account, *"*taking into consideration the available technology at the time of the processing and technological developments".

Again, it should be noted that the use of AI significantly increases the chances of "re-identification" - and thus of "overcoming" pseudonymisation.[621] For example, it is relatively easy to create profiles or augment existing profiles by linking the pseudonymised data records with other (possibly also pseudonymised) data. Such an "overlapping"[622] of two data sets, which in themselves do not allow conclusions to be drawn about the persons concerned, can therefore quickly lead to a re-identification of these persons.[623]

[620] Critical of the existing regulation *Schleipfer*, ZD 2020, 284 (291), according to which the GDPR overestimates the potential of pseudonymisation and thereby overlooks even more effective possibilities, which is why it would be desirable if the topic of pseudonymity, including all differentiations, were "intensively discussed" in the context of the upcoming evaluation of the GDPR.

[621] See, for example, *Sweeney*, Simple Demographics Often Identify People Uniquely, Carnegie Mellon University, Data Privacy Working Paper 3, 2000: If data omits, for example, names, social security numbers and addresses, but includes date of birth, gender and postcode, then 87% of the US population can nevertheless be uniquely identified; cf. on the whole also *Russell, Stuart/Norvig, Peter*: Artificial Intelligence - A Modern Approach, 4th ed., 2022, p. 1166 with further references.

[622] Thus *Hansen* in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 5 marginal no. 48.

[623] Cf. *Hansen* in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 5 marginal no. 48.

c) Profiling

The GDPR addresses a series of specific regulations to what is called profiling to protect the data subjects. In particular, the aim is to ensure greater transparency in processing.[624]

Article 4 No. 4 GDPR defines profiling as “any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements”. Generally speaking, profiling means "gathering information about an individual (or group of individuals) and evaluating their characteristics or behaviour patterns in order to place them into a certain category or group, in particular to analyse and/or make predictions about, for example, their: ability to perform a task; interests; or likely behaviour".[625] Particular importance is attached to the characteristic of automated personality assessment alluded to in Article 4 No. 4 GDPR, which in the present context is based on correlations and probabilities without there having to be a causal relationship.[626] In this context, profiling is characterised by the fact that new information and further insights into the personality of the data subject are generated by collecting, linking and analysing individual characteristics.[627] The data basis for profiling can be, for example, communication and use habits (activity in social networks, websites visited, etc). The instruments of profiling include, in particular, tracking.[628]

It is obvious that AI systems enable profiling: the existence of these systems and the potential availability of Big Data have significantly increased the possibilities for profiling and also allow real-time analysis. For example, the literature points out that AI systems in the service of insurers are able to determine the likelihood of illness of applicants based on their health records, but also on their habits (e.g. of

[624] Cf. only *Scholz*, in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 4 DSGVO marginal no. 1.
[625] *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 39.
[626] *Scholz*, in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 4 DSGVO marginal no. 9 with further references.
[627] Cf. only *Scholz*, in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 4 DSGVO marginal no. 6.
[628] Cf. only *Scholz*, in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 4 No. 4 DSGVO marginal no. 7 f.

diet or exercise) or social conditions.[629] In the present context, it should be noted in particular that the use of people analytics qualifies as profiling if it does not operate at an aggregate level, but rather makes assessments or predictions regarding individual employees.[630] However, AI-based text or speech analysis is also profiling, which in this case is intended to reveal certain personal or character traits of applicants.[631]

2. Principles for the processing of personal data

Article 5 GDPR formulates - in implementation of Article 8(2) of the Charter of Fundamental Rights[632] - a number of principles for the processing of personal data.[633] With regard to these principles, too, numerous questions arise in connection with the use of AI, as will be shown below. Before looking at these in more detail, however, a brief word is in order about the meaning of the principles.

a) General meaning of the principles

The principles of Article 5 GDPR constitute a general objective order of data protection law. They constitute directly applicable law.[634] However, some of the principles mentioned in Article 5 GDPR require a high degree of concretisation. They also always presuppose a balancing of different interests.[635] The importance of the principles of Article 5 GDPR lies in the fact that they contain objectives for the design of data processing systems and the implementation of data processing

[629] *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 39.

[630] Cf. only *Holthausen,* RdA 2021, 19 (26) with further references. On the whole, cf. also *Blum*, People Analytics - Eine datenschutzrechtliche Betrachtung moderner Einsatzszenarien für automatisierte, datenbasierte Entscheidungen, 2021, p. 249 ff.

[631] Cf. *Joos*, NZA 2020, 1216 (1217).

[632] Art. 8(2), first sentence, Charter of Fundamental Rights reads: "Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law".

[633] Cf. also Resolution of the 97th Conference of the Independent Data Protection Authorities of the Federation and the Länder, Hambach Castle 3 April 2019, Hambach Declaration on Artificial Intelligence - Seven Data Protection Requirements. As the title suggests, this formulates data protection law requirements that are predominantly related to the principles mentioned in Art. 5 DSGVO: 1. AI must not turn humans into objects (Art. 22 GDPR); 2. AI must only be used for constitutionally legitimised purposes and must not override the purpose limitation requirement (Art. 5(1)(b) GDPR); 3. AI must be transparent, comprehensible and explainable (Art. 5(1)(a) GDPR); 4. AI must avoid discrimination; 5. The principle of data minimisation applies to AI (Art. 5(1)(c) GDPR); 6. AI needs accountability; 7. AI needs technical and organisational standards. On the basis of the Hambach Declaration, the Commission then drew up concrete requirements for AI systems in a position paper in November, the implementation of which it recommends for a data protection-compliant design of AI systems; cf. position paper of the Conference of the Independent Data Protection Authorities of the Federation and the Länder of 6 November 2019.

[634] Cf. *Roßnagel*, in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 23 f.

[635] Cf. *Roßnagel*, in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 5 marginal no. 22.

operations.[636] As far as the development and use of new technologies, such as AI systems, is concerned, they provide a framework for data protection requirements, which in turn can offer some guidance to data controllers. [637]

However, it has been criticised that the EU legislature has not further developed the principles and adapted them to the new technical challenges. It is argued that the provisions of the GDPR are essentially tailored to the relationship between the controller and the data subject, and yet the situation has become much more complex due to AI and Big Data. The critics point out, for example, that many people are involved in varying roles, that the data are distributed across many places and only used when necessary and usually unnoticed, that multiple purposes are pursued and that the processing is often organised by the systems themselves.[638] Some judgements are harsh: "Ignoring specific threats to the Principles from the challenges of modern and future forms of data processing means that the Principles are diametrically opposed to these developments and should actually prevent them. It is much more likely, however, that the normative force of the factual will lead to future developments undermining the principles and causing them to lose their function." [639]

In fact, all of the principles mentioned in Article 5 of the GDPR are exposed to massive challenges in view of the development of AI and Big Data. Accordingly, the "AI-proofness" of the GDPR is subject to considerable doubts, which is why some even recommend a comprehensive modification of the Regulation.[640] In its evaluation of the GDPR, the European Commission also recognised that Article 5 of the GDPR faces particular challenges.[641]

[636] Thus *Roßnagel*, in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 21.

[637] Cf. *Jaspers/Schwartmann/Hermann* in: Schwartmann/Jaspers/Thüsing/Kugelmann (eds.), DS-GVO/BDSG, 2nd ed., 2020, Art. 5 Principles for the Processing of Personal Data, para. 95.

[638] Cf. *Roßnagel/Geminn/Jandt/Richtert*, Datenschutzrecht 2016 "Smart" genug für die Zukunft? Ubiquitous Computing und Big Data als Herausforderungen des Datenschutzrechts, 2016, p. 99.

[639] Thus *Roßnagel*, in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 30: "Die Ignoranz gegenüber spezifischen Gefährdungen der Grundsätze durch die Herausforderungen moderner und zukünftiger Formen der Datenverarbeitung führt dazu, dass die Grundsätze diesen Entwicklungen diametral entgegenstehen und diese eigentlich verhindern müssten. Viel wahrscheinlicher ist jedoch, dass die normative Kraft des Faktischen dazu führt, dass die zukünftigen Entwicklungen die Grundsätze unterlaufen werden und diese ihre Funktion einbüßen".

[640] For example, *Brink/Groß,* RuP 2019, 105 with the demand that the "Hambach Declaration on Artificial Intelligence" of the Conference of Independent Data Supervisory Authorities be included in the GDPR.

[641] Communication from the Commission to the European Parliament and the Council - Data protection as a pillar of citizens' empowerment and the EU's approach to the digital transition - two years of application of the General Data Protection Regulation, COM(2020) 264 final, p. 13: "Future challenges lie ahead in clarifying how to apply the proven principles to specific technologies such as artificial intelligence, blockchain, Internet of Things or facial recognition which require a monitoring on a continuous basis"; cf. on this also *Roßnagel*, MMR 2020, 657 (659); *Heberlein*, ZD 2020, 487 (490); cf. on the whole also European Data Protection Board, Response to the MEP Sophie in't Veld's letter on unfair algorithms of 29 Jan 2020 (GDPR as a "solid legal framework").

b) The individual principles of Article 5 GDPR

aa) Lawfulness, fairness and transparency

According to the principle of legitimacy, the processing of data requires a legal basis. However, the principle not only concerns the "whether" of data processing, but also the "how", so that the manner of data processing must also be lawful.[642]

The threats to this principle posed by AI systems are primarily linked to the central element of consent. A position that insists on consent being given by the data subject hardly seems tenable given the ubiquity of computer-based systems and the ability to process information drawn from different sources in real time.[643] Nor is it possible to see how procedures familiar from other contexts (such as consent by ticking a box) could provide a remedy.[644] Finally, it seems almost impossible for practical reasons to ask the data subjects (who, moreover, are regularly numerous) for consent beforehand in the case of Big Data analytics, especially since the user often does not even know who the data subjects are.[645]

Article 5(1)(a) GDPR also stipulates that personal data must be processed "fairly".[646] Article 5(1)(a) of the GDPR refers to fairness in the sense that the data subject must be informed about the existence of the processing operation and its purposes. In doing so, the controller should provide the data subject with "any further information necessary to ensure fair and transparent processing taking into account the specific circumstances and context in which the personal data are processed". Furthermore, "the data subject should be informed of the existence of profiling and the consequences of such profiling".[647] Already this aspect of fairness

[642] Cf. *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 5 marginal no. 38.
[643] Cf. also *Schürmann*, ZD 2022, 316 (318): "fraglich, ob eine Einwilligung in die Verarbeitung personenbezogener Daten zur Anwendung von KI überhaupt möglich ist" ("questionable that consent to personal data processing for use by AI is even possible").
[644] Certain hopes rest on so-called *Personal Information Management Systems* (PIMS); cf. on this *Botta*, MMR 2021, 946; cf. also *Conrad*, InTer 2021, 147.
[645] Closer *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 40; cf. also *Roßnagel/Geminn/Jandt/Richtert, Datenschutzrecht* 2016 "Smart" genug für die Zukunft? Ubiquitous Computing und Big Data als Herausforderungen des Datenschutzrechts, 2016, p. 102 ff; cf. on the whole also *Kollmar/El-Auwad,* K & R 2021, 73.
[646] The German version refers to *Treu und Glauben* (literally "good faith". It would have been advisable to use the term *Fairness* in the German-language version as well. In more detail *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 para. 46 f.
[647] GDPR, Recital 60. Overall, this aspect of fairness mainly serves to "underpin existing transparency obligations"; *Geminn*, ZD-Aktuell 2021, 05557.

raises questions in the present context, given the complexity of processing in AI applications, the uncertainty of the outcome and the multiplicity of uses.[648]

However, Article 5(1)(a) of the GDPR also addresses "substantive fairness". In this regard, Recital 71 urges controllers to “use appropriate mathematical or statistical procedures for the profiling, implement technical and organisational measures appropriate to ensure, in particular, that factors which result in inaccuracies in personal data are corrected and the risk of errors is minimised, secure personal data in a way that takes into account the potential risks involved for the interests and rights of the data subject and that prevents, inter alia, discriminatory effects on natural persons on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union membership, genetic or health status or sexual orientation, or that result in measures having such an effect”. Under this aspect as well, Article 5(1)(a) GDPR is recognisably very important in the context of AI and Big Data. Overall, however, fairness is under less pressure than other elements of Article 5 GDPR, but this is due simply to the manageable scope of application, which only relates to the way in which the right is exercised in the relationship between the controller and the data subject.[649]

"Transparency" requires that "information addressed to the public or to the data subject be concise, easily accessible and easy to understand, and that clear and plain language and, additionally, where appropriate, visualisation be used". The information may also be "provided in electronic form, for example, when addressed to the public, through a website".[650] Significant challenges also arise with regard to this principle due to AI applications. In the case of data processing under the sign of ubiquitous computing and AI, the " invisibility" of the collection is a "design feature of the technology and as such not a remediable flaw".[651] Apart from this, it is true that deep learning modelling often takes place in a black box. This means that millions upon millions of data points are input into the algorithm, which then identifies correlations between certain data features to produce an output. However, the process inside the box is mostly self-directed and hardly comprehensible to data scientists, programmers or users in general. This is especially true since most AI applications are based on neural networks, which are difficult to decipher. True, there are efforts to find technical ways to open the black

[648] See *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 44.
[649] Cf. *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 48, who does not rule out an increase in importance if "consent and weighing of interests are increasingly becoming a farce".
[650] Recital 58.
[651] Cf. *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 61.

box.[652] But recently, there has been an increase in comments warning against "false hopes".[653]

Against this background, it can come as no surprise that ensuring transparency of AI systems is considered a central challenge. The Data Protection Conference (DSK) has formulated a series of transparency requirements in a position paper. For example, “the respective responsible party must specify during planning and at least document for operation on which theoretical basis and with which method raw data are normalised and standardised; synthetic data are generated; a dataset is completed or errors are corrected; personal data are pseudonymised and/or anonymised; the total amount of raw or training data for a basic training is not exceeded; training data for the basic training of an AI component [...] is regulated; up to which development stage the ultimately valid test data may be used before the transition from the test status to the productive status takes place; legally prohibited, negative discrimination is prevented [...]; the relevance or representativeness of the training data for the knowledge domain is determined”.[654]

With regard to AI systems, it is particularly important that transparency pursuant to Article 25(1) and (2) GDPR can also be ensured through data protection-compliant system design and through data protection-friendly default settings pursuant to Article 25(1) and (2).[655] Certification procedures pursuant to Article 42 of the GDPR may also be considered.[656] We will come back to both of these in a moment.[657]

[652] On this so-called explainable AI, see most recently *Hamon/Junklewitz/Sanchez/Malgieri/De Hert*, Bridging the Gap Between AI and Explainability in the GDPR: Towards Trustworthiness-by-Design in Automated Decision-Making, IEEE Computational Intelligence Magazine, Feb 2022, 72: doi: 10.1109/MCI.2021.3129960; *Hacker/Passoth*, Varieties of AI Explanations under the Law. From the GDPR to the AIA, and Beyond (August 25, 2021). in: Holzinger/Goebel/Fong/Moon/Müller/Samek (eds.), Lecture Notes on Artificial Intelligence 13200: AI - beyond explainable AI, Springer, 2022: https://ssrn.com/abstract=3911324 or http://dx.doi.org/10.2139/ssrn.3911324.

[653] Thus explicitly *Marzyeh Ghassemi, PhD/uke Oakden-Rayner/ Andrew L Beam*: The false hope of current approaches to explainable artificial intelligence in health care, Viewpoint November 01, 2021, e745: DOI:https://doi.org/10.1016/S2589-7500(21)00208-9; most recently also *Krishna/Han/Gu/Pombra/Jabbari/Wu/Lakkaraju*, The Disagreement Problem in Explainable Machine Learning: A Practitioner's Perspective: https://arxiv.org/abs/2202.01602.

[654] Position paper of the Conference of the Independent Data Protection Authorities of the Federation and the Länder of 6 Nov 2019.

[655] For more details, see *Klingbeil/Kohm,* MMR 2021, 3.

[656] Cf. also *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 54.

[657] Cf. sections G.V.7. and 8.

bb) Purpose limitation

Article 5(1)(b) of the GDPR lays down the principle of purpose limitation. According to this, personal data may only be collected "for specified, explicit and legitimate purposes and not be further processed in a manner that is incompatible with those purposes". The principle of purpose limitation is by far the most important principle of German and European data protection law and distinguishes it from data protection law in other legal systems.[658] Moreover, it is required by Article 8(2), first sentence, of the Charter of Fundamental Rights.[659]

Article 5(1)(b) GDPR requires the clear definition of a legitimate purpose and limits further processing to this purpose. However, it is likely to become increasingly difficult to determine and limit the purpose of individual data processing operations in advance when data from very different sources are increasingly being evaluated and merged.[660] Accordingly, there are doubts as to whether the purpose can continue to be the appropriate criterion for distinguishing permissible data processing from impermissible data processing in the future. [661]

In any case, however, it is often impossible to predict what the algorithm will learn. The purpose can change as the machine learns and "evolves".[662] AI systems allow re-purposing, the processing of personal data for entirely new purposes. To take an example outside of employment law, data may be collected for the purpose of contract management, only to be processed later for advertising purposes.[663] In many cases, AI systems are designed for re-purposing. This applies, for example, to so-called knowledge discovery in databases (KDD), where decision rules emerge in a dynamic and unpredictable manner in the course of automated processing, without these rules always being fully predictable even by the

[658] Cf. only *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 63.

[659] Cf. *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 64.

[660] *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 111.

[661] *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 112: "If "smart" information technologies are to expand the senses of the user, they cannot only collect data for a specific purpose. Like the user's senses, they must perceive the entire environment. Only when this data has been collected and stored can a purpose-oriented selection and evaluation gradually take place." *Holthausen*, RdA 2021, 19 (24), who states "a fundamental tension between Big Data on the one hand and data protection law purpose limitation on the other hand " is also sceptical with regard to Big Data and believes that Big Data "in principle and in its conception does not provide for any purpose limitation (so-called multidimensionality of data processing)".

[662] See also The Norwegian Data Protection Authority, Artificial Intelligence and Privacy, Report January 2018, p. 18.

[663] Ex. after *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 45.

developers.[664] In this respect, the crucial question is whether the new purpose is "compatible" with the purpose for which the data was originally collected or whether there is an impermissible change of purpose.

Whether there is compatibility in this sense is decided according to Article 6(4) on the basis of five criteria: (a) the "link between the purposes for which the personal data have been collected and the purposes of the intended further processing"; (b) the "context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller"; (c) the "nature of the personal data"; (d) the "possible consequences of the intended further processing for data subjects"; and (e) the "existence of appropriate safeguards, which may include encryption or pseudonymisation." All these criteria require a great deal of concretisation, so that determining the permissibility of a change of purpose is accompanied by considerable legal uncertainty.[665]

Whether re-purposing is permissible will have to be answered differently from case to case. The existing uncertainties are reflected in positions that sometimes have no common ground. For example, according to some views in the literature, there are no concerns about personal data being used in the training of AI systems, as long as there is protection against misuse and sufficient security measures are taken.[666] In contrast, the DSK is much stricter in this respect. According to it, the processing of personal data for the training of AI components constitutes a separate processing purpose, so that only such data can be used that serve the directly identified purpose. The data may only be used for another purpose if the conditions for a change of purpose are met or if there is an explicit legal basis (so-called non-linking). [667]

[664] Cf. *Tabarrini*: Understanding the Big Mind, EuCML 2020, 135 (141) with concerns also from the point of view of transparency (Art. 5(1)(a)) and from the point of view of Art. 5(2) GDPR.

[665] So also *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 6 para. 4 marginal no. 35.

[666] *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 46 f.

[667] Position paper of the Conference of the Independent Data Protection Supervisory Authorities of the Federation and the Länder of 6 November 2019, p. 9; also *Roßnagel/Geminn*, ZD 2021, 487 (487), according to which personal data can only be used in a legally secure manner for training, testing and evaluating AI systems after prior anonymisation. The position paper then goes on to say: "It is particularly problematic when legal prohibitions on discrimination do not permit the use of certain data and instead highly correlated substitute variables are used. Such discrimination is conceivable if, for example, characteristics such as first name, weight, products purchased, etc. are used instead of gender. In this respect, the DSK demands that AI systems and their individual components be checked for their discriminatory properties at an early stage and on a permanent basis in order to recognise and, if necessary, avoid discrimination potentials".

cc) Data minimisation

According to Aricle 5(1)(c) of the GDPR, the principle of data minimisation applies. According to this, the processing of personal data must be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed".[668] It is obvious that a considerable area of tension arises in the present context: On the one hand, there is the principle that as little data as possible should be processed. On the other hand, there are AI and Big Data, which are based on nothing less than the most comprehensive data analysis possible to uncover previously unseen correlations.[669] To put it casually, the more data can be used, the better. To resolve this tension, some commentators have proposed a proportionality test. The question is whether the inclusion of additional personal data in a processing operation creates a benefit that outweighs the risks for the data subjects.[670] According to another view, the principle of data minimisation should not only have an impact on the choice of means, but also on the choice of purposes. This would mean that the less intrusive approach should always be chosen if the purpose can also be achieved through this.[671]

The fact that AI poses considerable challenges for the applicable law is shown particularly strikingly by the example of applications that make it possible for objects networked in the Internet of Things to create a "memory" in order to reconstruct their "life traces". If the corresponding data are compared with each other, not only can the common context of different Things be determined, but also the social context of their respective owners. And yet, as regards the (original) function of "memory support", all collected data can be considered "appropriate, substantial and of the necessary degree".[672] However, the literature also points out that AI systems often create added value precisely by accessing data in the background that has already been generated by other applications and merging it with current data. However, such dynamic inclusion of data from a wide variety of sources makes it difficult to "enforce a limit on the data to be collected for each

[668] See also Recital 78 on the measures to be taken to minimise data.

[669] See also *Kugelmann*, DuD 2021, 503 (506): "A fundamental contradiction exists between the need for AI systems to use as much training data as possible and the principle of data minimisation under Art. 5 GDPR".

[670] Thus *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 47.

[671] Thus *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 124: "Grundsatz der Datensparsamkeit als Gestaltungsziel"; cf. on the connection between data minimisation and purpose limitation also *Holthausen*, RdA 2021, 19 (24).

[672] *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 134; *ibid.*, DuD 2016, 561.

individual application".[673] Finally, with self-learning machines, as already mentioned above, the purposes of data processing can change. This also results in specific challenges with regard to data minimisation, as it is no longer possible to predict how much data will be needed for learning.[674]

The Data Protection Conference also has the goal of data minimisation in mind in its position paper already mentioned above. Specifically, it states: “The amounts of data required for training with an acceptable error in order to achieve the designated target values of the system behaviour should be estimated on a theory-based basis when specifying an AI system. While it is possible to use "arbitrary data" in "arbitrarily large quantities" to try to identify the essential features of a poorly understood knowledge domain, this increases the risks to the rights and freedoms of data subjects: If AI components are trained with categories of data whose relevance to the knowledge domain is not clarified, risks may subsequently arise when they are used in AI systems. These risks may be, for example, that based on the categories of data, such as gender, the AI system provides discriminatory or erroneous results”.[675]

dd) Accuracy

The principle of "accuracy" laid down in Article 5(1)(d) of the GDPR states that personal data "shall be accurate and, where necessary, kept up to date". If personal data are inaccurate in relation to the purposes for which they are processed, "every reasonable step must be taken" to ensure that the data are "erased or rectified without delay". It is obvious that this principle of "data quality"[676] has considerable significance in the area of AI applications. This is all the more true when one considers that data are often used to make statements about the persons concerned or even to make decisions based on them, which can be of existential importance for these persons.[677]

[673] Thus *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal No. 134.

[674] See The Norwegian Data Protection Authority, Artificial Intelligence and Privacy, Report January 2018, p. 18.

[675] Position Paper of the Conference of the Independent Data Protection Supervisory Authorities of the Federation and the Länder v. 06.11.2019, p. 9.

[676] Thus *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 5 marginal no. 136.

[677] See *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 48.

The dangers are likely to increase even further in the course of the expected further technical development. This applies, for example, to ubiquitous computing, in which the use of computers is distributed across a multitude of digital end devices and systems that are simultaneously integrated into everyday activities. The users of these systems are normally not even aware of the enormous amount of data being collected and processed about them, which is why users cannot demand their accuracy.[678] Quite apart from this, however, AI systems as self-learning systems are designed precisely for processing systems to develop independently. Over time, however, such a system inevitably becomes "a black box, both for the person responsible and even more so for the data subject, whose results can be perceived, but whose structure, rules and data are not known and could at best be reconstructed laboriously in individual cases".[679] Under these circumstances, effective control seems almost impossible. After all, the application of AI typically involves the recognition of abstract patterns. Although these then form the basis for decisions that take effect vis-à-vis a concrete person, they are not themselves person-related. To put it another way: The problem here is not the accuracy of data, but the fact that a person is "treated according to the statistical average" by the AI system [680]

ee) Storage limitation

The principle of "storage limitation" in Art 5(1)(d) GDPR, which requires, among other things, that personal data be "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed", is also coming under increasing pressure from AI applications. Collecting as much data as possible in order to make them fruitful for all possible purposes is hardly compatible with a requirement of limited storage.[681] It should also be kept in mind that data are often processed for very different and, what is more, changing purposes.[682] Finally, the literature also points out that the elimination of personal references demanded by the principle of storage limitation is increasingly becoming an illusion due to technical developments. Indeed,

[678] Cf. *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 147.
[679] Thus *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 148.
[680] Thus *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 149; cf. also *ibid.*, DuD 2016, 561.
[681] See *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 49.
[682] Cf. *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 5 para. 164 with the example that the information that a passenger has boarded a train can be used for ticket control, reservation system, hospitality system, bonus system, telecommunication system, travel planning system and other systems.

sensors, for example, can perceive persons directly and also recognise them directly despite anonymisation or pseudonymisation. Big Data analytics also makes it relatively easy to de-anonymise anonymised or pseudonymised data.[683] The following applies: The more extensive and detailed the data available for an analysis, the greater the likelihood that it will be possible to re-identify the person by comparing these characteristic data.[684]

ff) Interim result

As an interim result of the considerations made so far, it can be stated that the relationship between the principles mentioned in Article 5 GDPR on the one hand and AI on the other is, to put it mildly, extremely tense.[685] Even more vividly, the literature states: "If one considers the fundamental principles of data protection law and additionally takes into account the further requirements of accuracy, storage limitation, integrity and confidentiality of personal data named in Article 5(1) of the GDPR, it becomes clear that Big Data, by its very conception, largely runs diametrically counter to these principles".[686]

3. Lawfulness of the processing of personal data

The core question of data protection law is, evidently, under which conditions data processing is lawful.

a) Lawfulness of the processing according to Article 6 GDPR

Article 6(1) of the GDPR links the lawfulness of the processing of personal data to one of two conditions: that consent is given (Article 6(1)(a)) or that the processing serves one of the purposes mentioned in the provision (Article 66(1)(b)-(e)). Finally, under Article 5(1)(f)) GDPR, processing is also permissible if it is necessary to protect the legitimate interests of the controller or of a third party. In this context,

[683] Cf. *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 166.
[684] Cf. *Weichert*, ZD 2013, 251 (258).
[685] Cf. also e.g. *Weichert*, in: Däubler/Wedde/Weichert/Sommer (eds.), EU-DSGVO und BDSG, Art. 22 DSGVO marginal no. 5, according to whom all data-protection principles are being put to the test ("sämtliche Datenschutzprinzipien werden auf den Prüfstand gestellt ").
[686] *Holthausen*, RdA 2021, 19 (25) with reference to *Roßnagel*, ZD 2013, 562 (564).

in the case of further processing, not only must there be a further processing purpose compatible with the original purpose, but any new processing that goes beyond mere collection must be fully covered by a legal basis in paragraph 1.[687] As far as consent is concerned, it should again be recalled that this requirement inevitably reaches its limits in the era of ubiquitous computing and Big Data.[688] In contrast, as far as the grounds in Article 6(1)(b)-(e) are concerned, these are often unlikely to be relevant in the present context.

Irrespective of this, processing of personal data is, as stated, also permissible if it is "necessary for the purposes of the legitimate interests pursued by the controller or by a third party". However, this is not the case where "such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data". In this respect, the literature argues that it is harmless in principle if personal data are further processed in order to develop an algorithmic model or to "train" the AI system. However, it raises even greater concerns if personal data are processed for the purpose of deriving concrete conclusions about the data subject. Accordingly, for example, the "assessment of the performance of employees on the basis of extensive monitoring" is to be inadmissible.[689] However, it is difficult to determine where the boundaries lie: Article 6(1)(f) GDPR is worded very openly, which on the one hand guarantees a certain flexibility, but on the other hand makes the results of the balancing exercise hard to predict.[690] There is also little that is tangible in the recitals. Efforts by the European Parliament to concretise the text were just as unsuccessful as the Commission's proposal to concretise the text by means of delegated acts.[691] There have also been repeated calls for special legal regulations, in the literature but also from business circles,[692] but these have not yet come about.[693] In addition to all this, data processing is based solely on the initiative of the data controllers, who

[687] Cf. only *Albrecht* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 13 with further references.

[688] *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 5 marginal no. 40.

[689] Thus *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 50.

[690] More closely *Schantz* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 6 para. 1 marginal no. 86.

[691] Cf. *Schantz* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 6 para. 1 marginal no. 103 with further references.

[692] Cf. also *Schantz* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 6 para. 1 marginal no. 86 (with further references), who explicitly mentions the case of "augmented reality such as [e.g.] Google Glass".

[693] Nevertheless, a certain concretisation in the form of approved codes of conduct of associations is conceivable; cf. Art. 40 para. 2(b) GDPR.

initially assess its permissibility independently, but often tend to assume that the processing is lawful.[694]

The re-purposing already mentioned above has been regulated separately. According to Article 6(4) of the GDPR, a number of criteria must be taken into account when answering the question of whether "processing for another purpose is compatible with the purpose for which the personal data are initially collected". According to the (non-exhaustive) catalogue of criteria, the following must be taken into account: any link between the purposes; the context in which the personal data were collected, in particular with regard to the relationship between the data subjects and the controller; the nature of the personal data; the possible consequences of the intended further processing for the data subjects[695] and the existence of appropriate safeguards, (e.g. encryption or pseudonymisation). There is no standard procedure approved by the supervisory authorities, especially with regard to the latter criterion. Accordingly, with regard to ascertaining when the conditions for processing for compatible purposes are fulfilled, there are complaints of widespread legal uncertainty.[696]

b) Prohibition of the processing of sensitive data

The GDPR classifies certain data as particularly sensitive: According to Article 9(1) GDPR, the processing of personal data "revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade-union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited".

In this respect, too, AI applications pose particular challenges. One challenge is that data that are not initially linked to a specific person can be linked to that person through further processing. The second challenge is that certain observable behaviour or known characteristics of individuals (e.g. their online activities) can

[694] Cf. on this *Schantz* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 6 para. 1 marginal no. 87 with further references.
[695] Cf. on this *Roßnagel* in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 6 para. 4 marginal no. 58 with the assessment that "under the conditions of modern data processing in the context of Big Data, artificial intelligence, self-learning systems, context detection, Internet of Things and other applications of ubiquitous computing [...] the consideration of possible consequences of changes of purpose can only have a restrictive effect"; likewise *Frenzel*, in: Paal/Pauly, DS-GVO BDSG, 3rd ed. 2021, Art. 6 marginal no. 49.
[696] Cf. *Dehmel*, ZD 2020, 62 (65).

allow conclusions to be drawn about attitudes, state of health or sexual orientation, that is, can lead to the generation of sensitive data.[697] However, the prevailing view is that Article 9 GDPR does not require that the data be sensitive in and of themselves. Rather, the provision should apply if their sensitivity only becomes apparent indirectly from the overall context of the processing.[698] However, there are also other voices according to which a "revealing" within the meaning of Article 9(1) GDPR is to be denied in the case of "automatically attached data". If one considers, for example, the analytical potential of new AI systems, which can, for example, analyse photographs and voice recordings to identify emotional states and health data, it becomes clear "how much territory the concept of 'revealing' can cover when simple processing is being performed if the specific purpose of the processing is not concretised".[699]

c) Specific regulations in Section 26 BDSG

Specific regulations on employee data protection are contained in Section 26 of the BDSG, the wording of which has been adapted to the GDPR.[700] According to Section 26(1), first sentence, BDSG: “Personal data of employees may be processed for employment-related purposes where necessary for hiring decisions or, after hiring, for carrying out or terminating the employment contract or to exercise or satisfy rights and obligations of employees' representation laid down by law or by collective agreements or other agreements between the employer and staff council”. The central criterion for the lawfulness of processing – and one that applies equally to all the grounds for authorisation in paragraph 1 – is therefore "necessity".[701] This concretises the general provision of Article 6(1)(b) of the GDPR, which in particular binds the processing of personal data for the

[697] Cf. *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 53; fundamental criticism in *Barczak*, DÖV 2020, 997 with the observation that "hard *data mining* (is) currently being carried out in this area as well".

[698] *Petri* in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 9 marginal no. 11 (so-called "context-related or indirectly sensitive data"), who at the same time objects to the fact that such data "are only cautiously assigned to the processing prohibition under para. 1"; ibid., marginal no. 12 with further references.

[699] Thus *Matejek/Mäusezahl*, ZD 2019, 551 (553): "The scope of application of Art. Art. 9 GDPR with regard to the first category [data that "reveal" certain sensitive characteristics] is only opened when the source data are used to obtain such sensitive data on the said characteristics and not already when it is established that the data are in principle suitable for disclosing such characteristics. Instead of the passive formulation "to emerge", the provision should therefore be understood in a more narrowed way in relation to the concrete intention to process (in the sense of an active "bringing forth") by way of a teleological reduction.

[700] Cf. also BT-Drucks. 18/11325, p. 97. General information on the relationship between Section 26 BDSG and the GDPR: *Gräber/Nolden*, in: Paal/Pauly, DS-GVO BDSG. 3rd ed. 2021, Section 26 BDSG, para. 8 ff.

[701] In more detail *Seifert*, in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 88 marginal no. 56 f.

performance of a contract to the principle of necessity.[702] Section 26(2), first sentence, BDSG provides for the possibility of consent, but at the same time stipulates that "in assessing whether such consent was freely given”, the “employee’s level of dependence in the employment relationship and the circumstances under which consent was given shall be taken into account". Section 26(3), first sentence, BDSG stipulates that – by derogation from Article 9(1) GDPR – the processing of special categories of personal data as referred to in Article 9(1) GDPR for purposes of the employment relationship is permissible primarily "if it is necessary to exercise rights or comply with legal obligations derived from labour law [...] and there is no reason to believe that the data subject has an overriding legitimate interest in not processing the data". According to Section 26(4), first sentence, BDSG, the processing of personal data is also permitted on the basis of collective agreements. Thus, works agreements and collective agreements can also permit the processing of employee data.[703]

The core prerequisite for the lawfulness of processing is, as said, its "necessity". In the examination of necessity that arises from this, "the conflicting fundamental rights positions must be weighed up in order to establish practical concordance". Specifically, "the employer's interests in data processing and the employee's right to privacy must be balanced in a way that takes both interests into account as far as possible".[704] The discussion about the permissibility of predictive policing of employees, for example, shows that the legality of the processing will often be a matter of dispute.[705] According to some views, this is inadmissible if it involves the creation of a personality profile.[706] However, the limits are sometimes defined differently. For example, it is sometimes argued that predictive policing is only inadmissible if it involves a "screening of the personality", whereby the borderline to such a screening is only crossed "if the collected data are close to the core of the right of personality and these are processed into a prognosis that is also close to the core of the right of personality".[707] The different views illustrate the imponderables of a proportionality test, in which not only data minimisation under

[702] *Seifert*, in Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 88 marginal no. 21.

[703] See only *Wybitul*, NZA 2017, 413 (417).

[704] BT-Drs. 18/11325 of 24 Feb 2017, p. 97.

[705] Already today, so-called fraud monitoring is sometimes carried out, in which the probability of committing crimes against the employer's assets is determined for certain activities. Predictive policing, on the other hand, is about using AI to forecast which employees are particularly susceptible to committing compliance violations; see *Rudkowski*, NZA 2020, 72 (73); *Haußmann/Thieme, NZA* 2019, 1612 (1615).

[706] *Rudkowski*, NZA 2020, 72 (74).

[707] Thus *Dzida*, NZA 2017, 541 (545), who therefore considers a Big Data analysis to be "legally problematic" if it is intended to determine whether employees are "generally inclined to commit criminal offences".

Article 5(c) GDPR but also the employer's entrepreneurial freedom, within which economic aspects may also come into play, must be taken into account.[708]

4. Information obligations and rights of the data subjects

a) Information obligations

Articles 13 and 14 of the GDPR contain information requirements for the collection of personal data. The former provision stipulates an information obligation when personal data is collected from the data subject, and the latter provision stipulates an information obligation when the personal data was not collected from the data subject.

In the present context, the provision in Articles 13(3) and 14(4) GDPR is of particular importance, as it is aimed at cases in which "the controller intends to further process the personal data for a purpose other than that for which the personal data were collected". The provision obviously serves to safeguard the principle of purpose limitation (Art. 5(1)(b) and Art. 6(3)(3), (4)).[709] However, the provisions in Articles 13(2)(f) and 14(2)(g) GDPR are also of particular importance. According to these provisions, the controller must inform the data subject about "the existence of automated decision-making, including profiling" and, "at least in those cases, provide meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject".[710] This corresponds to the identical provision in Article 15(1)(h) of the GDPR, which will be explained below; the obligation to provide information here differs from the disclosure duty provided for there only in that it is directed towards the future, whereas Article 15(1)(h) refers to automated decisions that have already taken place. [711]

[708] Cf. only *Gola*, in: Gola/Heckmann, Bundesdatenschutzgesetz, 13th ed. 2019, § 26 BDSG marginal no. 16 with further references.

[709] Cf. *Dix* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 13 marginal no. 20.

[710] *De lege ferenda* for an extension of these duties to inform to fundamental rights-sensitive algorithm-based procedures *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, pp. 285 et seq. and 293 et seq., respectively.

[711] Cf. *Dix* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 13 marginal no. 16.

b) Right to information

According to Article 15(1) of the GDPR, there is a right to information as to whether personal data have been processed; if this is the case, the data subject may request a number of specific details.[712] With regard to AI systems, Article 15(1)(h) of the GDPR is particularly relevant, which, as already mentioned, corresponds to the provision in Articles 13(2)(f) and 14(2)(g). According to this, the right of access also refers to "the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject".

The regulation raises questions. For example, it is debated whether there is a right to information with regard to profiling even if the requirements of Article 22(1) GDPR are not met.[713] However, the scope of application of the regulation is also unclear because it is not clear what the phrase "at least in those cases" refers to.[714] Accordingly, it is uncertain when a claim for "meaningful information about the logic involved" exists. In this respect, some in the literature call for a risk assessment: "The more serious the effects of the supported decisions can be for data subjects, the greater the risk of unseen adoption by the decision-maker and the more the results of the automated data processing are incorporated into the decisions to be made, the more interested data subjects are in the logic followed by the automated data processing and the scope this type of data processing can have”.[715]

Irrespective of this, however, it remains difficult to determine what exactly the responsible party owes and whether there is also an obligation to disclose calculation formulas.[716] It is obvious that particular difficulties arise with machine learning in view of the complexity of the procedures and the limited traceability of the results. In this respect, it is sometimes assumed that under current law, every person concerned has a right to access the internal documentation of the creation of a machine learning model. However, it must be ensured that the information

[712] Cf. on Art. 15 GDPR most recently *Leibold*, ZD-Aktuell 2021, 05313.
[713] Cf. *Sesing*, MMR 2021, 288 (289).
[714] See also *Sesing*, MMR 2021, 288 (290), who sees the existence of automated decision-making as such as the "only meaningful point of reference"; on the background to the regulation *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2020, p. 290.
[715] Thus *Sesing*, MMR 2021, 288 (290).
[716] See also *Sesing*, MMR 2021, 288 (291) with an overview of the state of the dispute.

interest of the person concerned is weighed against the interest in maintaining business secrecy and the operationality of the model.[717] *De lege ferenda,* "in view of the manifest risk of discrimination in the use of ML, a (ML-specific) right of the persons concerned to obtain information on the statistical distribution of the output of the ML model between the different protected groups (e.g. score distribution men vs. women vs. diverse persons)". Only in this way could the persons concerned ultimately examine whether there was any statistical unequal treatment at all and thus possibly (indirect) discrimination.[718]

Whether Article 15(1)(h) of the GDPR also gives rise to a right to a detailed explanation of automated individual decisions*,* as is found in German law, for instance in connection with the justification of administrative acts (Section 39(1) VwVfG), is a subject of debate in the literature, not least given the limited explicability of the results found, for example, through the use of algorithms or neural networks.[719] Proponents of such a claim point out that this is the only way to "counteract the informational asymmetry exacerbated by new technologies such as artificial intelligence".[720]

The relationship to the protection of trade and business secrets, which the obligated party may invoke, raises problems in the context of Article 15(1)(h) of the GDPR (as well as the claims under Art. 13(2)(f) and Art. 14(2)(g)).[721] This protection should be taken seriously, as "a blanket obligation to disclose the source code for software applications could pave the way for the exploitation of another's intellectual performance".[722] However, this is unlikely to be of much practical help anyway.[723] The disclosure obligation (and the obligations to provide information) also focuses on information about the logic involved as well as the scope and effects of such processing for the data subject, whereby "logic involved" means the structure and process of the data processing.[724] In this respect, however, one will have to expect the controller – if only due to the asymmetry of information that has

[717] Cf. *Hacker*, NJW 2020, 2142 (2144) with further references.
[718] Cf. *Hacker*, NJW 2020, 2142 (2144).
[719] See also *Sesing*, MMR 2021, 288 (292); cf. also *Cabral*, in: Hallinan/Leenes/De Hert (eds.), Data Protection and Privacy Data Protection and Artificial Intelligence, 2021, p. 29 ff.
[720] Thus *Dix* in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 15 marginal no. 25; largely the same applies to information obligations; cf. also ibid., Art. 13 marginal no. 16.
[721] General on the legal protection of algorithms *Söbbing*, ITRB 2019, 192.
[722] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2020, p. 289; similarly *Joos*, NZA 2020, 1216 (1218) and reference to Recital 63, according to which the right to information "should not adversely affect the rights or freedoms of others, including trade secrets or intellectual property and in particular the copyright protecting the software", but "the result of those considerations should not be a refusal to provide all information to the data subject".
[723] Likewise *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2020, p. 289.
[724] Cf. *Dix* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 13 marginal no. 16.

been exacerbated by the use of AI anyway – to explain the involved logic of a technique of automated decision-making or profiling to the data subject in such a way that the latter can exercise his or her rights under Article 22 GDPR.[725] There is much to be said for matching the extent of information owed to the context and the risk levels of the system.[726]

c) Further rights of data subjects

Article 17(1) of the GDPR is the basis for a right to deletion ("right to be forgotten").[727] Article 20 of the GDPR contains a right to data portability, although the scope of application of this provision in relation to AI systems is not settled. For example, there is no consensus on the question of whether there is also a right to data transfer if the controller of a Big Data system can establish a link to the data subject through a data link.[728] However, it is also doubtful whether the right also extends to data derived from the collected data about the data subject.[729] Article 21(1) of the GDPR contains a right of the data subject to object to the processing of personal data relating to him or her, whereby – particularly interesting in the present context – profiling is explicitly mentioned (Article 21(1), first sentence, GDPR).

5. The prohibition of automated decisions

According to Article 22(1) GDPR, "the data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her".[730] The provision should be read together with Recital 71. There, not only are "e-recruitment practices without any human intervention" mentioned as a

[725] Accurately *Dix* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 13 para. 16: "A controller who is not able to explain in a meaningful way to the data subject the logic involved in an automated decision-making or profiling technique cannot use it in a legally compliant way".

[726] Cf. in this respect also *Krafft/Zweig*, Transparenz und Nachvollziehbarkeit algorithmenbasierter Entscheidungsprozesse - Ein Regulierungsvorschlag aus sozioinformatischer Perspektive, 2019.

[727] On the - quite controversial - concept of "deletion" *Dix* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 17 marginal no. 5.

[728] Thus *Dix* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 20 marginal no. 6, who believes that this could also concern the use of fitness trackers, smart watches or smartphone apps; *Werkmeister/Brandt*, CR 2016, 233 (237), by contrast, is narrower.

[729] See *Sartor*, The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 57.

[730] Art. 22(4) GDPR additionally stipulates that decisions under paragraph 2 may not, in principle, be based on special categories of personal data under Art. 9(1).

possible case of application of the provision, but also, under "profiling", the automated processing of personal data "to analyse or predict aspects concerning the data subject's performance at work, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements" is explicitly mentioned.[731] Contrary to what the wording of Article 22(1) GDPR and the legal system suggest, this is not solely a right of the data subject. Rather, according to the prevailing opinion, Article 21(1) of the GDPR is at the same time a prohibition, or more precisely a prohibition with a reservation of consent.[732]

a) Purpose of the provision

The purpose of Article 22(1) GDPR is largely obscure.[733] However, the considerations made in connection with the predecessor provision, Article 15 of the Data Protection Directive,[734] are likely to be largely transferable to Article 22(1) GDPR.[735] The Commission stated in its original proposal that the provision was about protecting the data subject from being "made the subject of decisions by public- and private-sector institutions involving the assessment human conduct on the sole basis of an automatic processing of personal data forming a data or personality profile of the data subject". This provision is intended to protect "the interest of the data subject in participating in the making of decisions which are of importance to him." And further: "The use of extensive data profiles of individuals

[731] According to *Weichert*, in: Däubler/Wedde/Weichert/Sommer (eds.), EU-DSGVO und BDSG, Art. 22 DSGVO marginal no. 42a, the provision will be applicable even if, with its nudge, an employer pursues an intention or achieves a result that is directly controlling.

[732] Cf. only *Atzert* in: Schwartmann/Jaspers/Thüsing/Kugelmann, DS-GVO/BDSG, 2nd ed., 2020, Art. 22 marginal no. 1 f.; more recently also *Djeffal,* ZaöRV 2020, 847 (861), who interprets the provision itself as a "law-by-design-obligation", which requires developers to include the rights, freedoms and legitimate interests of data subjects in the design process. In this context, it is not sufficient to apply a given set of rules. Rather, a socio-technical impact assessment must be carried out and all aspects must be weighed up; ibid., 868; cf. on the whole also *Bygrave,* in: Kuner/Bygrave/Docksey/Drechsler (eds.), The EU General Data Protection Regulation - A Commentary, 2020, Art. 22 note 2 (with further references).

[733] In its preliminary ruling of 1 Oct 2021, VuR 2022, 70, the Wiesbaden Administrative Court commented on the purpose of Article 22(1) of the GDPR as follows: "The legislature's concern is to prevent decision-making from taking place without an individual assessment and evaluation by a human being. The data subject should not be at the mercy of an exclusively technical and opaque process without being able to comprehend the underlying assumptions and assessment standards and, if necessary, to intervene by exercising his or her rights. Thus, in addition to protection against discriminatory decisions based on supposedly objective data processing programs, the aim of the regulation is also to create transparency and fairness in decision-making. Decisions on the exercise of individual freedoms should not be left unchecked to the logic of algorithms. This is because algorithms work with correlations and probabilities that do not necessarily follow a causality and also do not necessarily lead to results that are "correct" according to human insight. Rather, erroneous, unfair or discriminatory conclusions can be drawn from the systematisation of accurate individual data, which - if they become the basis for decision-making - considerably affect the freedom rights of the person concerned and degrade him or her from the subject to the object of a depersonalised decision. This is particularly true if the data subject is not aware of the use of algorithms or - if he or she is - cannot overlook which data are included in the decision, with what weight and through which methods of analysis".

[734] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ 23.11.1995, p. 31.

[735] See *Bygrave*, in: Kuner/Bygrave/Docksey/Drechsler (eds.), The EU General Data Protection Regulation - A Commentary, 2020, Art. 22, A. Rationale and Policy Underpinnings.

by powerful public and private institutions deprives the individual of the capacity to influence decision-making processes within those institutions, should decisions be taken on the sole basis of his 'data shadow'."[736] The amended proposal then adds another idea to this: "The danger of the misuse of data processing in decision-making may become a major problem in the future: the result produced by the machine, using more and more sophisticated software, and even expert systems, has an apparently objective and incontrovertible character to which a human decision-maker may attach too much weight, thus abdicating his own responsibilities".[737]

At the same time, there is widespread agreement that Article 22(1) of the GDPR is closely related to the protection of human dignity.[738] The purpose of Article 22(1) of the GDPR is to ensure that "human beings do not degenerate into mere objects of fully automated data processing procedures".[739] This formulation can be found in the literature and largely corresponds to the so-called "object formula", which the BVerfG applies in connection with Article 1(1) of the Basic Law.[740] It is also occasionally stated that decision-makers should not be allowed to base "decisions that are legally relevant or even harmful for those affected 'exclusively' on rational considerations or a schematisation of facts of life".[741] The former view is only partially convincing, as one would then have to see in Article 22(1) GDPR a right of the data subject to (potentially) "irrational" decisions. In this respect, it cannot be denied that "purely machine" decision-making does not necessarily only offer disadvantages compared to human decision-making.[742] Just as human beings

[736] Commission of the European Communities, Proposal for a Council Directive on the Protection of Individuals with regard to the Processing of Personal Data, 13 Sept 1990, COM(90) 314 final, p. 29.

[737] Commission of the European Communities, Amended Commission Proposal for a Council Directive on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 15 Oct 1992, COM(1992) 422 final, p. 26.

[738] See only *Bygrave*, in: Kuner/Bygrave/Docksey/Drechsler (eds.), The EU General Data Protection Regulation - A Commentary, 2020, A. Rationale and Policy Underpinnings (with further references).

[739] Thus *Atzert* in: Schwartmann/Jaspers/Thüsing/Kugelmann, DS-GVO/BDSG, 2nd edition, 2020, Art. 22 marginal no. 2 with further references; cf. also *Geminn*, DÖV 2020, 172 (176) with the question of whether a technical system is even capable of appreciating the subject quality of human beings and treating them as anything other than an object. On the special position of the provision, which does not regulate the processing of personal data in the narrower sense, but the decision-making based on it and the application of a certain processing result *Scholz* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, 1st ed. 2019, Art. 22 marginal no. 4.

[740] Cf. only BVerfGE 9, 89 and C. I. 1.); also *Kunig/Kotzur*, in: von Münch/Kunig, Grundgesetz-Kommentar, 7th ed. 2021, Art. 1 marginal no. 33 with further references.

[741] *Atzert* in: Schwartmann/Jaspers/Thüsing/Kugelmann, DS-GVO/BDSG, 2nd edition, 2020, Art 22 marginal no. 2.

[742] Cf. in this respect, for example, *Schulz,* in: Gola/Heckmann, Bundesdatenschutzgesetz, 13th ed. 2019 Art. 22 marginal no. 2. *Atzert* in: Schwartmann/Jaspers/Thüsing/Kugelmann, DS-GVO/BDSG, 2020, 2nd ed., Art. 22 marginal no. 6, however, points out that the "decisions of AI are also based on algorithms that have to be programmed and implemented by natural persons who do not work without error", which is why the "human factor is also inherent in exclusively automated decision-making per se".

occasionally decide out of "compassion",[743] they also often decide out of antipathy.[744] In contrast, the latter view seems plausible, since a "schematisation of life situations", as is inherent in decision-making by AI systems, does not represent a "genuine" individual case decision.[745] Though one occasionally reads on the purpose of Article 22(1) of the GDPR that "the accountability for an onerous value judgment should always in substance lie with a natural person",[746] this is not true. Rather, it is correct that AI systems, at least in their current state, cannot make any "value decisions" at all.[747]

It is obvious that Article 22(1) of the GDPR is of central importance in the present context: are there previously unheard of possibilities not only for collecting data, but also for linking and analysing it. In the Internet of Things, data on all areas of life and from the most diverse sources are available and are being exploited. Nor are they restricted to the "classic" internet any more: one has only to consider so-called wearables,[748] or sensor data in the context of the Internet of Things.[749] As far as the aspect of linking data is concerned, we need only refer to the possibility of so-called scoring, in which future behaviour is predicted with a probability value. This is spreading into more and more areas of life including the assessment of job applicants and employees.[750]

b) Uncertain issues

As clear as the legal dignity and the practical significance of Article 22(1) of the GDPR are, it is also clear that the application of the provision gives rise to a plethora of doubts. The first question concerns the scope of application of the

[743] Cf. *Atzert* in: Schwartmann/Jaspers/Thüsing/Kugelmann, DS-GVO/BDSG, 2nd edition, 2020, Art. 22 marginal no. 6: "With the help of Art. 22, the legislator ensures that even in the age of ubiquitous digitalisation, there remains room for human discretion and compassion and thus the consideration of atypical or hardship cases continues to be guaranteed".

[744] Remarkably, the Art 29 Group has even acknowledged that an "automated decision-making process [can] potentially enable greater consistency or fairness in the decision-making process (e.g. by reducing the potential for human error, discrimination and abuse of power)"; cf. WP 29 WP 251 Rev. 01 v. 3.10.2017, p. 13.

[745] Cf. also *Scholz* in Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 22 marginal no. 10, who additionally notes that in Big Data analytics "correlations and probabilities (are) calculated, behind which there need not be causalities", and "forecasts based solely on statistical probabilities [...] (can) turn out to be individually wrong".

[746] Thus *Scholz* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 22 marginal no. 3.

[747] Cf. also *von Graevenitz*, ZRP 2018, 238 (240): "It also remains questionable whether a machine will ever be able to incorporate evaluative points of view in a way that is possible for human decision-makers and as is generally expected of them"; cf. also Enders, JA 2018, 238 (240). also *Enders*, JA 2018, 721 (725) with the observation that "even with a self-learning effect of the AI [...] the question (remains) who will provide the first, or more accurately one would have to say the first, initiating value decisions and set the standards for the subsequent "self-learning" of the AI.*"*.

[748] On this from a US perspective *Ajunwa*, Algorithms at Work: Productivity Monitoring Applications and Wearable Technology as the New Data-Centric Research Agenda for Employment and Labor Law, September 10, 2018, St. Louis U. L.J. 2019, 21: https://ssrn.com/abstract=3247286.

[749] Cf. *Scholz* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 21 marginal no. 8.

[750] *Scholz* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 21 marginal no. 9.

provision: While in the prevailing opinion Article 22(1) of the GDPR covers all automated data processing procedures that lead to a final decision, according to other opinions, the provision should only be applicable if the decision is aimed at evaluating individual personal characteristics[751] and, in any case, the merely selective inclusion of personal data in automated decision-making should not be covered by the prohibition.[752]

At first glance, defining what is meant by an automated decision poses considerably fewer problems. It should be relatively clear that one can only speak of a "decision" if there is a formative act with final effect.[753] It will also hardly be possible to speak of an "automated decision" if it is a simple if-then decision.[754] Beyond that, however, the picture quickly becomes unclear. There is no need to justify the fact that Article 22(1) of the GDPR is only aimed at automated decisions and not also at automatically generated proposals for a decision. Accordingly, it is obvious that the provision does not apply if the machine's output is "checked by a human being – on the basis of further criteria – in the sense of a final assessment and thus translated into a decision of its own" and that this must also apply if the suggestion is ultimately followed.[755] However, it remains unclear what exactly the "translation into a decision of one's own" requires.

A similarly difficult question is when a decision is based *solely* on automated processing.[756] This is to be assumed not only if no review by a human being is intended from the outset and thus does not take place, but also if the human being merely confirms or adopts the automated specification.[757] Instead, according to the prevalent opinion, it should be required that the human being has the necessary data basis for such a review and possesses the necessary qualification. In addition, it is said that the human being should have a margin of discretion to be able to deviate from the automated decision if necessary.[758] All of this assumes that the human being has the actual possibility of reviewing the situation, which is likely to

[751] Cf. only *Scholz* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 22 marginal no. 19 with further references on the state of the dispute (and fn. 50).
[752] Thus *Höpfner/Daum*, ZfA 2021, 476 (485).
[753] Thus *Scholz* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 22 marginal no. 17.
[754] Thus again *Scholz* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 22 marginal no. 18 with reference to the protective purpose of the statute.
[755] Thus *Scholz* in Simitis/Hornung/Spiecker gen. Döhmann, (eds.) Datenschutzrecht, 1st ed. 2019, Art. 22 marginal no. 28.
[756] Recital 71 cites an "online recruitment procedure without any human intervention" as an example of a decision based solely on automated processing. Further details on this *Binns*/*Veale*, International Data Privacy Law 2021, 319.
[757] Cf. *Scholz* in: Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 21 marginal no. 26.
[758] Cf. *Scholz* in: Simitis/Hornung/Spiecker gen. Döhmann, (eds.) Datenschutzrecht, 1st ed. 2019, Art. 21 para. 27 (citing Art. 29 Group, 17/EN WP251rev. 01, p. 10).

cause mass automated decisions to fail due to Article 22(1) of the GDPR, because they make human monitoring impossible. Even if one follows this, however, there still remains a considerable grey area. This is shown in particular by the fact that there is disagreement on the question of whether it satisfies the requirements of Article 22(1) GDPR "if the review is limited to filtering out implausible decisions".[759] It is true that even program decisions that are likely to be correct may turn out to be individually wrong.[760] However, it is precisely the question of whether Article 22(1) of the GDPR requires that more than just an "evidence check" be carried out in relation to automated decisions.[761]

Finally, the requirement that the automated decision "has legal effect on the data subject or similarly significantly affects him or her" is not at all unproblematic. This also applies in the present context. However, it is obvious that a "legal effect" is to be affirmed in private law, such when a contract is terminated.[762] However, as regards instructions under labour law, it must be noted that these are only aimed at concretising the duties arising from the employment contract to begin with; the employer can only "specify in more detail" the content, place and time of the work according to Section 106, first sentence GewO. The fact that in individual cases it may be unclear when an actual impairment is "considerable" (and not merely a "nuisance", for example),[763] hardly requires justification.

6. Data protection through technology design and data protection-friendly default settings

Article 25 of the GDPR describes a new instrument, namely data protection by design and data protection by default.[764] According to this, the controller must take the appropriate technical and organisational measures,[765] in particular to effectively

[759] In the negative *Scholz* in: Simitis/Hornung/Spiecker gen. Döhmann, (eds.) Datenschutzrecht, 1st ed. 2019, Art. 21 marginal no. 27; affirmaing BeckOK DatenschutzR/v. *Lewinski*, DSGVO Art. 22 marginal no. 25.1.

[760] Thus *Scholz* in: Simitis/Hornung/Spiecker gen. Döhmann, (eds.) Datenschutzrecht, 1st ed. 2019, Art. 21 marginal no. 27.

[761] In this, that the human being controls all decisions here, even if only "marginally", lies the difference to mere random control, which is probably generally considered insufficient; cf. *Scholz* in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 21 marginal no. 27.

[762] *Scholz* in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 21 marginal no. 34.

[763] On this point *Scholz* in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 21 marginal no. 35.

[764] Cf. for example *Baumgartner/Gausling*, ZD 2017, 308.

[765] On technical and organisational measures in connection with Art. 32 GDPR *Joos/Meding*, CR 2020, 834. Art. 32 GDPR, like Art. 25 GDPR, concretises the requirements of Art. 24 GDPR; cf. only *Hansen* in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 25 marginal no. 14.

implement the principles of Article 5 GDPR.[766] The goal is data protection "built into" the processing of personal data,[767] which deserves interest especially in the context of AI and on which certain hopes are also focused.[768]

In Germany, the so-called Standard Data Protection Model (SDM) was developed as a method "to transform the legal requirements (especially) from Article 5 GDPR into technical and organisational measures, which are described in detail in the SDM reference measures catalogue. It thus supports the transformation of abstract legal requirements into concrete technical and organisational measures".[769] Furthermore, the Data Protection Conference (DSK) has recommended that, "in analogy to Article 25 GDPR all data protection requirements should be taken into consideration in the system development stage".[770]

As convincing as the idea of "built-in data protection" is at first glance, not least from the point of view of effectiveness, it is important to warn against exaggerated expectations. The fact that in reality restraint is called for results on the one hand from the fact that the requirements of the GDPR cannot be fulfilled by technology alone and on the other hand also from the fact that in view of the different official language versions of the regulation, it is not at all clear to what extent the regulation even requires a "technology that promotes data protection".[771] Irrespective of this, the idea seems to have hardly found its way into the practice of software and hardware development so far. [772]

Critics argue that the provision in Article 25 of the GDPR is "technology-neutral". It is true that the provision aims at a "freedom-promoting design of technology". However, they argue, it is too abstract and, moreover, contains too many reservations. For this reason, those responsible "use the technology neutrality of this regulation in order not to have to deal with the task of designing systems in

[766] For more details on the measures that can be considered in this respect, see Recital 78.

[767] Cf. *Hansen* in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 25 marginal no. 1.

[768] Cf. also *Klingbeil/Kohm*, MMR 2021, 3.

[769] Das Standard-Datenschutzmodell – Eine Methode zur Datenschutzberatung und-prüfung auf der Basis einheitlicher Gewährleistungsziele, Version 2.0b von der 99. Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder am 17. April 2020 beschlossen, p. 5 f.: https://www.datenschutzzentrum.de/uploads/sdm/SDM-Methode_V2.0b.pdf.

[770] Cf. position paper of the Conference of the Independent Data Protection Authorities of the Federation and the Länder of 06.11.2019.

[771] Cf.only *Hansen* in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 25 marginal no. 16.

[772] Thus *Hansen* in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 17.

accordance with data protection".[773] The concrete proposals for a reform of the GDPR therefore include "promoting and designing technical data protection (privacy by design)".[774] Data protection that is built into the systems would then "not have to be painstakingly reviewed in individual cases – a task that is no longer even possible".[775]

7. Certification

According to Article 42(1), first sentence, of the GDPR, "[t]he Member States, the supervisory authorities, the Board and the Commission shall encourage, in particular at Union level, the establishment of data protection certification mechanisms and of data protection seals and marks, for the purpose of demonstrating compliance with this Regulation of processing operations by controllers and processors".[776] Behind this regulation lies the attitude that data protection is effective primarily when it involves the data controllers. For this there need to be alternative data protection concepts, including data protection audits and data protection certifications, which aim to create market incentives for data protection-compliant behaviour.[777]

As with Article 25 of the GDPR, the certification is only addressed to data controllers (Art. 4 No. 7) and processors (Art. 4 No. 8). In contrast, the regulation does not address the manufacturers and providers of products and services. And yet it is precisely the manufacturers of IT products who could ensure the actual availability of data protection-friendly solutions on the market through conception, research and development. Accordingly, they should also be given the opportunity for certification.[778]

[773] *Roßnagel*, MMR 2020, 222 (227): "Mahnt z. B. ein betrieblicher Datenschutzbeauftragter eine bestimmte Maßnahme datenschutzgerechter Systemgestaltung an, kann er nie nachweisen, dass diese von der abstrakten Anforderung eines 'Privacy by Design' gefordert wird." ("If, for example, a company data protection officer demands a certain measure of data protection-compliant system design, he can never prove that this is required by the abstract requirement of privacy by design.")

[774] See also *Vásquez*, DSRITB 2021, 149.

[775] *Roßnagel/Geminn/Jandt/Richtert*, Datenschutzrecht 2016 "Smart" genug für die Zukunft? Ubiquitous Computing und Big Data als Herausforderungen des Datenschutzrechts, 2016, p. XVI; cf. on the whole also ibid, p. 134 ff.

[776] See also Recital 100.

[777] Cf. *Scholz* in: Simitis/Hornung/Spiecker gen. Döhmann (ed.), Datenschutzrecht, 1st ed. 2019, Art. 42, marginal no. 4.

[778] Cf. *Scholz* in: Simitis/Hornung/Spiecker gen. Döhmann (eds.), Datenschutzrecht, 1st ed. 2019, Art. 42, marginal no. 19; likewise *Haußmann/Thieme*, NZA 2019, 1612 (1618): "Auch ohne verbindliche Verfahren würde es Betriebsparteien helfen, wenn die Hersteller sich mit den zentralen Fragen des Arbeitnehmerdatenschutzes im System vertraut machen und die Antworten auf gängige Fragen in leicht verständlichen Formaten mitliefern". ("Even without binding procedures, it would help operating parties if manufacturers familiarised themselves with the central issues of employee data protection in the system and included the answers to common questions in easily understandable formats".) The authors complain that the relevant provisions of the GDPR are interpreted differently from one Member State to another, especially in the area of employee data protection.

Independent of this, however, the problem still remains under current law that certification is designed as a voluntary measure. The practical effectiveness of the instrument therefore depends on market forces generating sufficient pressure. Since this is quite doubtful, instead of ex ante certification, there are various calls for the provision of a regulatory framework for auditing that accompanies the entire life cycle of an algorithm-based system.[779]

8. Data protection impact assessment

Article 35 of the GDPR contains a provision on the so-called data protection impact assessment.[780] The first sentence of Article 35(1) GDPR reads: “Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data”. According to Article 35(3), a data protection impact assessment is required, for example, in the case of a systematic and extensive assessment of personal aspects of natural persons based on automated processing, including profiling, but also in the case of large-scale processing of special categories of personal data pursuant to Article 9(1) GDPR. In particular, where the effects of automated decisions or preparations for decisions are likely to result in high risks to the rights and freedoms of natural persons, a data protection impact assessment must be carried out.[781] In the case of a systematic and comprehensive assessment of personal aspects of natural persons, the standard example from Article 35(3) GDPR will often also be present.[782] It is currently being discussed whether and to what extent Article 35 GDPR could be a model for a comprehensive technology impact assessment for certain AI applications.[783]

[779] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 442.
[780] Further detail on this *Schürmann*, ZD 2022, 316 (319).
[781] For a comprehensive analysis see *Kaminski*/*Malgieri*, Algorithmic Impact Assessments under the GDPR: Producing Multi-layered Explanations (September 18, 2019), International Data Privacy Law, 2020, forthcoming, U of Colorado Law Legal Studies Research Paper No. 19–28. https://ssrn.com/abstract=3456224.
[782] Cf. position paper of the Conference of the Independent Data Protection Authorities of the Federation and the Länder of 06.11.2019.
[783] Cf. *Ruschemeier*, NVwZ 2020, 446 (446).

9. Fundamental deficits of the current data protection law

The GDPR provides a legal framework for the application of AI. However, it is often unclear which concrete requirements are to be placed on them. The principles of Article 5 of the GDPR are very general and abstract and therefore require a great deal of concretisation, which is atypical for a (directly applicable) regulation.[784] For some of the principles, a "stable" doctrine is still missing. As a rule, a balancing of different interests must be carried out, the results of which are difficult to predict.[785] This openness of the GDPR offers many advantages in view of a dynamically developing technology,[786] but at the same time poses considerable challenges for the legal practitioner.

Whether the existing data protection law, and in particular the existing legal framework in the form of the GDPR, is sufficient to sufficiently contain the application of AI in terms of data protection law must be called into question in view of the many problems highlighted above. There is a great deal of scepticism in the literature. For example, it is sometimes argued that "the basic orientation of data protection law [...] originates from the era of punch cards". But today, dangers arise less from the fact that individual personal data is collected or processed, but rather from the fact that "an algorithm taps mass data from the information stream and then draws conclusions about the personality or behaviour of an individual person by assigning him or her to a comparison group on the basis of specific characteristics".[787] The GDPR is also obviously based on the assumption that one or more data controllers can be identified for all data. However, this premise is also increasingly subject to doubts, which incidentally arise not least from the frequent linking of AI and blockchain technology, since the latter is precisely designed to achieve decentralisation by replacing a single actor with many different actors.[788] Against this background, it can then no longer come as a surprise that some are calling for a fundamental revision of the GDPR.[789]

[784] Accurately *Roßnagel*, ZD 2018, 339 (342).

[785] Cf. again *Roßnagel*, ZD 2018, 339 (342).

[786] In this respect, *Kugelmann*, DuD 2021, 503 (503), for example, is confident: "Data protection law offers security insofar as the legal basis, in particular the General Data Protection Regulation and the national data protection laws, are established.

[787] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 265.

[788] See only *Finck*, Blockchain and the General Data Protection Regulation, 2019, p. 101.

[789] For example, *Roßnagel/Geminn/Jandt/Richtert,* Datenschutzrecht 2016 "Smart" genug für die Zukunft? Ubiquitous Computing und Big Data als Herausforderungen des Datenschutzrechts, 2016, p. XVII, according to which the regulation is based on "an exaggerated understanding of technology neutrality". There are also calls for a "broad debate involving not only political and administrative authorities, but also civil society and academia". This debate, according to one of the desiderata expressed in the literature, must address the question of "what standards should apply to AI processing of personal data, particularly to ensure the acceptability, fairness and reasonability of decisions on individuals": *Sartor*, The

In this context, a reorientation is sometimes recommended that breaks away from the current reference to persons and strives more for a (preventive) "risk regulation",[790] as can now be seen in the draft of an "AI Act". Indeed, due to the possibilities of combining and evaluating data from different sources, the differentiation between personal data on the one hand and mere factual data or anonymised data on the other hand seems increasingly problematic and to a certain extent downright obsolete.[791] The recommended alternative would largely involve "procedural requirements for the general handling and processing of data".[792] The starting point for regulation would no longer be (solely) personal data. Instead, the focus would increasingly be on the means of analysis.[793] Particularly with regard to Big Data, there are calls to supplement the existing "concept of danger prevention" with the "concept of precaution"; the measures proposed in this regard include, for example, the obligatory prediction of possible personal references or limitations on consent that can have negative consequences for third parties.[794] Also discussed is a departure from the concept of protection of personal data and the establishment of protection against "any individual or public negative consequences of information processing […] all automated information processing should trigger at least an obligation to assess what impact it is likely to have".[795]

In any case, it seems advisable to focus more on the idea of "data protection through technological design" expressed in Article 25 of the GDPR, if only because this would address the problem at its root. However, certain functions would have to be specified in a binding manner in order to avert the objection that can arise under current law that a certain measure is not required by the provision.[796] Understood in this way, the idea of anchoring legal rules and their compliance in the code of programs, which is already provided for in Article 25 of the GDPR, has an enormous scope. In any case, the assessment is that the effectiveness of the law can ultimately only be ensured in this way. Legal science is invited to "make

impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020, p. 80 in his study for the European Parliament. The author also makes a number of reform proposals.

[790] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 265.

[791] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 266.

[792] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 264. At the same time he advises, following the existing regulations in Art. 24(1) and Art. 25(1) (privacy by design), Art. 32(1) GDPR (IT security by design) and Art. 35(1) first sentence GDPR, to "focus more on the risk content of the processing activity"; ibid, p. 268.

[793] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 269.

[794] *Roßnagel*, Big Data - Small Privacy? - Conceptual Challenges for Data Protection Law, ZD 2013, 562 (566 f.).

[795] Thus *Purtova*, The Law of Everything. Broad Concept of Personal Data and Future of EU Data Protection Law, Law, Innovation and Technology 2018, 40 (79 f.).

[796] Cf. *Roßnagel*, Technik, Recht und Macht, MMR 2020, 222.

its knowledge of norms and organisation available in a formalised way so that computer science procedures can access and integrate this knowledge".[797]

Other demands – based on the regulation of other risk technologies – aim, among other things, at increased transparency requirements[798] (in particular by expanding the existing information obligations)[799] as well as an "accompanying legality control",[800] which should take into account the "mutability of algorithms".[801] In this respect, not surprisingly, many of the considerations already presented in connection with anti-discrimination law come up again.[802]

Furthermore, it seems plausible to call for a stronger focus on collective legal protection (not only) in data protection law. Already in terms of substantive law, there is much to be said for "understanding the protection of privacy not only as a task of individual pursuit of rights", but also for "conceiving and developing it more strongly, including as a collective protection of goods".[803] Quite independent of this, however, the existing asymmetry of information must be taken into account, which makes it seem obvious that further protection mechanisms beyond the existing instruments of collective protection (in Article 80(1) and (2) GDPR) need to be established.[804]

VI. Occupational health and safety

The use of AI raises numerous occupational safety and health issues that can only be touched on here. This is all the more true when one considers that AI is often employed together with robotic technologies, which leads to a whole series of additional problems.

[797] *Herberger*, NJW 2018, 2825 (2828); cf. also *Brownsword,* Law, Technology and Society - Reimagining the Regulatory Environment, 2019.

[798] Cf. in this respect *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 283 ff.

[799] Cf. *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 284 ff.

[800] Cf. *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 363 ff.

[801] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 363.

[802] Cf. G. IV. 3.

[803] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 370.

[804] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 274, who calls for "new mechanisms of state control and market supervision"; see also *De Stefano*, Valerio/*Wouters*, *Mathias*: AI and digital tools in workplace management and evaluation – An assessment of the EU's legal framework, May 2022, p. 65 f.; *de Hert*/*Lazcoz*, Radical Rewriting of Art. 22 GDPR on machine decisions in the Ai era, European Law Blog.

The use of human resource analytics, for example, is anything but unproblematic from an occupational health and safety perspective. It is often accompanied by numerous promises, including toward employees. For example, they are supposed to be enabled to better recognise individual opportunities to optimise their performance and thus to advance their career and personal development in a more targeted manner. Regardless of how one feels about these promises, however, the considerable physical and especially psychosocial risks that go hand in hand with them should not be underestimated. This is especially true if employers, with the tools of observation and subsequent analysis they have at their disposal, always "know more" than the worker, who can easily develop a feeling of powerlessness and "being at the mercy" of the employer. According to the European Agency for Safety and Health at Work (EU-OSHA): "How can workers be sure that decisions are being made fairly, accurately and honestly if they do not have access to the data that their employer holds and uses? OSH risks of stress and anxiety arise if workers feel that decisions are being made based on numbers and data that they have neither access to nor power over".[805]

The increasing use of robotics is beneficial in many respects because it makes it possible to use machines where there are ergonomic risks for humans or dangers from the use of chemicals. Hazards to workers can thus be significantly minimised.[806] What's more, the use of robots may for instance even make it possible for older workers to continue doing their jobs. In addition, it is often the case that routine tasks can easily be transferred to robots.[807] Recently, for example, there have been reports of robots learning the body language of their "human workmates" (and thus being better able to adapt to them).[808] AI also appears on the scene in the form of collaborative robots (cobots). These are becoming increasingly popular with companies.[809] Amazon, for example,

[805] European Agency for Safety and Health at Work, OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, Discussion Paper, 2019, p. 4.

[806] European Agency for Safety and Health at Work, Foresight on new and emerging occupational safety and health risks associated with digitalisation by 2025, European Risk Observatory Report, 2018, p. 89. The fact that the use of robots is suitable for relieving people of strenuous work is made clear by the example of so-called exoskeletons. These dock onto the body from the outside, so to speak, and support people in performing work tasks either purely mechanically (passive exoskeleton) or (additionally) with the help of digital data evaluation (active exoskeleton); see *Martini/Botta*, NZA 2018, 625.

[807] Further positive examples in *Beyerer/Müller-Quade et al.* Protecting AI systems, preventing misuse - measures and scenarios in five application areas, white paper, n.d., p. 28.

[808] Cf. *Knight*, This Warehouse Robot Reads Human Body Language – Machines that understand what their human teammates are doing could boost productivity without taking jobs, 28 Jun 2022. https://www.wired.com/story/warehouse-robot-reads-body-language/.

[809] A study published in June of 2022 shows a 45% increase in the global market for cobots in the year 2021, reaching a record high of 31, 325 units delivered: https://
www.k-zeitung.de/cobots-markt-legt-2021-um-45-zu/.

reportedly has about 100,000 AI-assisted cobots in use, and has managed to reduce the training time for new employees to two days.[810] But cobots are not without problems in terms of occupational health and safety. For example, the complex interplay of sensors, hardware and software and their connectivity is a source of danger; physical hazards for workers are particularly likely if connections are unstable or faulty.[811] There are also concerns that the increasing mobility and decision-making autonomy of cobots based on self-learning algorithms could make their actions less predictable for the people working with them. There is also a risk that workers will fall behind the pace and level of work of a cobot and face increasing pressure to keep up. Finally, the increasing use of cobots reduces contact with colleagues, with potentially significant psychosocial consequences.[812]

Special challenges also arise in connection with the increasing use of augmented or virtual reality. While its advantages are obvious – consider that they make it possible to carry out maintenance work remotely – people working in these environments can experience cognitive disturbances, even disorientation. And if meetings increasingly take place in "virtual reality", this may increase efficiency and reduce the travel budget, but it comes at the expense of "real" social interaction.[813] Incidentally, this should also be considered in relation to the vision – which is increasingly taking shape in practice – of (hybrid) working in the so-called metaverse.[814]

The use of so-called chatbots, technical dialogue systems with which communication can take place via text input or speech, also raises questions. It is true that chatbots can relieve people of repetitive tasks. But it is advised that employees be trained sufficiently, so that they understand the role and function of chatbots in the workplace and can better interpret the support they are able to provide.[815]

[810] *Moore*, OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, European Agency for Safety and Health at Work Discussion Paper, 2019, p. 6.

[811] *Moore*, OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, European Agency for Safety and Health at Work Discussion Paper, 2019, p. 6.

[812] European Agency for Safety and Health at Work, Impact of Artificial Intelligence on Occupational Safety and, Health, Policy Brief, 2021, p. 1; cf. on the whole also *Jansen/van der Beek/Cremers/Neerincx/van Middelaar,* Emergent Risks to Workplace Safety, Working in the Same Spot as a Cobot, Report for the Ministry of Social Affairs and Employment, 2018, p. 50 et seq.

[813] European Agency for Safety and Health at Work, Foresight on new and emerging occupational safety and health risks associated with digitalisation by 2025, European Risk Observatory Report, 2018, p. 37.

[814] On this cf. only *Purdy*, How the Metaverse Could Change Work, 5 April 2022. https://hbr.org/2022/04/how-the-metaverse-could-change-work.

[815] *Moore*, OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, European Agency for Safety and Health at Work Discussion Paper, 2019, p. 7.

Other areas where AI is used in the workplace are also viewed critically in the literature. This applies, for example, to digital assistance systems such as wearables, glasses with virtual reality functionality or tablets, used in particular in connection with internal logistics and lean production.[816] In this respect, the danger of work intensification and performance compression is pointed out in particular, as employees constantly receive instructions from these systems.[817] Overall, it is emphasised in the literature that digital assistance systems enable real-time control of work processes and individual guidance of employees and thus indicate "tendencies [...] towards control of work".[818] Quite aside from this, there is a risk – in the longer term – of deskilling, as employees are increasingly reduced to carrying out "modular activities".[819] There are also fears of an increasing loss of work autonomy,[820] as is also the case in call centres, for example, when employees have to follow a given script word for word.[821] Concerns are also raised by the fact that employees are often exposed to constant monitoring by AI,[822] to which, incidentally, a company's customers are often also subjected.[823] The working conditions of Amazon employees in the USA are a prime example of this. The company's logistics centres are fully monitored by cameras.[824] But the company does not seem to want to stop there. Some time ago, Amazon patented a wristband that tracks exactly where warehouse workers place their hands and can steer them in a different direction through vibrations. The patent states that "ultrasonic tracking of a worker's hands can be used to monitor the performance of assigned tasks".[825] The inventiveness displayed in this area seems limitless: for example, literature reports of a company that has transformed the classic name badge into a

[816] See also, for example, *Butollo/Jürgens/Krzywdzinski, Martin,* WZB Discussion Paper, No. SP III 2018-303, 2018, p. 11 ff.

[817] *Moore*, OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, European Agency for Safety and Health at Work Discussion Paper, 2019, p. 7; also *Adrian Todoli Signes*, Making algorithms safe for workers: occupational risks associated with work managed by artificial intelligence, p. 6 f.: https://doi.org/10.1177/10242589211035040.

[818] Cf. *Butollo/Jürgens/Krzywdzinski*, From lean production to Industrie 4.0. More autonomy for employees? WZB Discussion Paper, No. SP III 2018-303, 2018, p. 17.

[819] *Moore*, OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, European Agency for Safety and Health at Work, Discussion Paper, p. 7 f.

[820] See only *Todoli-Signes*, Making algorithms safe for workers: occupational risks associated with work managed by artificial intelligence, p. 6 ff.: https://doi.org/10.1177/10242589211035040.

[821] See also *Doellgast/O'Brady*, Making call centre jobs better: The relationship between management practices and worker stress A Report for the CWA, 2020, p. 4.

[822] Cf. most recently *Todoli Signes*, Making algorithms safe for workers: occupational risks associated with work managed by artificial intelligence, p. 4 ff.: https://doi.org/10.1177/10242589211035040.

[823] See *Levy/Barocas*, Refractive Surveillance: Monitoring Customers to Manage Workers, International Journal of Communication 2018, 1166, noting that companies are increasingly using customer data to manage employees.

[824] Cf. on this most recently *Gurley*, Internal Documents Show Amazon’s Dystopian System for Tracking Workers Every Minute of Their Shifts. https://www.vice.com/en/article/5dgn73/internal-documents-showamazons-dystopian-system-for-tracking-workers-every-minute-of-their-shifts; cf. also Gilbert/Thomas/Pissarides/Al-Izzi/Miller/Burnell, The Amazonian Era: How algorithmic systems are eroding good work, Institute for the Future of Work, 2021.

[825] Thus *Hanley/Hubbard*, Eyes Everywhere: Amazon's Surveillance Infrastructure and Revitalising Worker Power, 2020, p. 9 with further references.

monitoring device. It has a microphone that records conversations, a Bluetooth and infrared sensor that monitors where an employee is and an accelerometer that records their movements. The accompanying software collects data on how much time each employee spends talking and the relationship between talking and listening.[826] Studies have described the impact of this monitoring practice on workers.[827] They sometimes refer to the so-called Hawthorne effect,[828] which refers to changes in the behaviour of people who become aware that they are being watched.[829] In particular, constant surveillance is described as damaging workers' self-esteem and their ability to communicate with colleagues.[830]

So-called gig work, work mediated by platforms, is also viewed critically by the vast majority of commentators. This applies, for example, to employees of delivery services, where there are complaints of time and performance pressure, which emanates not least from the remuneration system (payment according to the number of deliveries, for example) and the assessment of performance by customers. The resulting occupational health and safety risks, especially the psychosocial risks,[831] are obvious.[832] At the same time, it is precisely here that we see serious problems in the enforcement of occupational health and safety regulations by the competent authority.[833] The European Agency for Safety and Health at Work has also identified the extensive algorithmic management by platforms as one of the sources of this type of risk: “[U]sing algorithms to allocate, monitor and evaluate work and the performance and behaviour of platform workers affects the power balance between platform workers, platforms and clients and undermines the autonomy, job control and flexibility of platform workers [...]. This can lead to stress, anxiety, exhaustion and depression, and worsens platform workers’ physical and mental health, safety and overall well-being”.[834]

[826] See *Bales/Stone*, An Invisible Web at Work: Artificial Intelligence and Electronic Surveillance at the Workplace, Berkeley Journal of Employment & Labour Law, 2020, 1 (18).
[827] See for example *Ockenfels-Martinez/Boparai*, The Public Health Crisis Hidden in Amazon Warehouses, Oakland, CA. Human Impact Partners and Warehouse Workers Resource Center, 2021: https://humanimpact.org/hipprojects/amazon/.
[828] Thus *Hanley/Hubbard*, Eyes Everywhere: Amazon's Surveillance Infrastructure and Revitalizing Worker Power, September 1, 2020: https://www.openmarketsinstitute.org.
[829] https://en.wikipedia.org/wiki/Hawthorne_effect
[830] See *Singh*, Employee Surveillance Rises Alongside Work-from-Home Rates, December 18, 2021: https://icetonline.com.
[831] See only *Bérastégui*, Exposure to psychosocial risk factors in the gig economy: a systematic review. ETUI Report, European Trade Union Institute, 2021: https://www.etui.org/sites/; see also European Agency for Safety and Health at Work, Digital platform work and occupational safety and health: a review, 2021.
[832] *Moore*, OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, European Agency for Safety and Health at Work, Discussion Paper, p. 8. Another problem is the lack of "fairness" of working conditions; see for example *Sühr/Biega/Zehlike*, Two-Sided Fairness for Repeated Matchings in Two-Sided Markets: A Case Study of a Ride-Hailing Platform, Applied Data Science Track Paper 2019.
[833] See only European Agency for Safety and Health at Work, Actions by labour and social security inspectorates for the improvement of occupational safety and health in platform work, Policy Case Study 2022.
[834] See European Agency for Safety and Health at Work, Occupational Safety and Health in Digital Platform Work: Lessons from Regulations, Policies, Actions and Initiatives, Policy Brief 2021, p. 2.

These and other dangers are increasingly being recognised and have already led to some laudable private initiatives.[835] The Commission has also explicitly recognised the "potentially … significant impact on the physical and mental health" of workers in its proposal for a directive to improve working conditions in platform work.[836] Accordingly, under Article 7(2) of the Directive, digital work platforms are required to “(a) evaluate the risks of automated monitoring and decision-making systems to the safety and health of platform workers, in particular as regards possible risks of work-related accidents, psychosocial and ergonomic risks; b) assess whether the safeguards of those systems are appropriate for the risks identified in view of the specific characteristics of the work environment; (c) introduce appropriate preventive and protective measures”. In addition, "They shall not use automated monitoring and decision-making systems in any manner that puts undue pressure on platform workers or otherwise puts at risk the physical and mental health of platform workers".[837]

To summarise up to this point: Technological development, including the increasing use of AI, offers enormous opportunities for the protection of workers' lives and health. But there are also risks that should not be underestimated. In the literature, the following risks are named with regard to automation, robotisation and AI: Performance compression; loss of autonomy; mixing of work and private life; increasing complexity of tasks; constant monitoring; pressure to adapt; and the loss of human interaction in the workplace.[838] As one expert summarised it: “Stress, discrimination, heightened precariousness, musculoskeletal disorders, and the possibilities of work intensification and job losses have already been shown to pose psychosocial risks, including physical violence, in digitised workplaces [...] Indeed, AI exaggerates OSH risks in digitalised workplaces because it can allow increased monitoring and tracking and thus may lead to micromanagement, which is a prime cause of stress and anxiety”.[839]

[835] See for example Partnership on AI, Framework for Promoting Workforce Well-being in the AI-Integrated Workplace, 2020.

[836] Proposal for a Directive of the European Parliament and of the Council on improving working conditions in platform work of 9.12.2021, 2021/0414 (COD), p. 2; see also Commission Staff Working Document - Impact Assessment Report - Accompanying the document Proposal for a Directive of the European Parliament and of the Council to improve the working conditions in platform work in the European Union of 10.12.2021, SWD(2021) 396 final/2 (Recital 38).

[837] According to Art. 6(1)(b), digital workplace platforms must inform platform workers about, among other things, automated decision-making systems that have a significant impact on their safety and health at work.

[838] See European Agency for Safety and Health at Work, Foresight on new and emerging occupational safety and health risks associated with digitalisation by 2025, European Risk Observatory Report, 2018, p. 37.

[839] *Moore*, OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, European Agency for Safety and Health at Work, Discussion Paper, p. 2 f.; see also *this*, The Threat of Physical and Psychosocial Violence and Harassment in Digitalized Work, ACTRAV Bureau for Workers' Activities, ILO, 2018.

In proposing a directive to improve working conditions in platform work, the Commission has acknowledged the impact on workers' physical and mental health. But the use of AI is not limited to the field of platform employment, so regulations should be enacted here as well – or rather, must be enacted, under the general principle of equality.[840] After all, the European Commission has not only presented a proposal for a regulation on AI, but also a proposal for a regulation on machine products,[841] which is to replace the Machinery Directive 2006/42/EC, and also contains legally binding conditions for the use of AI.[842]

VII Liability issues

Liability issues are among the most significant and at the same time most difficult problems that arise in connection with the use of AI systems. These questions naturally also arise in labour law, although to a large extent nothing applies here that is different to general civil law. The overriding importance of civil liability (and specifically tort law) results from the fact that it serves to protect rights and legal assets, with Section 823(1) of the German Civil Code (BGB) explicitly naming life, body, health, freedom and property.[843] At the same time, tort law determines the scope of development that the individual has in personal and economic terms.[844] In this respect, tort law stands in the "tension between the protection of legal interests [of the potentially injured party] and freedom of action [of the potential tortfeasor]".[845] To put it another way, tort law has a preventive function in addition to its compensatory function, in the sense that individuals will regularly orient their behaviour towards the goal of avoiding claims for compensation. In the present context, however, this means quite simply that the modalities of the production and

[840] *Todoli Signes,* Making algorithms safe for workers: occupational risks associated with work managed by artificial intelligence, p. 16: https://doi.org/10.1177/10242589211035040, calls for an overarching occupational safety and health regulation that should start with the programming of an algorithm. However, the discussion of the issue of AI in the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, EU Strategic Framework for Health and Safety at Work 2021-2027, Occupational Safety and Health in a Changing World of Work v.28.6.2021, COM(2021) 323 final, p. 7, seems uninspired..6.2021, COM(2021) 323 final, p. 7, which states that "However, new technologies also bring a number of challenges: i) both because of the increasing irregularity in when and where work is carried out ii) and because of the risks associated with new tools and machinery".

[841] Proposal for a Regulation of the European Parliament and of the Council on machinery products v. 21.04.2021, COM(2021) 202 final.

[842] See for example *Mattiuzzo/Vock/Mössner/Voß*, Sichere Maschine mit - oder trotz - künstlicher Intelligenz, ARP 2021, 188.

[843] Cf. only Staudinger/Hager, Vorbemerkungen zu §§ 823 BGB marginal no. 10.

[844] Cf. in this respect also the resolution of the European Parliament of 20 October 2020 with recommendations to the Commission on the regulation of civil liability in the use of artificial intelligence (2020/2014(INL)), to be presented in more detail below, at Recital B.

[845] Thus *Larenz/Canaris*, SchuldR BT, 13th ed. 1994, p. 350.

use of AI systems depend very substantially on which potential liability consequences are to be taken into account by the tortfeasor.[846] In this context, it can be assumed that an increase in liability creates an incentive for all potential tortfeasors to reduce the risk of damage occurring as far as possible.[847]

In the following, the issue of non-contractual liability will be examined first, followed by contractual liability. The main reason for this is that current legal policy efforts to create specific civil liability rules for AI are primarily aimed at tort liability.

1. Non-contractual liability

a) "Tortiousness" of the machine and "autonomy risk"

The fact that the question of whether and to what extent the machine itself can be a tortfeasor, or, translated into the categories of the German Civil Code (BGB), has the capacity to commit a tort, is discussed in great detail in the context of liability law makes it clear how fundamental many liability issues are. That this question is posed is not surprising, considering what was said above about the question of legal capacity.[848] And indeed, the discussion on the legal capacity of machines was triggered by a resolution of the European Parliament, which aims at nothing other than shaping civil liability for AI systems.

If this question is raised in liability law in particular, it is because robots cannot be denied a certain degree of "autonomy" under certain conditions and thus also an ability to harm others. This is not even remotely comparable to human self-determination. But a certain "freedom" can hardly be denied when a machine "decides" between two or more options without this already being mandated by the controlling software.[849] And those who object that the "decision" is ultimately

[846] Cf. also in this respect the European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for the use of artificial intelligence (2020/2014(INL), where already the first Recital (A.) contains the indication that "the concept of "liability" plays two important roles in everyday life, namely to ensure that a person who has suffered personal injury or damage to property is entitled to seek and obtain compensation from the party who is proven to be liable for the personal injury or damage to property, and to create economic incentives for natural and legal persons to avoid causing personal injury or damage to property in the first place or to factor the risk of paying compensation into their behaviour"; cf. on the whole also *Evas*, Civil liability regime for artificial intelligence - European added value assessment, Study European Research Service, 2020, p. 4 ff. (with further references).

[847] See also *Wagner*, VersR 2020, 717 (718) and later: "A well-adjusted liability law generates incentives for the development of technologies that promise a safety gain compared to the established solutions and discourages the use of technologies that cause higher damage costs than they generate in additional benefits"; ibid. (723).

[848] Cf. G. II.

[849] *Wagner*, VersR 2020, 717 (720) speaks of a "minimal concept of autonomy".

always "pre-programmed" should consider that the "learning ability" of algorithms consists precisely in the ability to improve their decision rules: By developing them further, the system "outgrows the control of its programmer".[850] This even applies in a very special way to artificial neural networks that imitate the human brain, since the program opts one way or another depending on the "state" of the network, with the generated result being fed back into the system so that the links are constantly evolving.[851] Accordingly, there is no getting around the fact that "for the first time in history artefacts can decide for themselves what to do or not do", which is why one cannot easily equate the "behaviour" of the digital system with the behaviour of the person who created it or uses it.[852] However, in the current state of technical development, liability on the part of the machine itself must be rejected for the same reasons as mentioned above: As yet, machines still lack the capacity for autonomous decision-making that could put them on equal footing with natural persons in terms of liability law.

b) Individual questions of fault liability

If we now turn to the preconditions of fault liability, it quickly becomes apparent that every one of its preconditions is put to a serious test when the use of AI is involved, and that there are indeed gaps in liability.[853] This is also the conclusion of a group of experts set up some time ago by the European Commission to examine the current liability law in the Member States for possible weaknesses.[854]

aa) Causality

According to the experts, it is often hard to clarify even the initial question of causality, that is, whether the damage that occurred was caused by the person sued as a tortfeasor. Thus it is often difficult to establish or prove that the cause of

[850] *Wagner*, VersR 2020, 717 (720).

[851] Also *Wagner*, VersR 2020, 717 (720): "The only decisive factor is that the behaviour of the system is not predictable for outsiders, including its creators, but that in the concrete situation of action the system itself generates a decision".

[852] *Wagner*, VersR 2020, 717 (724).

[853] For example, *Veith,* Künstliche Intelligenz, Haftung und Kartellrecht - Zivilrechtliche Verantwortlichkeit beim Einsatz von Künstlicher Intelligenz und Implikationen für das Kartellrecht, 2021, p. 124; in contrast, *Hofmann,* Der Einfluss von Digitalisierung und künstlicher Intelligenz auf das Haftungsrecht, CR 2020, 282, is more "relaxed".

[854] Expert Group on Liability and New Technologies, New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019; cf. also the following comparative law study by *Karner/Koch/Geistfeld*, Comparative Law Study on Civil Liability for Artificial Intelligence, 2020, which also includes the legal situation in the USA.

the damage was a faulty algorithm.[855] This is all the more true because in cases of this kind there can be a considerable distance in time and space between the conduct and the infringement of the legitimate interest.[856] The difficulties would increase even more if the algorithm was modified further and further by machine learning techniques and had therefore become more removed from the "original".[857] In the expert group's analysis, clarifying causality in such cases is often only possible after costly analyses. The fact that the injured party may have a claim for reimbursement in connection with those expenses is little consolation because it is uncertain who can be considered the debtor (and thus the addressee of a claim for damages).[858] Accordingly, in such cases it is worth considering the introduction of strict liability[859] that is not linked to fault, but to the principles of risk causation and risk control.[860]

In connection with causality, there is often talk of the so-called "networking risk", which must be taken into account when using AI. While the "autonomy risk" just described is linked to the fact that with autonomous AI systems, one can or must rightly speak of "misconduct" of the system itself, the term "networking risk" is linked to the fact that AI systems are as a rule based on complex connections between computers, so that when damage occurs, it proves difficult to reconstruct the course of the damage and, above all, to reliably identify the damaging party. Digitalisation thus "shakes the foundations of a premise […], namely the practical possibility of being able to name spheres of responsibility and to delimit different subjects' spheres of responsibility from one another".[861] Legal answers must be found to this as well.

Nevertheless, one may find it comforting with regard to German law that it does provide solutions when several tortfeasors can be considered for an act of infringement. According to Section 830(1), first sentence, para. 2 BGB, in the case of so-called cumulative causality ("if several have caused damage by a jointly

[855] Expert Group on Liability and New Technologies, New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 20; cf. on the problem also *Müller-Hengstenberg/Kirn,* CR 2018, 682; *Teubner*, AcP 2018, 155 (157) even says in this respect that "the most difficult liability gap to correct arises in the case of multiple causality in the case of damaging acts of several computers if these are networked with each other".

[856] Cf. only *Zech*, ZfPW 2020, 198 (206 f.).

[857] Cf. on the so-called "autonomy risk" due to the fact that the "behaviour" of a technical system in the damage situation is not predetermined by a human being but determined by the software, *Hofmann*, CR 2020, 282 (283) following *Teubner*, AcP 2018, 155 (163).

[858] Expert Group on Liability and New Technologies - New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 20.

[859] See indeed Expert Group on Liability and New Technologies - New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 21.

[860] Cf. *Larenz/Canaris*, SchuldR BT, 13th ed. 1994, p. 605.

[861] Thus *Wagner*, VersR 2020, 717 (725).

committed tortious act"), there is a liability of each individual tortfeasor and, what is even more important, according to Section 830(1), second sentence BGB, a liability of each individual tortfeasor even in the case of doubts about causality ("if it cannot be determined which of several participants has caused the damage by his act"). However, this does not lead any further if it is already unclear whether actions were only causal in interaction with others or each on its own. For Section 830(1), sencond sentence BGB, it must be established that one of several actors caused the violation of the legitimate interest or the damage. In addition, it must also be established that each of them committed an act that was specifically suitable for causing the violation of the legitimate interest or the damage. In contrast, the provision does not apply if it cannot be determined whether the violation of the legal interest or the damage was only caused by the interaction of both (possible cumulative causality).[862] As a result, under current law, the "networking risk" can only be countered to a very limited extent.

bb) Illegality and culpability

However, problems also arise with regard to the prerequisites of illegality and culpability. In view of the novelty of the phenomenon of AI, it is unclear what the concrete content of the "traffic (safety) obligations" that must be assumed in this area should be,[863] or how one should substantiate the "care required in traffic" that is to be taken into account in connection with the determination of fault.[864] In line with this, the expert group mentioned above concluded: "Emerging digital technologies make it difficult to apply fault-based liability rules, due to the lack of well established models of proper functioning of these technologies and the possibility of their developing as a result of learning without direct human control".[865] Even more scepticism is sometimes encountered in the literature. For example, the assessment is occasionally expressed that, especially in view of self-learning algorithms and the formation of complex systems, causality, but also culpability, can often no longer be reconstructed. Some therefore call for a "systemic liability", according to which all participants in a system would be

[862] Cf. *Zech*, ZfPW 2019, 198 (207 f.); also *Wagner*, VersR 2020, 717 (733).

[863] Cf. for example *Schmid*, CR 2019, 141 on an obligation for "integrated product monitoring" for automated and networked systems.

[864] See also *Zech*, ZfPW 2019, 198 (210 f.); cf. also *Haagen*, Verantwortung für Künstliche Intelligenz – Ethische Aspekte und zivilrechtliche Anforderungen bei der Herstellung von KI-Systemen, 2021, p. 187 ff.; *Herbosch*, The Diligent Use of AI Systems: A Risk Worth Taking?, EuCML 2022, 14.

[865] Expert Group on Liability and New Technologies - New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 23.

fundamentally liable;[866] this would represent a step back from the "model of individual attribution of causality and fault".

cc) AI as "vicarious agent”

Liability for third parties also raises questions in connection with the use of AI systems. In Germany, Section 831 of the German Civil Code (BGB) applies here, according to which "whoever orders another to perform a task" is liable to pay compensation if "the other person unlawfully [causes damage] to a third party in the course of performing the task". It seems obvious that such liability could be of eminent importance in the present context. For if the "principal" is liable for the "vicarious agent", then it stands to reason that the person who uses an AI system to perform certain tasks should also be liable. The EU expert group found the argument "quite convincing" that "using the assistance of a self-learning and autonomous machine should not be treated differently from employing a human auxiliary".[867] With regard to Section 831(1) BGB, it should be added that the qualification as a performing agent presupposes dependence and being bound by instructions,[868] whereby it is sufficient "that the principal can restrict or withdraw the activity of the agent at any time or determine it in terms of time and scope".[869] Since this is to be assumed in the present context, it seems more than obvious to apply Section 831 BGB also with regard to AI and in this way to arrive at a liability of the person who uses such a system. [870] The fact that the use of "non-human aids" may be safer in many areas (and the occurrence of damage may therefore be less likely) than with the assistance of natural persons, as the expert group points out,[871] cannot change this.[872]

Even if one can choose the path of an (analogous) application of Section 831(1) BGB, in the end not too much is gained. The provision does not allow for a strict

[866] Thus *Spiecker gen. Döhmann*, CR 2016, 698 (703).

[867] Expert Group on Liability and New Technologies - New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 25. It is in line with this that parts of the literature in Germany at least consider an analogous application of section 831(1)BGB to be justifiable; in this sense *Denga*, CR 2018, 69 (74f.); *Horner/Kaulartz*, CR 2016, 7, 8 f.; also *Grützmacher,* CR 2016, 695; cf. on the whole also *Müller-Hengstenberg/Kirn*, CR 2018, 682 (686).

[868] Cf. only MünchKomm/Wagner, 8th ed. 2020, § 831 BGB marginal no. 14.

[869] For example BGH, NJW 1966, 1807.

[870] Cf. also *Wagner*, VersR 2020, 717 (730): "Thus they correspond to the role of "helper" or "servant" better than humans".

[871] Expert Group on Liability and New Technologies - New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 24.

[872] Interesting in this context, by the way, are recent considerations in the US-American literature to start with the position of employees, who are often replaced by AI, in the question of "liability for AI" and to construct this liability analogously to the liability of the employer for the vicarious agents used by him; so *Diamantis*, Employed Algorithms: A Labor Model of Corporate Liability for AI (October 19, 2021). 72 Duke L.J.: https://ssrn.com/abstract=3945882.

attribution of the "misconduct" of the "digital vicarious agent" to the principal. Rather, as is well known, Section 831(1) BGB regulates a liability for presumed fault or for a presumed breach of the duty of care.[873] Accordingly, the principal always has the option of procuding exonerating evidence pursuant to section 831(1), second sentence, BGB, which should usually be possible, and usually even better than with the involvement of a "human assistant", because the user cannot influence the behaviour of an autonomous digital system anyway.[874]

c) Legal policy initiatives

In view of the numerous question marks that arise with regard to AI systems on the basis of current liability law, the question arises as to whether and to what extent liability law should be redesigned in the light of AI. In the following, a corresponding initiative of the European Parliament and parallel efforts of the European Commission will be presented. Both initiatives are based to a certain extent on the recommendations of the above-mentioned expert group; these should therefore be presented first. Furthermore, it should be pointed out that the planned AI Regulation is also likely to have an impact on the question of civil liability, because it seems anything but far-fetched, for example, to adopt from its provisions standards for what is to be regarded as conduct in breach of duty and culpable conduct within the framework of Section 823(1) BGB.[875]

aa) Recommendations of the expert group

According to the expert group set up by the European Commission, certain properties are characteristic of AI systems. They are highly complex, both in terms of the interaction of different hardware and software components[876] and in terms of the "internal complexity" of algorithms. Accordingly, the contributions to

[873] *Larenz/Canaris*, SchuldR BT, 13th ed. 1994, p. 475.

[874] So also *Wagner*, VersR 2020, 717 (730), who accordingly predicts that in addition to the duties of the principal to select, supervise and instruct under section 831 of the Civil Code, there will be a development of "digital organisational duties" under section 823(1) of the Civil Code and considers an abolition of the exculpatory evidence under subsection (1) sentence 2 to be "urgently required" in terms of legal policy (ibid., 736).

[875] For more details see *Grützmacher*, CR 2021, 433 (442).

[876] To this could be added that for numerous applications an "intertwining of the individual actions of humans and algorithms" can be observed; according to *Teubner,* AcP 2018, 155 (189), who believes for this "compound risk" that "the risks arising here due to the almost unresolvable intertwining of the individual actions of humans and algorithms [...] can be better countered by identifying the human-algorithm association, the hybrid as such, as a common point of attribution for actions, rights and obligations".

accountability are often difficult to determine.[877] The complexity of AI systems is accompanied by a considerable lack of transparency. "The more complex emerging digital technologies become," the expert group states, "the less those taking advantage of their functions or being exposed to them can comprehend the processes that may have caused harm to themselves or to others." The systems are also "open" in the sense that they are designed for constant updating and interaction. AI systems are also "unpredictable" in the sense that it is often difficult to anticipate their impact at the time they go live. Furthermore, they are "data-driven", that is highly dependent on external data. Finally, they are "vulnerable" because they need to be updated and are geared towards interaction.[878]

In light of these particularities, the expert group recommends a two-tier liability system consisting of strict liability[879] and fault liability. As far regards the former, the expert group considers it as an "appropriate response to the risks posed by emerging digital technologies, if, for example they are operated in non-private environments and may typically cause significant harm". The experts conclude that the person who controls the risk and benefits from the application should be liable. This indeed addresses aspects that also characterise the elements of strict liability in Germany. Here, too, the person who causes or controls the danger is liable.[880] And here, too, the "idea of the coherence of advantage and corresponding risk" plays a significant role.[881] However, the considerations of the expert group have drawn some criticism. For example, it is argued that the distinction between applications with a lower and higher risk propensity, which ultimately underlies them, is at least in need of clarification and that some applications are downright "black swans" with regard to the risks they pose.[882]

As far as fault-based liability is concerned, the experts advocate the recognition of a number of "duties of care" on the part of AI system operators. These should

[877] In this respect, *Teubner*, AcP 2018, 155 (163)) speaks of a "networking risk", cf. also *Hofmann*, CR 2020, 282 (283).

[878] Cf. Expert Group on Liability and New Technologies -New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 33 f.

[879] *Evas*, Civil liability regime for artificial intelligence - European added value assessment, Study European Research Service, 2020, p. 12 ff.

[880] Cf. *Teubner*, AcP, 2018, 155 (184), who insists on the difference that in the present context "liability for the unlawful misconduct of the autonomously deciding software agent" is at issue, rather than liability for the lawful use of dangerous equipment. The risk of digital decision-making autonomy is "in principle of a different nature than the risks that are relevant in the previous cases of strict liability". Accordingly, the author calls for an independent "digital assistance liability", whereby with regard to the addressee of liability, an orientation towards the "attribution and liability unit" should take place, in which the case law for motor vehicle endangerment liability combines driver, owner and insurance; ibid, 31; cf. also *Zech,* Entscheidungen digitaler autonomer Systeme: Empfehlen sich Regelungen zu Verantwortung und Haftung?, Gutachten A zum 73. Deutschen Juristentag, 2022, A 77 ff.

[881] *Larenz/Canaris*, SchuldR BT, 13th ed. 1994, p. 605.

[882] Thus *Bertoloni*, Artificial Intelligence and Civil Liability - Study requested by the JURI Committee, 2020, p. 77 f.

relate, among other things, to the "choosing the right system for the right task and skills" and also include a duty to monitor the system and maintain it.[883] Moreover, the injured party should be helped by facilitating evidence, for which the probability that the technology at least contributed to the damage and an existing asymmetry of information between the operator and the injured party would play a role.[884] In individual cases, the burden of proof should even be reversed, especially if proving a violation would involve disproportionate effort and costs for the injured party.[885]

Finally, as regards possible vicarious liability, according to the experts, the same conditions should apply in principle for this as for liability for the misconduct of a (human) third party: "If harm is caused by autonomous technology used in a way functionally equivalent to the use of human auxiliaries, the operator's liability for making use of the technology should correspond to the otherwise existing vicarious liability regime of a principal".[886] Translated into the categories of German law, this would mean that the AI system would have to have "acted" unlawfully and in principle also culpably.[887] In the estimation of the expert group, this should only not apply if the performance of the AI system is superior to human performance. In this case, liability would require that the user of the system had failed to use a less error-prone system. This would fit in with the fact that Section 831(1) BGB is ultimately linked to the fact that the principal did not carefully select the assistant or did not carefully supervise and instruct that assistant.[888]

It has already been pointed out above that the considerations of the expert group have not met with unmitigated approval in the literature. As already mentioned, the division of AI into "low risk" and "high risk" has been criticised, which, according to the experts, should decide on the intervention of strict liability.[889] However, there are also fundamental reservations about an approach that seeks to regulate AI across the board. Instead, there are calls in the literature for an approach that differentiates between the individual systems and thus seeks specific solutions. [890]

[883] Expert Group on Liability and New Technologies - New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 44.

[884] Expert Group on Liability and New Technologies - New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 49 f.

[885] Expert Group on Liability and New Technologies - New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 52.

[886] Expert Group on Liability and New Technologies - New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies, 2019, p. 45.

[887] Cf. *Larenz/Canaris*, SchuldR BT, 13th ed. 1994, p. 479.

[888] *Larenz/Canaris*, SchuldR BT, 13th ed. 1994, p. 481.

[889] See *Bertolini*, Artificial Intelligence and Civil Liability - Study requested by the JURI Committee, 2020, p. 77.

[890] See in this respect *Bertolini*, Artificial Intelligence and Civil Liability - Study requested by the JURI Committee, 2020, p. 89. The author himself advocates a so-called risk management approach, on the basis of which liability is allocated to the

bb) Resolution of the European Parliament

On 20 October 2020, the European Parliament adopted a resolution containing recommendations for the design of a civil liability regime; attached to the resolution is a concrete draft regulation.[891] According to its Article 1, the draft regulation envisages the liability of the operator of an AI system.[892] With regard to manufacturers, on the other hand, a revision of the Product Liability Directive is suggested,[893] which, according to the plans of the European Parliament, is to be transformed into a regulation.[894] The European Parliament here employs a broad concept of "AI systems"; this encompasses a "large group of different technologies, including simple statistics, machine learning and deep learning".[895]

As regards the challenges posed by AI systems in terms of liability law, the Parliament sees them firstly in the fact that AI systems" could lead to situations in which their opacity could make it extremely expensive or even impossible to identify who was in control of the risk associated with the AI-system, or which code, input or data have ultimately caused the harmful operation". This factor could "make it harder to identify the link between harm or damage and the behaviour causing it, with the result that victims might not receive adequate compensation".[896] But Parliament also sees other challenges, namely those arising from "the connectivity between an AI-system and other AI-systems and non-AI-systems, their dependency on external data, their vulnerability to cybersecurity breaches as well as from the design of increasingly autonomous AI-systems using, inter alia, machine-learning and deep-learning techniques".[897]

party "which is best able (i) to identify a risk, (ii) to control it and minimise it through its decisions, and (iii) to manage it - ideally by pooling and distributing it among all other parties - possibly through insurance and/or compensation funds [...]"; ibid., p. 99.

[891] Resolution of the European Parliament of 20 October 2020 with recommendations to the Commission on a civil liability regime for the use of artificial intelligence (2020/2014(INL); cf. on this, for example, *Müller-Hengstenberg/Kirn,* MMR 2021, 376; *Heiss*, Europäische Haftungsregeln für Künstliche Intelligenz, EuZW 2021, 93.

[892] Cf. Art. 1, according to which the regulation lays down rules for civil liability claims by natural and legal persons against operators of AI systems.

[893] See also *Bertolini*, Artificial Intelligence and Civil Liability - Study requested by the JURI Committee, 2020, p. 60 ff.

[894] Cf. in this respect Recital 8, which contains some substantive desiderata: a definition of "products" clarifying whether and to what extent digital content and digital services fall within its scope; the possible adaptation of terms such as "damage", "defect" and "producer"; a definition of "producer" covering manufacturers, developers, programmers, service providers as well as back-end operators; the possible introduction of a reversal of the burden of proof "in clearly defined cases and after a detailed assessment". In doing so, Parliament stresses that "any update of the product liability framework should be accompanied by an update of Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety (16) to ensure that AI systems comply with the principle of 'safety and security by design'".

[895] Recital F.

[896] Recital H.

[897] Recital I.

A closer look at the proposal reveals the following: First, Parliament explicitly opposes leaving everything as it is under liability law, as the"complexity, connectivity, opacity, vulnerability, the capacity of being modified through updates, the capacity for self-learning and the potential autonomy of AI-systems, as well as the multitude of actors involved" pose particular challenges that need to be addressed.[898] Parliament then argues against giving legal personality to AI systems, arguing that personal injury or damage to property is "nearly always the result of someone building, deploying or interfering with the systems".[899] While it concedes that causality problems and difficulties of proof do arise, it considers that these could be addressed "by making the different persons in the whole value chain who create, maintain or control the risk associated with the AI-system liable".[900] Finally, it states that it must be "clear that whoever creates, maintains, controls or interferes with the AI-system, should be accountable for the harm or damage that the activity, device or process causes". [901]

Finally, the above-mentioned division into two parts is central:[902] Article 4 of the Draft Regulation is to apply to "high-risk AI systems". Paragraph 1 provides that the operators of such systems are subject to strict liability for all personal injury or damage to property "caused by a physical or virtual activity, device or process driven by an AI-system".[903] In contrast, all other AI systems are subject to fault-based liability under Article 8(1) of the Draft Regulation. Liability under Article 8(2) of the Draft Regulation is excluded if the operator proves that the damage was caused through no fault of its own.[904] However, operators " shall not be able to exonerate themselves from liability by arguing that [...] the harm or damage was caused by an autonomous activity, device or process driven by their AI system". Only in the case of force majeure is the operator not liable. The case of contributory

[898] Under 6.
[899] In Recital 6 of the draft regulation, it is even explicitly emphasised that "AI systems have neither legal personality nor human conscience and that their only task is to serve humanity".
[900] Under 7.
[901] Recital (8).
[902] Critical *Sousa Antunes*, Civil Liability Applicable to Artificial Intelligence: A Preliminary Critique of the European Parliament Resolution of 2020, December 5, 2020: https://ssrn.com/abstract=3743242.
[903] According to Art. 4 para. 2 p. 1, "all high-risk AI systems and all critical sectors where they are used" are to be listed in the Annex to the Regulation. Art. 4 para. 3 p. 1 excludes an exemption from liability for cases in which operators "argue that they acted with due diligence or that the damage was caused by autonomous activities, devices or processes controlled by their AI system". Exclusion of liability can only be considered in cases of force majeure according to Art. 4 para. 3 sentence 2.
[904] Cases mentioned in this respect are those where "a) the AI system [...] was activated without the operator's knowledge, while all reasonable and necessary measures were taken to prevent such activation outside the operator's control, or b) all the following measures [...] were taken with due diligence: Selection of an appropriate AI system for the task and capabilities in question, proper commissioning of the AI system, monitoring of activities and maintenance of operational reliability by regularly installing all available updates".

negligence is regulated by Article 10 of the Draft Regulation. According to Article 11, first sentence of the Draft Regulation, several operators of an AI system are jointly and severally liable.

It is noteworthy that, as already mentioned, liability under Article 1 of the Draft Regulation is to attach to the user and not, for example, the manufacturer of AI. This has been criticised in the literature with reference to the fact that the user, in contrast to the manufacturer, usually has little influence on the "behaviour" of the system.[905] The question is of fundamental importance, as it raises the problem of the distribution of responsibilities between user and manufacturer in general. In the present context, however, it does not really arise because the Regulation – in Article 3 lit e and f – distinguishes between "frontend operators" and "backend operators" and defines the latter as "the manufacturer of the digital autonomous system, or in any case the most important actor within the group of those persons (who) can be considered as manufacturers".[906] The definitions of "frontend operator"[907] and "backend operator"[908] show that the boundaries are not always easy to draw.

[905] Cf. *Wagner*, ZEuP 2021, 545 (551); on the whole, cf. most recently also European Data Protection Board, letter to the European Commission on adapting liability rules to the digital age and artificial intelligence (AI), 25.02.2022, which calls for a "clear distribution of roles" with regard to the addressees of liability.

[906] Thus *Wagner*, ZEuP 2021, 545 (552); on the background ibid., 571.

[907] "'Frontend operator' means any natural or legal person who exercises a degree of control over a risk connected with the operation and functioning of the AI-system and benefits from its operation".

[908] "'Backend operator' means any natural or legal person who, on a continuous basis, defines the features of the technology and provides data and an essential backend support service and therefore also exercises a degree of control over the risk connected with the operation and functioning of the AI-system".

cc) Report of the Commission

In 2020, the European Commission presented a report addressing, among other things, the issue of AI liability and product safety.[909] In it, the Commission identified, among other things, connectivity, autonomy, data dependence and opacity of AI systems as challenges for existing product safety law.[910] With regard to the aspect of "autonomy" in particular, the Commission stated that "[t]here may be also situations in the future where the outcomes of the AI systems cannot be fully determined in advance. In such a situation, the risk assessment performed before placing the product on the market may no longer reflect the use, functioning or behaviour of the product.".[911] It is also noteworthy that the Commission's report states that "the future "behaviour" of AI applications could generate mental health risks for users deriving, for example, from their collaboration with humanoid AI robots and systems at home or in working environments". It concludes that "mental health risks should be explicitly covered within the concept of product safety in the legislative framework".[912] With regard to liability, one of the main problems identified is what was referred to above by the term "networking risk" developed in the literature. The Commission's report states: “AI applications are often integrated in complex IoT environments where many different connected devices and services interact. Combining different digital components in a complex ecosystem and the plurality of actors involved can make it difficult to assess where a potential damage originates and which person is liable for it. Due to the complexity of these technologies, it can be very difficult for victims to identify the liable person and prove all necessary conditions for a successful claim, as required under national law. The costs for this expertise may be economically prohibitive and discourage victims from claiming compensation”.[913]

Not least with a view to reversing the burden of proof in favour of the injured party, the Commission carried out a public consultation, which was concluded in January

[909] Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee, Report on the security and liability implications of Artificial Intelligence, the Internet of Things and Robotics, 19.2.2020, COM(2020) 64 final; cf. also the White Paper "On Artificial Intelligence - A European Approach to Excellence and Trust", 19.2.2020, COM(2020) 65 final/2.

[910] Report, p. 6 ff.

[911] Report, p. 8; cf. on the "autonomy" of AI also *Buiten/de Streel/Peitz*, EU Liability Rules for the Age of Artificial Intelligence, 2021, p. 26 ff.

[912] Report, p. 10.

[913] Report, p. 17.

2022. The Commission intends to present its conclusions in the third quarter of this year.

2. Contractual liability

a) "Traffic" duties as contractual duties to protect

The above remarks were aimed at the area of non-contractual liability, that is, what is called tort and strict liability in Germany. The European Parliament's Draft Regulation just presented also has only this area in mind. Article 2(3) of the Draft Regulation explicitly states: “This Regulation is without prejudice to any additional liability claims resulting from contractual relationships, as well as from regulations on product liability, consumer protection, anti-discrimination, labour and environmental protection between the operator and the natural or legal person who suffered harm or damage because of the AI-system and that may be brought against the operator under Union or national law”.

Non-contractual liability is also relevant for employees, especially when they claim damages from someone other than their contractual partner. However, it is also relevant when the employee claims damages from the employer. In this respect, the so-called accumulation principle applies under German law. This means that – unlike in France, for example – contractual and tortious claims are not mutually exclusive, but can co-exist.[914] However, the former – and this is the crucial point – are not limited to the explicitly assumed primary duties. Rather, the duties of care of tort law, the so-called "traffic duties", are integrated into the contract as duties of protection. It follows from Section 241(2) BGB that there is no difference in content between tortious duties of care and contractual duties of protection. The breach of a tortious duty of care thus automatically triggers contractual liability, which cumulatively occurs alongside tort liability.[915]

[914] Cf. only MünchKomm/Wagner, 8th ed. 2020, before § 823 BGB, marginal no. 82.
[915] Cf. MünchKomm/Wagner, 8th ed. 2020, § 823 BGB, marginal no. 449.

b) Digital aider and abettor liability?

However, even according to the principle of cumulation, contractual and tortious liability each follow their own rules, and in this respect, with regard to the former, there are, as is well known, the special features that the debtor of a contractual obligation must allow the fault of his vicarious agent to be imputed to him according to Section 278 BGB, while the possibility of exculpatory evidence is open to him in the case of the tortious vicarious liability of Section 831(1) BGB. In the present context, this then leads to the question of whether Section 278 BGB is applicable to digital vicarious agents by analogy. This question is predominantly answered in the negative, and rightly so. Section 278 BGB defines the liability of the debtor in such a way that the debtor is responsible for the fault of the vicarious agent as if it were the debtor's own fault. Section 278 BGB thus presupposes fault on the part of the vicarious agent.[916] However, since the question of fault cannot (yet) be meaningfully posed with regard to machines, an analogy to Section 278 BGB must be ruled out.[917]

A brief look at (contractual and tortious) liability shows two things in particular. On the one hand, liability in connection with AI raises specific and, above all, entirely new problems (keywords are "autonomy risk" and "networking risk") and, on the other hand, the existing law only provides some adequate answers.

[916] This is, of course, not uncontroversial; cf. for example the comments in MünchKomm/Grundmann, § 278 BGB marginal no. 50 (and fn. 230).

[917] Cf. in this respect only Staudinger/Caspers, 2019, § 278 BGB marginal no. 5 with further references; likewise *Kumkar*, K & R 2020, 801 (806 f.), also *Heiderhoff/Gramsch,* ZIP 2020, 1937 (1939) with the argument that it is "not very plausible" to "provide for a threshold beyond which the change from simple machine liability to attribution according to § 278 BGB begins"; in contrast, *Hacker*, RW 2018, 243 (255) is in favour of a differentiation according to the degree of "autonomy" of the system. In contrast, *Lohmann/Preßler*, RDi 2021, 538, argue for an analogous application of § 278 BGB.

VIII. Workplace co-determination

Workplace co-determination is a central element of employee protection in Germany. The following remarks show that the importance of co-determination will increase in the future because AI and Big Data affect recognised purposes of co-determination in a special way and also provide further arguments for safeguarding effective co-determination.

1. Recognised purposes of co-determination

Looking first at the widely recognised purposes of co-determination, it should be noted that – unlike collective bargaining autonomy – it serves less to equalise unequal positions of power than to take account of the employee's position as a "member of the workforce".[918] The guiding principle is thus the idea that in the employment contract the employee "undertakes to take on a merely generally defined function (task) within an external area of work or life",[919] which leads to a dependency which must be limited with a view to the individual's right to self-determination.[920] And the fact that the legislature has taken the path of establishing a collective representation of interests is not only explained by practical considerations, but also by the fact that the works constitution is also about "balancing divergent interests of the employees among themselves";[921] that the employer would be overburdened with such a balancing is obvious.

These interrelationships become particularly clear if one recalls the report of the Co-Determination Commission from 1970.[922] In this report, the focus was on corporate co-determination. However, it still contains valid considerations on the foundations of co-determination and thus also on workplace co-determination. The starting point of these considerations is the protection of human dignity and respect for the right of personality as a central value decision. The report states: “The

[918] GK-BetrVG/Wiese, 12th ed. 2021, Einl., marginal no. 75.
[919] GK-BetrVG/Wiese, 12th ed. 2021, Einl., marginal no. 76.
[920] Also GK-BetrVG/Wiese, 12th ed. 2021, Einl., marginal no. 76.
[921] GK-BetrVG/Wiese, 12th ed. 2021, Einl., marginal no. 77; on the value-based foundation of the works constitution also *Reichold*, Betriebsverfassung als Sozialprivatrecht, 1995, p. 486 ff.
[922] BT-Drucks. VI/334, p. 56 ff.

subject of this value decision is the appropriate shaping of the position of the individual in the organised community of a company. In terms of content, the value decision is based on the fundamental commitment to the dignity of the person, to the inviolable and inalienable human rights as the basis of every human community and to the right to free development of the personality (Articles 1 and 2 Basic Law). First of all, in general terms, it means that the subordination of the employee to external management and organisational power in the company is only compatible with his self-determination, the possibility granted to him by law to choose his own purposes and to develop his own initiatives, as long as it finds its counterpart in the form of freedom of participation in the decisions that regulate and shape the work process".[923]

The Co-determination Commission derives from this value decision that a works constitution is necessary, since the employment contract cannot capture the "social and societal reality of the enterprise as a social association".[924] Co-determination "gives legal expression to the social reality of the enterprise as a common, purpose-oriented organisation and as a common condition of all the employees concerned, which the individual contract is not able to give it".[925] The Co-determination Commission considered that by concluding the employment contract, the employee submitted to both the "entrepreneurial planning competence" and the employer's right to issue instructions. According to the Commission, the employee was faced with "structural circumstances" which "could not be dealt with by the legal means of voluntary cooperation, that is, by means of contractual freedom alone". And further: "Just as for the conclusion or non-conclusion of the employment contract, the employee can also not be referred to his 'freedom' to conclude contracts or refuse to conclude them for submission to the planning competence of the enterprise and a right of instruction of the employer. Such equal freedom is lacking in both cases and thus also the full realisation of the consensual principle characterising the contract. The economic compulsion to conclude an employment contract continues in the necessity to consent to the planning competence of the enterprise, its concretisation through the right to issue instructions and thus the existence and exercise of powers of command. Thus, in the Commission's view, the power to issue instructions cannot

[923] BT-Drucks. VI/334, P. 56.
[924] BT-Drucks. VI/334, P. 58.
[925] BT-Drucks. VI/334, P. 59.

be justified solely by the employee's contractual consent. It is not a result of mutual agreement between the contracting parties, but exists independently of this”.[926]

In this context, the Co-determination Commission considered an essential task of co-determination to be the "establishment of a compulsion for cooperation between employer and employee" and specifically the "introduction of compulsory argumentation and discussion", in particular with regard to the exercise of the employer's authority to issue instructions resulting from the employment contract. In this respect, the Commission considered that employee co-determination was "also necessary because it is only in this way that the particular expertise and perspective specific to the position and activity of employees in the enterprise can be reliably incorporated into the decision-making processes in the enterprise".[927] At the same time, it also considered co-determination to be "a suitable and necessary means of supplementing the protective function of labour contract law and occupational health and safety law in the internal company sphere".[928]

a) Membership concept

There is no doubt that the idea of the employee's "membership" in the "social association" of the company, as put forward by the Co-Determination Commission, can be continued under the conditions of digitalisation. It cannot be overlooked that in the course of digitalisation, company boundaries are often dissolving and the "social association" is becoming looser to the same extent as the company boundaries are becoming blurred. But no one will claim that the bond of membership has been cut. It may also be that the interests of workers will clash less if, for example, as a result of the increasing use of the home office, the number of personal encounters decreases. However, it will hardly be possible to deny that new conflicts (both within the workforce and vis-à-vis the employer) will immediately take the place of old ones; just consider that in the context of home office the question must be answered by whom, under what conditions and to what extent it should be allowed. However, these conflicts must then also be resolved.

[926] BT-Drucks. VI/334, P. 61.
[927] BT-Drucks. VI/334, P. 67.
[928] BT-Drucks. VI/334, P. 67.

The "membership idea" is thus still fully viable as a justification for the existence of a works constitution.[929]

b) Human dignity and personality rights

If we now turn to the protection of human dignity and respect for the right of personality as a "central value decision" for co-determination in companies, we must first emphasise that this value decision is "timeless" and thus undoubtedly continues to be valid. In addition, however, the protection of human dignity and the protection of personality are facing completely new challenges due to AI and Big Data. Although human dignity and the right to privacy should always be considered in tandem and therefore many overlaps can be observed, this will be examined separately for both areas in the following.

aa) Protection of human dignity

The guarantee of human dignity is a defining feature of the Basic Law, and even more, "the 'state idea' of the Federal Republic of Germany as a legal community finds its normative expression most readily in Article 1 of the Basic Law and there above all in the guarantee of human dignity in Paragraph 1".[930] Accordingly, human dignity is superior to other concerns of constitutional rank, a primacy that is reflected not least in the resistance to weighting that human dignity shows in conflict with other legal interests, as postulated by the majority of scholars.[931] This is confirmed by the inclusion of the guarantee of human dignity in the "eternity guarantee" of Article 79(3) of the Basic Law. It follows from all this that respect for human dignity has lost none of its importance even fifty years after the work of the Co-Determination Commission. Accordingly, nothing has changed with regard to the insight that the employee's entry into a "company community" must be accompanied by the opportunity to have a say. In all of this, it must also be taken into account that Article 1(1) not only imposes an obligation on the state to respect human dignity (para. 1, second sentence, first alternative) that takes into account

[929] It should only be noted in passing that the increasing dissolution of the boundaries of the workplace further underlines the need for collective representation of interests, since a certain "isolation", which is almost inevitable in the course of the increasing spread of work that is not tied to a company, alone leaves no other option for the effective representation of workers' interests than the option of bundling them in a collective interest representation body.
[930] Thus Maunz/Dürig/Herdegen, Art. 1 GG marginal no. 4.
[931] Also Maunz/Dürig/Herdegen, Art. 1 GG marginal no. 4.

the defensive dimension of the guarantee of human dignity,[932] but also an obligation to protect (para. 1, second sentence, second alternative), which is aimed at "securing the conditions of a dignified existence and taking precautions against violations of dignity by private individuals".[933]

It is obvious that in view of the possibilities that AI and Big Data are already opening up today, human dignity is, if not at stake, then at least exposed to massive challenges. Particularly illustrative in this respect is the subjection of humans to the decision-making power of machines, which prompted the European legislature to lay down the regulation in Article 22(1) of the GDPR. The connection between this provision and the protection of human dignity is shown, as already mentioned above, in the fact that the literature often formulates the purpose of Article 22(1) of the GDPR in a way that is strongly reminiscent of the so-called "object formula" developed by the Federal Constitutional Court for Article 1(1) of the Basic Law.[934] However, human dignity also appears to be affected on this side of automatic decisions, if one considers the pronounced "manipulative abilities" of many AI systems. It is not without reason that critics of nudging also point out that it is not in line with human dignity.[935]

However, human dignity is also affected in other dimensions. Three aspects of the guarantee of human dignity are particularly striking: protection against the "exploration of involuntary processes",[936] the protection of the spatial-objective private sphere and protection against the exploration of personality traits.

As concerns the first area of protection, there is agreement that the state is prevented from exploiting a deficit in the control of the will over bodily processes by means of state power. This means, for example, that the use of polygraphs is prohibited, as these are used for the direct registration of physiological processes and allow corresponding evaluations.[937] In general, the use of a polygraph to measure and record bodily processes that are beyond voluntary control, such as blood pressure, pulse or respiratory rate, during interrogations in criminal proceedings violates human dignity (as well as the provisions of ordinary law in

[932] Cf. only Maunz/Dürig/Herdegen, Art. 1 GG marginal no. 75.

[933] Thus Maunz/Dürig/Herdegen, Art. 1 GG marginal no. 78.

[934] Cf. only *Atzert* in: Schwartmann/Jaspers/Thüsing/Kugelmann, DS-GVO/BDSG, 2nd edition, 2020, Art. 22 marginal no. 2 with further references.

[935] See only *McCrudden*, Nudging and human dignity, 06 January 2015: https://verfassungsblog.de/nudging-human-dignity-2/.

[936] Cf. Maunz/Dürig/Herdegen, Art. 1 GG marginal no. 85.

[937] Cf. Maunz/Dürig/Herdegen, Art. 1 GG marginal no. 85

Sections 136a and 69(3) of the Code of Criminal Procedure) and largely leads to a completely unsuitable means of evidence.[938] No other principle, then, can apply to procedures based on AI and Big Data, insofar as these also amount to an "exploration of involuntary processes".

As far as the protection of the spatial-objective private sphere is concerned, it has not yet been possible to define a core area of private life that would set insurmountable limits to state access.[939] However, there is every reason to consider human dignity to be affected in this respect, if, for example, evaluations based on the interaction of AI and sensor technology follow employees "right into the farthest corner of the company".

With regard to protection against the exploration of personality traits, however, two things should be noted: first, implications for human dignity arise in particular when it is a question of revealing "deeper layers of personality traits",[940] and second, that this results in barriers, in particular for predictive tests.[941] In this respect, however, with regard to AI and Big Data, it should be remembered once again that AI is increasingly being used to reveal (inner) characteristics of people. It should also be remembered that AI is often about nothing less than predicting people's behaviour as reliably as possible. Predictive policing is probably the best example in this respect.[942]

Finally, human dignity also appears to be affected by AI and Big Data from the point of view of the recognition of a "social claim to respect".[943] It is worth recalling, for example, the warnings of experts that in the working world of tomorrow, sufficient "interaction of human employees with other employees" should be ensured and at least a "core network of human employees" should be maintained at every decision-making level in order to "create empathy" and "improve the preservation of human autonomy".[944] However, warnings that people could lose their social and other competences in the long run when working alongside

[938] Cf. only Werner, in: Weber, Rechtswörterbuch, 27th edition 2021, Lügendetektor; cf. from the case law BGH, NJW 1999, 657; BGH, NJW 2003, 2527.

[939] Cf. only Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 90.

[940] Cf. Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 93.

[941] See also Maunz/Dürig/Herdegen, Art. 2 GG marginal no.94, also with reference to a "right not to know".

[942] Cf. on this e.g. *McDaniel*/*Pease*, Predictive Policing and Artificial Intelligence, 2021, p. 290, who however raise a number of doubts about predictive policing and bemoan its considerable susceptibility to error.

[943] Cf. Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 1. 117.

[944] IEEE, Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System, 1st ed., 2019, p. 102.

machines should also be taken seriously. This would also be extremely problematic from the point of view of protecting human dignity.

In summary, it can be stated at this point that the protection of human dignity is facing serious challenges under the conditions of AI and Big Data,[945] which make it appear urgent to also take precautions for sufficient protection in the form of collective representation of interests and co-determination.

bb) Right of personality

However, nothing has changed with regard to the importance of the right of personality, which the Co-Determination Commission also refers to.[946] This follows from the fact that the right to free development of the personality is "inextricably linked" to the protection of human dignity,[947] so that nothing else can apply to the latter from the outset. The question that arises here is therefore not whether the right of personality under Article 2(1) of the Basic Law can continue to claim unrestricted validity, which is undoubtedly the case, but rather whether recent technological developments are accompanied by special challenges for the protection of personality.

To answer this question, it seems advisable to briefly recall the content of this right, The general right of personality, which the courts have developed from Article 1(1) and Article 2(1) of the Basic Law,[948] “supplements, as an "unnamed" right of freedom, the special ("named") rights of freedom, which, such as freedom of conscience or freedom of opinion, also protect constituent elements of personality. Its task is to guarantee, in the sense of the supreme constitutional principle of "human dignity" (Article 1(1) Basic Law), the narrower personal sphere of life and the preservation of its basic conditions, which cannot be conclusively covered by

[945] It should only be noted in passing that questions also arise from the point of view of whether and to what extent machines (should) participate in the human dignity guarantee; cf. on this *Geminn,* DÖV 2020, 172; cf. on this, however, also IEEE, Ethically Aligned Design - A Vision for Prioritising Human Well-being with Autonomous and Intelligent System, 1st ed., 2019, p. 19 f., according to which "autonomous and intelligent systems (should) not be granted rights and privileges that correspond to human rights".

[946] In this context, it must also be noted for the right of personality that this is not only a classic right of defence, but also leads to corresponding duties of protection on the part of the state; cf. only Maunz/Dürig/Herdegen, Art. 2 GG marginal nos. 132 ff.

[947] Or, in other words: "In the value system of fundamental rights, Article 2(1) makes it indisputable what the dignity of the human being (Article 1(1)) primarily consists of in terms of content (material): - in the "free development of his or her personality""; Preliminary Edition, Article 2(1).

[948] For more details on the dogmatic foundations, see Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 127 ff.

the traditional concrete guarantees of freedom".[949] From the beginning, it was recognised in case law that the need for protection of the personal sphere of life arises "above all in view of modern developments and the new threats to the human personality associated with them".[950] In terms of ordinary law, it is particularly noteworthy that according to Section 75(2), first sentence BetrVG, the employer and the works council must protect and promote the free development of the personality of the workers employed in the enterprise. The provision is an expression of the state's obligation to protect the holders of fundamental rights from a disproportionate restriction of their fundamental rights by private autonomous regulations.[951]

The general right of personality cannot be laid out in detail here. In relation to the relationship of the individual to the state, it includes, among other things, the guarantee of objectively and spatially defined areas of refuge as well as protection against state spying in the narrower sphere of life. State interference in these areas is not inadmissible per se, but requires justification, the (graduated) standards for which are derived from the so-called sphere theory developed by the BVerfG.[952] As far as the first-mentioned cases are concerned, the "spaces of refuge" to be guaranteed according to this theory also include an area "in which the individual is left to his or her own devices unobserved"[953] or can claim to be simply "left alone".[954] This is not limited to the purely domestic sphere, but also includes other locations in which the person concerned can recognisably assume that he or she is not subject to observation by third parties.[955] As far as the latter cases are concerned, that is, in particular the "narrower sphere of life", the guarantee also aims in particular at the "protection of the psychological inner sphere", so that the individual is protected from state "exploration of the world of thoughts and feelings".[956] In this context, too, the literature often refers to the use of lie detectors, in which "the person giving evidence […] is quasi X-rayed by technology and mechanically […] relativised with his statement as his own personality representation" and "in this respect (becomes) a mere appendage of an

[949] Thus BVerfG, NJW 1980, 2070 (and II. 2a); cf. also BAG, NZA 2014, 551 (and para. 40).
[950] Thus BVerfG, NJW 1980, 2070 (and II. 2a); cf. also *Holthausen*, RdA 2021, 19 (27).
[951] In this respect cf. in particular BAG, AP BetrVG 1972 § 87 Überwachung No. 41 (and B.I.2a)).
[952] For details, see Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 157 ff.
[953] BVerfG, NJW 1995, 1015 (and B I. 3.).
[954] BVerfG, NJW 1969, 1707 (and C. II. 1b)); cf. on the whole also Richardi/Maschmann, 17th ed. 2022, § 75 BetrVG marginal no. 50.
[955] Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 149.
[956] Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 154.

apparatus".[957] A special manifestation of the general right of personality is the right to informational self-determination, which includes the right to decide for oneself on the disclosure and use of personal data.[958] The recognition of this right in jurisprudence is also a "manifestation of a protection of personality that adapts to modern developments",[959] whereby it is recognised that there are increasing threats to this right by private individuals[960] and thus the problem of recognising corresponding state obligations to protect against encroachments by private individuals is moving into the centre of attention.[961] It is true that the right of personality under civil law – which is to be observed by the employer – is not identical with the fundamental right derived from Article 2(1) in conjunction with Article 1(1) of the Basic Law. Rather, there are differences in the determination of the scope, in the weighing of conflicting interests and in the concretisation in individual cases. Nevertheless, the solutions should hardly differ from each other in the result, since an interpretation of the right of personality under civil law in contradiction to the dogmatics of fundamental rights is ruled out.[962]

The fact that the general right to privacy faces particular challenges in all of the aforementioned aspects of protection due to recent technological developments should hardly require further justification. In this respect, it is also worth recalling the examples of AI applications listed at the beginning of this study, which, in particular due to the interplay of sensor technology, Big Data and AI, result in possibilities for control and monitoring that extend into the realm of the "world of thoughts and feelings". Even without having to go into this again in detail here, it can hardly be denied that the possibilities opened up by the state of the art have considerably increased the risks of a violation of the general right of personality (and a disregard for human dignity).

The fact that the "vulnerability of personality law" has become considerably more acute, especially due to recent technical developments, is also made clear by a recent decision of the Federal Constitutional Court (BVerfG) that dealt with data

[957] Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 155 with the addition that "due to the de facto self-exculpation compulsion [...] it is also not possible to effectively consent to the use of a polygraph and thus waive the exercise of the fundamental right".

[958] For more details, see for example Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 173 ff.; cf. also Richardi/Maschmann, 17th ed. 2022, Sec. 75 BetrVG marginal no. 60 ff.

[959] Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 173: "The individual, who regularly does not even know what is being collected about him or her where, where he or she leaves electronic traces, can become a mere object of state agencies or economic marketing strategists in the case of systematic data collection".

[960] Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 173, who speaks of the fact that in this respect "the perspective (has) shifted in a specific way".

[961] More detailed Maunz/Dürig/Herdegen, Art. 2 GG marginal no. 189 ff.

[962] Cf. only MünchArbR/Reichold, 4th ed. 2018, section 94 marginal no. 2.

mining, or, in the words of the court, the use of procedures and methods "with the help of which already existing large databases, mostly based on statistical-mathematical procedures, are independently analysed for connections in order to generate 'new knowledge'".[963] In this respect, the BVerfG assumes that it has "in principle an increased burdening effect" if "data from different […] sources stored in a file are used by way of linking to generate new knowledge and correlations",[964] especially since these could have "a considerable relevance to personality".[965] In this context, the weight of intervention is increased even further if "these new findings can be used directly for operational purposes".[966]

cc) Summary

When summarising the considerations at this point, two things emerge: that the considerations of the Co-Determination Commission on the significance of human dignity and personal rights still deserve full approval, but that at the same time they have lost none of their topicality, to say the least, in view of the technical possibilities opened up today. Placing co-determination in the context of the protection of human dignity and respect for the personal rights of workers seems more important today than ever. Indeed, the self-determination of the employee, which is based on this, demands that limits be set to a "power of management and organisation", in the exercise of which employers today have very different and, above all, much more intensive instruments at their disposal than was the case in the past.

c) Collective representation of interests

Lastly, an important change of direction in workplace co-determination is the bundling of the representation of employee interests in a unified interest representation body, the works council. In the view of the co-determination commission, this "gives legal expression to the social reality of the enterprise as a

[963] BVerfG, NVwZ 2021, 226 (and para. 74) and reference to BT-Drs. 17/11582, p. 3.
[964] BVerfG, NVwZ 2021, 226 (and para. 109).
[965] BVerfG, NVwZ 2021, 226 (and para. 110).
[966] BVerfG, NVwZ 2021, 226 (and para. 111); generally, cf. also *Golla*, NJW 2021, 667 (668 f.) with further references and the comment that the intensity of the encroachment "also increases due to the possibility of creating complex links between the data". The processing of personal data "requires justification primarily because information can be linked in a complex manner with modern technical aids and combined to form personality profiles" and "through the software-based processing and linking possibilities [...] even previously possibly irrelevant information (gains) new content".

common, purpose-oriented organisation and as a common condition of all the employees concerned, which the individual agreement is not able to give it".[967]

However, one should not stop there. If it is true that often not even the employer is in a position to understand the "decision-making processes" of the AI it uses, then this must apply all the more to the employee who is exposed to it. For the latter, it must in any case remain as stated above, namely that "the technical specifications and design elements (are) too complex for the individual to be able to penetrate the exact functioning of a software system".[968] This "excessive demand" on the individual in the face of AI has also led to calls for a re-regulation of individual areas of regulation, which were not conclusively assessed above, but which seem highly plausible in view of the findings on the individual regulatory issues. At this point, I would like to call to mind two findings in connection with central regulatory areas in the present context, namely data protection law and discrimination law: The first finding is that in view of the weakness of protection based on the collection and processing of individual personal data alone, a new conception seems advisable that "relies more strongly on procedural requirements for the general handling of data files and data processing".[969] The second is that the current anti-discrimination law also suggests a new approach in view of the existing need for evidence for the data subject alone, which takes measures such as algorithm impact assessments, auditing and certification more clearly into consideration. However, both findings allow only one conclusion for the problem at issue here, namely that the weak position in which individuals find themselves vis-à-vis AI suggests that adequate collective rights should be granted.[970] This is all the more true because only in this way can preventive protection be guaranteed, in which workers can help shape things instead of first having to endure interventions and then being forced into the role of claimants.[971] If one adds to all this the fact that a wide range of reform proposals aim to strengthen collective legal protection against AI, then the value of employee representation by the works council becomes fully clear.

[967] BT-Drucks. VI/334, P. 59.

[968] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 269.

[969] *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 264.

[970] For the area of the GDPR, the importance of collective protection is underlined by *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019, p. 274.

[971] The above considerations are reflected - as far as the European level is concerned - in particular in demands from the literature, which aim to meet the challenges posed by AI by strengthening trade unions or increasing co-determination - although with regard to the latter, mainly co-determination at company level is considered; cf. only *Spencer/Cole/Joyce/Whittaker/Stuart*, Digital Automation and the Future of Work, European Parliamentary Research Service, 2021, p. 53 f. with further references.

d) Acceptance and expertise

Finally, it should also be borne in mind that the acceptance of a technology that is associated with many risks can only be hoped for if collective participation is guaranteed, which is why "participation on an equal footing" is ultimately also in the well-understood interest of the employer side itself. In addition, it is important to bear in mind two ideas of the Co-Determination Commission, namely that co-determination ensures that "the special expertise and approach specific to the position and activity of employees in the company are reliably incorporated into the decision-making processes in the company"[972] and that an essential advantage of co-determination is the "introduction of argumentation and discussion constraints". If one considers, for example, that the selection of representative test data is of crucial importance for ensuring "non-discriminatory" AI, then there is every reason to involve employee representatives in this task. And if it is true that "mixed" development teams are more likely to be able to guarantee "discrimination-free" applications than teams composed exclusively of AI experts and if, accordingly, the recommendation seems to be made in almost every case to form precisely these "mixed" teams,[973] then, with regard to the problem at hand here, this also allows only one conclusion, namely that in this respect the employee side must also be involved via its interest representation body. The "argumentation and discussion constraints" that are installed in the process can only improve the result.

Even more significant, however, seems to be the problem already mentioned above that effective control of AI presupposes two things: an expertise in computer science and a knowledge of the actual circumstances which, for example, "underlie the selection and application of algorithms, models and criteria of differentiation".[974] While the works council will hardly ever have the former knowledge itself, which is why it must be ensured that it has access to sufficient expertise,[975] it is virtually predestined to participate in the latter. In this context, it is also interesting to note that the discussion about "ethical AI" is predominantly conducted on an abstract level and that corresponding demands are thus often without concrete reference

[972] BT-Drucks. VI/334, P. 67.

[973] See only *Byrum*, Build a Diverse Team to Solve the AI Riddle, May 18, 2020: https://sloanreview.mit.edu/article/build-a-diverse-team-to-solve-the-ai-riddle/.

[974] Thus *Orwat*, Discrimination Risks through the Use of Algorithms, 2020, p. 127.

[975] *Thieltges,* ZfP 2020, 3 (29), gives an impressive account of the difficulties works councils face in providing meaningful support for AI projects in the workplace.

to the context in which AI is used.[976] In contrast, it would be desirable for the discussion to be as closely intertwined as possible with the reality of life.

2. New challenges

The considerations made so far are all characterised by the fact that they remain within the framework of the purposes "traditionally" assigned to co-determination. However, recent technological developments give cause to consider whether employee participation might not also be necessary for other reasons. In particular, employers are delegating more and more tasks to adaptive systems. These systems may not decide – or be allowed to decide – for themselves. However, there are indications that they are playing an increasingly important role in decision-making processes on the employer side. Accordingly, as in other contexts, the works council must be involved in order to ensure that the employer's actions comply with the law, amd in particular to ensure that the prohibition of automated decisions on the part of the employer contained in Article 22(1) of the GDPR, which is fundamental in the present context, is observed. This safeguarding appears all the more urgent as difficult questions of delimitation may arise in the context of Article 22(1) GDPR. Since the provision is only directed at "exclusively automated decisions", questions of doubt will often arise in practice if human involvement is "only marginal". It may then be necessary to clarify, for example, whether the human assisting the machine has the necessary data basis, has sufficient professional qualifications and also has the leeway to deviate from the automated decision.[977] Quite independently of this, however, it also seems plausible that people could be increasingly inclined to recognise the "superior knowledge" of the machine and to follow its "decision proposals".[978] In this respect,

[976] Cf. *Birhanem*/*Ruane*/*Laurent*/*Brown*/*Flowers*/*Ventresque*/*Dancy*, The Forgotten Margins of AI Ethics, 13. https://arxiv.org/pdf/2205.04221.pdf: “Even though AI Ethics is a fast growing and broadly construed field of enquiry, its pace is no match for the rate at which algorithmic systems are being developed and integrated into every possible corner of society. Thus, the field holds a crucial place in ensuring that algorithmic systems are just and equitable; in bringing to light algorithmic failures, whom they fail; as well as in holding responsible bodies accountable. If AI Ethics is to protect the welfare and well-being of the most negatively impacted stakeholders, it needs to be guided by the lived experiences and perspective of such stakeholders. The field also needs to treat AI Ethics as non-divorcible from day-to-day life and something that can’t emerge in a historical, social, cultural, and contextual vacuum.”

[977] Cf. *Scholz* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht (ed.), 1st ed. 2019, Art. 22 marginal no. 27.

[978] Experiments show that people often blindly trust robots even when they have every reason not to; cf. only *Robinette/Li/Allen/ Howard/Wagner*, Overtrust of robots in emergency evacuation scenarios, in: Bartneck (ed.), The Eleventh ACMIEEE International Conference on Human Robot Interaction, Piscataway, NJ, 2016, pp. 101-108; cf. *Hardré*, in: Tettegah/Espelage (ed.), Emotions, Technology, and Behaviors – A Volume in Emotions and Technology, 2016, p. 85. They also seem to be relatively easily led into risky behaviour; cf. *Hanoch/Arvizzigno/Hernandez García/Denham/Belpaeme/Gummerum*, The Robot Made Me Do It: Human-Robot Interaction and Risk-Taking Behavior, Cyberpsychology, Behavior, and Social Networking 2021: doi: 10.1089/cyber.2020.0148; cf. on the whole also *Martini*, Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2020, p. 87. See also *Brink*/*Schwartmann*/*Oetjen*/*Voss*, FAZ 18 July 2022 p. 18.

too, additional safeguards in the form of the participation of the collective interest representation body are indicated.

Above all these considerations, however, one should not lose sight of the fundamental problem. This lies in the fact that the use of AI, at least in the long term, carries the danger of a "depersonalisation" of the employment relationship[979] and also of disenfranchisement, not only on the part of the employees ("under-demanding" and "subordinating" of the human to the machine, etc, which addresses the protection by Article 1 and Article 2 Basic Law), but also, which is no less important in the present context, on the part of the employers. If – possibly as a result of a creeping process – "the ethics of action and responsibility shifts from humans to artificial intelligence – or initially to the humans in the environment of the AI, whose programming performance overlaps the actions of the user",[980] then protection under co-determination law must also be activated in this respect. It is true that the current law (Section 613 second sentence BGB) only contains the assessment that another employer may not be imposed on an employee against his or her will. However, one must also bear in mind the risk that the employer, through the extensive use of AI, will delegate a wide range of decision-making powers, and thus ultimately also parts of its responsibility towards the employees, to machines. In this respect, too, the involvement of the works council is urgently required to ensure that this limit is not exceeded.

The fact that the scenario just mentioned is a future scenario, but as such is by no means conjured from thin air, becomes clear when one considers, for example, the current debate under company law about the significance of AI. In the relevant literature, various development stages of the collaboration of humans and machines are described, ranging from assisted intelligence (AI assists humans), augmented intelligence (AI assists in a way that a human could not), amplified intelligence (the machine helps decide), autonomous intelligence (the machine decides independently according to human guidelines) to autopoietic intelligence (the machine is able to develop and expand the area of its decision-making responsibility independently).[981] To say the least, it cannot be ruled out that the machine is in the process of gradually climbing these levels – which is why, for

[979] See also *Allen/Master*, Technology Managing People - the legal implications, 2021, p. 78: "The increased reliance on technology to make management decisions risks profoundly undermining the personal nature of the employment relationship. Humans have the potential to provide empathetic and nuanced responses within decision-making, which is currently beyond AI-powered tools".

[980] Cf. *Schliesky*, NJW 2019, 3692 (3696) on the example of assistance systems.

[981] See *Hilb*, Toward Artificial Governance? The role of artificial intelligence in shaping the future of corporate governance, Journal of Management and Governance, 2020, 851 (861).

example, it is not at all far-fetched to consider the possibility of so-called self-driving corporations, in which corporate decisions are made by AI systems, as is already happening in the corporate law literature.[982]

3. Co-determination and platform work

Remarkably, the European Commission has also explicitly recognised the importance of collective representation in its proposal for a directive to improve working conditions in platform work, albeit (initially) only for this area.[983] According to Article 9(1) of the draft, Member States shall “without prejudice to the rights and obligations under Directive 2002/14/EC […] ensure information and consultation of platform workers’ representatives or, where there are no such representatives, of the platform workers concerned by digital labour platforms, on decisions likely to lead to the introduction of or substantial changes in the use of automated monitoring and decision-making systems referred to in Article 6(1)[984]”.[985] In so providing, the Regulation explicitly serves the goal of "promoting social dialogue on algorithmic management systems".[986] The regulation underlines the importance of worker participation in connection with AI. The fact that it would have only limited significance for German law is a different matter.[987]

[982] See *Armour/Eidenmüller*, Self-driving Corporations, European Corporate Governance Institute - Law Working Paper No. 475/2019. In other contexts, too, there are reminders to address possible regulatory issues in good time; for example, for the evolutionary development of robots, most recently *Eiben/Ellers/Meynen/Nyholm*, Robot Evolution - Ethical Concerns, frontiers in Robotics and AI, 03 November201: doi: 10.3389/frobt.2021.744590.

[983] Proposal for a Directive of the European Parliament and of the Council on Improving Working Conditions in Platform Work of 09.12.2021, COM(2021) 762 final.

[984] According to this, automated monitoring systems are systems "used to electronically control, monitor or evaluate the work performance of platform workers", whereas automated decision-making systems are used "to make or support decisions that significantly affect the working conditions of those platform workers, in particular their access to work assignments, their earnings, their occupational safety and health, their working time, their promotion and their contractual status, including the restriction, suspension or termination of their account".

[985] See also Recital 39, noting that "the introduction of automated monitoring and decision-making systems through digital work platforms or significant changes in the use of these systems [...] directly affect(s) the organisation of work and the individual working conditions of platform workers".

[986] COM(2021) 762 final, p. 4 and 20. It is also worth noting in this respect that the GDPR grants individual rights to data subjects but "does not take into account the important collective aspects of labour law, including the role of workers' representatives, information and consultation of workers and the role of labour inspectorates in the enforcement of workers' rights"; COM(2021) 762 final, p. 8.

[987] Cf. on this *Krause*, NZA 2022, 521 (530).

4. Individual regulations

In the following, we will take a closer look at individual regulations that are of particular relevance in the present context.[988]

a) Right of co-determination for technical equipment

Of central importance for the use of AI systems is obviously the right of co-determination according to Section 87(1) No. 6 BetrVG,[989] which is directed at "the introduction and use of technical equipment intended to monitor the behaviour or performance of employees". If one considers in particular the ability of these systems to handle enormous amounts of data (Big Data) and the possibility thus opened up to derive statements for the (future) behaviour of employees, then one cannot avoid the realisation that ensuring sufficient co-determination with regard to technical equipment is more important today than ever before.[990] As far as the scope of application of Section 87(1) No. 6 BetrVG is concerned, it should be noted that in answering the question of whether a technical device is intended to monitor the behaviour or performance of employees, the BAG focuses solely on objective suitability. It is therefore not necessary that the employer pursues the purpose of monitoring by introducing and using a technical device. Rather, it is sufficient if the technical device is objectively suitable for monitoring according to its design or a program connected with it.[991] In this context, Section 87(1) No. 6 BetrVG establishes, according to widespread opinion, a right of initiative. However, according to the case law, the works council should not be able to demand the introduction of a technical device, since this can only be a regulation to counteract dangers for employees.[992]

aa) Prerequisites of the right of co-determination

Regarding the prerequisites of the right of co-determination, it should be noted that the interpretation of the provision by the Federal Labour Court (BAG) cannot be

[988] The right of co-determination under section 87(1) no. 7 BetrVG will not be discussed in detail here; cf. only *Gäbert*, AuR 2021, 9. The rules on safeguarding employment under Sections 111 f BetrVG will also be excluded here, as they are not "CI-specific"; cf. however *Klebe*, SR 2019, 128 (131 f.).

[989] See also *Schwarze*, in: *Ebers/Heinze/Krügel/Steinrötter* (eds.), Künstliche Intelligenz und Robotik, 2020, § 8 Arbeitsrechtliche Probleme von KI und Robotik p. 270 (300 ff.).

[990] Illustrative of this is *Klebe*, NZA-Beil. 2017, 77 (82).

[991] Cf. in this respect only BAG, NZA 2017, 657 (para. 22 below): "Technical devices are 'intended' for monitoring if they are objectively suitable for collecting and recording behavioural or performance information about the employee; the employer's subjective intention to monitor is irrelevant [...]"; cf. also BAG, NZA 2019, 1009 (para. 24 below).

[992] BAG, NZA 1990, 407; approvingly *Richardi/Maschmann*, in: Richardi, 17th ed. 2022, § 87 BetrVG marginal no. 530.

ignored. It is true that concerns from company practice are to be taken seriously, which are based, for example, on the fact that every software update triggers a right of co-determination of the works council.[993] However, based on the case law, this does not change the fact that the employer must involve the works council and observe the requirement for consent.[994] Nor does the provision contain any reference to a materiality threshold,[995] nor is there any basis for granting the employer a right to provisionally introduce a technical device.[996] It is true that delays that may occur in the exercise of co-determination rights may be unacceptable from the employer's point of view. But this does not mean that co-determination rules can simply be set aside. And even if it may be particularly painful in times of increasing digitalisation of companies if there is a loss of time in this area in particular, two things would have to be demanded of those who assert the necessity of restrictions on co-determination in this respect: on the one hand, that they specifically state why co-determination no longer appears to be acceptable here, and on the other hand, that they at least provide indications as to how the cases are to be delimited in which the works council's co-determination rights are to be set aside. As far as can be seen, neither the one nor the other has been achieved so far.[997] Irrespective of this, it should be pointed out that the problem is crying out for regulations in the form of (framework) agreements between the company partners; a path that has indeed[998] been addressed by the Enquete Commission, for example, but also in the literature[999] and which some authors occasionally urge should be followed.[1000] For example, the final report of the Enquete Commission states: "In order to speed up the approval process, it would be possible, for example, for the employer and the works council to conclude a principle-based framework agreement and application-specific individual agreements. One advantage of this solution would be to reduce the effort involved in concluding the numerous individual agreements. The framework agreement

[993] Cf. again *Henssler*, NZA Supplement 2020, 3 (7); cf. also *Haußmann/Thieme*, NZA 2019, 1612 (1617) with proposed amendments *de lege* ferenda; extremely crit. most recently also *Krülls,* RdA 2021, 279 with a comparison with Austrian law.

[994] Also *Klebe,* NZA supplement 2017, 77 (82); in general *Richardi*, in: Richardi, 17th ed. 2022, § 87 BetrVG marginal no. 55.

[995] See also *Schreiner*, DB 2019, 554.

[996] Cf. Klebe, NZA Supplement 2017, 77 (and footnote 51).

[997] Another question is that it may be doubtful to what extent e.g. people analytics applications trigger a right of co-determination under section 87 (1) No. 6 BetrVG. The level of aggregation of the processing is of decisive importance, but also whether and to what extent the collected data lead to assessments which - at least potentially - result in individual personnel measures. On the other hand, it is probably irrelevant if the application is aimed at making predictions, as this objective does not exclude the existence of monitoring; for more details on the whole, see *Götz*, Big Data im Personalmanagement - Datenschutzrecht und betriebliche Mitbestimmung, 2020, p. 190 et seq.

[998] Report of the Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale, BT-Drucks. 19/23700, p. 321.

[999] Cf. most recently *Ludwig/Hinze*, NZA 2021, 1444 (1445).

[1000] Also on this point *Henssler*, NZA Supplement 2020, 3 (7) should regulate points that come up as questions during the introduction of all applications".[1001]

According to what has just been said, attempts at a restrictive interpretation of Section 87(1) No. 6 BetrVG must be opposed. At the same time, however, it should be pointed out that this co-determination provision is likely to become even more important in the future as a result of technical developments themselves. If in the past it was understandable to disregard Section 87(1) No. 6 BetrVG when processing mere status data (such as name, address, marital status, professional qualifications and curriculum vitae), because these were not directly linked to specific conduct, this is likely to be correct only to a limited extent given the current state of the art, which allows the merging and evaluation of data from very different sources. Even if one disregards the fact that the evaluation regularly takes place in a black box (and thus the "basis for decision-making" is at best only partially traceable), there is no denying the fact that status data are not only capable of analysis as such, but are also quite meaningful, and even more so when they are combined with other data.[1002] Against this background, there is much to be said for also applying Section 87(1) No. 6 BetrVG to status data stored for the purpose of processing contracts.[1003]

bb) Right of initiative

As to a possible right of the works council to introduce a technical device, there is no getting around the fact that the BAG expressly rejected such a right of initiative in connection with Section 87(1) No. 6 BetrVG some time ago,[1004] although, according to the court's established case law, the works council's co-determination rights under Section 87(1) BetrVG in principle also include a right of initiative of the works council.[1005] As justification, the court referred to the purpose of Section 87(1) No. 6 BetrVG: that "the dangers of an infringement of the right of personality and the right of employees to the free development of their personality that may

[1001] Report of the Enquete-Kommission, BT-Drucks. 19/23700, p. 321. In addition, the Commission calls for "the principles and contents of traditional company agreements to be further developed or rethought on the basis of Section 87(1) No. 6 BetrVG". And further: "It is a matter of strengthening process orientation and making it more agile, and of basing the analysis of effects and evaluation on standards and scientific findings".

[1002] More closely *Klebe/Klengel*, NZA 2021, 1144 (1147).

[1003] This is indeed the case with *Klebe/Klengel*, NZA 2021, 1144.

[1004] BAG, NZA 1990, 407.

[1005] This is also expressly the case in BAG, NZA 1990, 407 (and 2. a)).

emanate from technical monitoring devices […] can be countered by a co-determined regulation on the introduction and more detailed use of such devices". The right of co-determination thus has a "defensive function". However, it contradicts this function "if the works council itself – for whatever reasons – demands the introduction of such a technical monitoring device", and this applies "irrespective of whether the interests of the employees are actually affected by a technical monitoring device, whether the works council sees such an impairment of interests or whether it strives to exclude this by the more detailed design of the co-determined regulation on the use of the technical monitoring device".[1006]

Of course, this ruling must also be respected. However, there are serious reservations about it, which go as far as the legal methodological question of whether the BAG's decision was not in fact a judicial development of the law and at the same time exceeded its limits.[1007] It is also doubtful that the BAG really took all potential interests into account, since it cannot be ruled out from the outset that, in addition to the protection of the employee's personality, there are other interests worthy of protection which outweigh it (and which could in principle be taken into account within the "internal limits of company autonomy").[1008] Although this cannot be weighed up in detail here, one aspect that speaks in favour of the affirmation of a right of initiative of the works council is the fact that it would be surprising if the right of initiative for the use of AI systems rested exclusively with the employer and the works council thus had no legally secured means whatsoever of bringing to bear whatever advantages the use of AI may have for the benefit of the employees.[1009] The LAG Hamm recently affirmed the works council's right of initiative in the introduction of electronic time recording.[1010]

[1006] BAG, NZA 1990, 407 (407 f.).

[1007] Thus LAG Berlin-Brandenburg, BeckRS 2015, 68190 (and 3.3) in its detailed discussion of the BAG's decision.

[1008] Indeed, LAG Berlin-Brandenburg, BeckRS 2015, 68190 (and 3.5) with further references.

[1009] Similarly, *Greiner/Kalle*, RdA 2021, 76 (82), pointing out that the provision (now) also covers devices "whose introduction and use can also be advantageous from the employee's point of view", and that it is not convincing to reduce the role of the works council "generally to reacting to the employer's actions"; critical of the case law of the BAG, however, also *Klebe/Schmidt/Klengel*, in: Gräfl/Lunk/Oetker/Treibinger, (eds.), 100 Jahre Betriebsverfassungsrecht, 2020, p. 303; also Krause, ibid, p. 353.

[1010] LAG Hamm, NZA-RR 2021, 602 with annotations. *Schmidt*. Against this decision an appeal on points of law is pending before the BAG (Case no. 1 ABR 22/21).

b) Design of workplace, workflow and working environment

Section 90 BetrVG provides for information and consultation rights of the works council.[1011] Subsection 1 lists the following as subjects of these rights: the planning of new constructions, conversions and extensions of factory, administrative and other operational premises (No. 1), of technical installations (No. 2), of work procedures and work processes (No. 3) and of workplaces (No. 4).[1012] In this context, "work processes" is understood as the technology used to modify the object of work in order to fulfil the work task, whereas "work sequences" are the temporal and spatial arrangement of the work projects in the operation.[1013] Section 90(2), first sentence, BetrVG concretises the content of the right of participation such that the employer has to consult with the works council on "the envisaged measures and their effects on the employees, in particular on the nature of their work, as well as the resulting demands on the employees in a timely manner so that proposals and concerns of the works council can be taken into account in the planning". This makes it clear that the employer does not simply have to inform the works council of its plans, but that the works council must be involved in the planning process.[1014] According to Section 90(2), second sentence BetrVG, the employer and the works council " are to bear in mind the established findings of ergonomics relating to the tailoring of jobs to meet human requirements".

In this respect the *Betriebsrätemodernisierungsgesetz* (Works Council Modernisation Act) of 14 June 2021 has – at least from the point of view of the legislature[1015] – (only) brought the "clarification" that the "working procedures and work processes" mentioned in No. 3 include the use of AI. Section 90(1) No. 3 BetrVG means that the use of AI triggers co-determination under this provision if AI has an impact on work procedures and work processes, which is likely to be the case on a regular basis.[1016] Apart from this, the use of AI can (still) also be relevant

[1011] See also *Schwarze*, in: Ebers/Heinze/Krügel/Steinrötter (eds.), Künstliche Intelligenz und Robotik, 2020, § 8 Arbeitsrechtliche Probleme von KI und Robotik p. 270 (297 f.).

[1012] § Sec. 90 (1) No. 4 BetrVG concerns the technical organisation of the workplace and its environment and, according to general opinion, fulfils the function of a limited general clause, since already the objects mentioned in Sec. 90 (1) Nos. 1 - 3 BetrVG are subject to the duty of participation from the point of view that they have an influence on the design of the workplace, the work process or the working environment; cf. only BeckOK ArbR/Werner, § 90 BetrVG marginal no. 7; *Annuß*, in: Richardi, 17th ed. 2022, § 90 BetrVG marginal no. 16.

[1013] Cf. only *Annuß*, in: Richardi, 17th ed. 2022, section 90 BetrVG marginal no. 13.

[1014] Also *Annuß*, in: Richardi, 17th ed. 2022, § 90 BetrVG marginal no. 13.

[1015] Cf. the explanatory memorandum, BT-Drs. 19/28899, p. 15, 23.

[1016] Cf. only *Wankel*, in: Däubler/Klebe/Wedde, Betriebsverfassungsgesetz, 18th ed. 2022, § 90 BetrVG marginal no. 15a.

under Section 90(1) No. 2 BetrVG if AI is part of a "technical system":[1017] There seems to be consensus that the planned use of robots as a "technical system within the meaning of Section 90(1) No. 2 BetrVG[1018] triggers the rights under Section 90 BetrVG.[1019] There is nothing to suggest that anything else should apply to other AI systems.

c) Selection guidelines

Section 95 BetrVG addresses guidelines on the selection of personnel for recruitment, transfers, regrouping and dismissals. Pursuant to Section 95(1), first sentence, BetrVG, these require the consent of the works council, although in operations with more than 500 employees the works council may demand that guidelines be drawn up on the professional and personal requirements and social aspects to be taken into account in measures under Section 95(1), first sentence. With this, Section 95 of the Works Constitution Act contains the strongest right of participation in matters pertaining to personnel planning, with Section 95(2), first sentence, even granting the works council a right of initiative in in this respect.[1020]

The Works Council Modernisation Act added a new paragraph 2a to Section 95 of the BetrVG. According to this, subsections (1) and (2) also apply "if artificial intelligence is used in the establishment of the guidelines according to these subsections". The addition is aimed, for example, at the case where an AI application draws up selection guidelines independently or within a framework provided by a third party.[1021] Since the selection policy would be attributable to the employer in both cases if the employer has initiated the use of AI, the new regulation does not change the substantive legal situation. Nevertheless, it is a justified clarification of a question of attribution that is becoming more and more important with the increasing use of AI.

d) Vocational training

[1017] Cf. ErfKomm/Kania, 22nd ed. 2022, § 90 BetrVG marginal no. 4; also *Ludwig/Hinze,* NZA 2021, 1444 (1445); cf. on the whole also Frank/Heine, NZA 2022, 1448 (1449).

[1018] Cf. only BeckOK ArbR/Werner, § 90 BetrVG marginal no. 3.

[1019] Cf. only *Kohte*, NZA 2015, 1417 (1419); *Günther/Böglmüller*, in: Arnold/Günther, Arbeitsrecht 4.0, 1st ed. 2018, ch. 4, marginal no. 104.

[1020] Cf. only *Thüsing*, in: Richardi, 17th ed. 2022, section 95 BetrVG marginal no. 50.

[1021] Cf. the explanatory memorandum, BT-Drs. 19/28899, p. 23.

Pursuant to Section 97(2) BetrVG, the works council has a right of co-determination in relation to in-company training measures.[1022] This concerns situations in which the employer plans or implements measures "which result in the activity of the employees concerned being modified and their professional knowledge and skills no longer being sufficient to fulfil their tasks". In this case, the works council can then have a say in the introduction of in-company vocational training measures. According to some, the regulation – still often overlooked today[1023] – is a suitable basis for qualification initiatives, especially in times of digital upheaval.[1024] This is especially true with regard to the use of AI systems. The central prerequisite for the provision to apply is[1025] that the employees' activities change due to a measure taken by the employer and that their professional knowledge and skills are therefore no longer sufficient to fulfil their tasks. If, on the other hand, the employees have the necessary knowledge and skills and only need to be retrained, the provision does not apply.[1026] It is obvious that the use of AI systems in particular changes the workplaces concerned and thus the activities of the employees to such an extent that mere instruction is no longer sufficient.[1027] It should be noted in all this that co-determination under Section 97(2) BetrVG does not rule out the applicationof Sections 111 et seq. BetrVG,[1028] although the details of the relationship between co-determination in further training on the one hand and a balance of interests and social plan on the other have not yet been conclusively clarified.[1029]

e) Involvement of experts

Pursuant to Section 80(3) BetrVG, the works council may "consult experts in the performance of its duties, subject to further agreement with the employer, to the extent necessary for the proper performance of its duties". It is true that according to the clear wording of the provision, the works council can only call in an expert "only after further agreement with the employer". However, there is agreement that the works council has a right to call in an expert to the extent necessary for the

[1022] See also *Schwarze*, in: Martin Ebers/Christian Heinze/Tina Krügel/Björn Steinrötter (eds.), Künstliche Intelligenz und Robotik, 2020, § 8 Arbeitsrechtliche Probleme von KI und Robotik p. 270 (276). See also *Krause*, NZA 2022, 737.

[1023] *Göpfert/Seier*, NZA 2019, 588 (588) even speak of a "Sleeping Beauty existence" of the provision.

[1024] *Göpfert/Seier*, NZA 2019, 588 (594) attribute to the provision the "potential to become a "central norm"".

[1025] With regard to the prerequisites of the right of co-determination, some things are unclear; cf. only *Thüsing,* in: Richardi, 16th ed. 2018, section 97 BetrVG marginal no. 11 et seq.

[1026] At least according to *Göpfert/Seier,* NZA 2019, 588 (589).

[1027] So also *Göpfert/Seier*, NZA 2019, 588 (589).

[1028] Cf. *Göpfert/Seier*, NZA 2019, 588 (592); also *Röder/Gebert*, NZA 2017, 1289 (1293).

[1029] Likewise *Göpfert/Seier*, NZA 2019, 588 (592).

proper performance of its duties. According to case law, the employer can only be obliged to agree to the involvement of an expert if it is considered necessary in the specific situation in which the works council has to fulfil its tasks.[1030] If the employer refuses to reach an agreement despite the necessity of calling in the expert, the works council can have the employer's consent replaced by a labour court decision.[1031]

The Works Council Modernisation Act added two sentences to para. 3: "If the works council has to assess the introduction or application of artificial intelligence in order to carry out its duties, the involvement of an expert shall be deemed necessary in this respect. The same applies if the employer and the works council agree on a permanent expert in matters pursuant to sentence 2". The new provision is intended to provide works councils with "simplified access to special expertise on related issues so that they can carry out their duties in this respect".[1032] Considering that Section 80(3) BetrVG in its former version left considerable room for disagreement between the parties on the involvement of an expert, it is clear that the amendment of the provision offers a not inconsiderable added value. For if the involvement is to be regarded as necessary by virtue of an irrebuttable legal presumption, which this is,[1033] disputes between employer and works council are ruled out. However, the fact that an agreement between the employer and the works council is still required in these cases is a cause for concern,[1034] as it means that there is still a risk that disputes between the employer and the works council will arise with regard to the person of the expert and/or questions of fees.[1035] Even more worrying, however, is that there is still a risk of disputes over the question of how the topic on which the expert is to give an opinion is to be determined in concrete terms.[1036]

[1030] Cf. only BAG decision of 16 Nov 2005, NZA 2006, 553 (and para. 31) with further references.

[1031] Cf. only BAG decision of 25 Jun 2014, NZA 2015, 629 (and para. 20).

[1032] Cf. the explanatory memorandum, BT-Drs. 19/28899, p. 23.

[1033] Cf. only *Buschmann*, in: Däubler/Klebe/Wedde, Betriebsverfassungsgesetz, 18th ed. 2022, § 80 BetrVG marginal no. 158c.

[1034] This is expressly stated in the explanatory memorandum, BT-Drs. 19/28899, p. 23.

[1035] See also *Frank/Heine*, NZA 2021, 1448 (1449); *Reinartz*, NZA-RR 2021, 457 (467); sceptically also *Schulze*, ArbRAktuell 2021, 211 (213).

[1036] Cf. Richardi/Thüsing, 17th ed. 2022, § 80 BetrvG marginal no. 103, who moreover complains about the lack of definition of the term "artificial intelligence" by the legislator; cf. on this also *Frank/Heine,* NZA 2021, 1448; *Horstmeier*, BB 2022, 116 (120).

f) The regulatory instrument of the works agreement

As in other cases, when it comes to co-determination in the context of AI, the works agreement is the method of choice.[1037] However, the conclusion of works agreements is also an obvious choice in substance. On the one hand, the legal requirements in the form of the GDPR are, as the above analysis has shown, rather thin, very abstract throughout and in need of a great deal of filling out. Consequently, it makes sense from the outset to concretise these through corresponding collective agreements. On the other hand, the technical development is in such a state of flux that one should not wait for full answers from the legislature – especially against the background of an ever-increasing inclination of entrepreneurs to use AI and Big Data. This is all the more true as individual aspects can hardly be considered in isolation. For example, AI, sensor technology, video and audio technology, geotracking, pattern recognition processes and Big Data analytics are increasingly intertwined.[1038] In this respect, it is no coincidence that the GDPR not only refers in Article 88(1) – implemented by the German legislature in Section 26 of the BDSG – to the possibility of "adopting more specific provisions by means of collective agreements to ensure the protection of rights and freedoms with regard to the processing of personal employee data in the employment context", but also explicitly refers to "works agreements" in Recital 155. Works agreements allow for employee data protection tailored to the needs of the company.[1039] According to Article 88(2) GDPR and Section 75(2) BetrVG, this requires a balancing of the employee's personal rights and the employer's legitimate interests, taking into account the circumstances of the individual case,[1040] for which the relevant business partners are virtually predestined.[1041]

However, it is not only – in the abstract – the rapid technical development and the "openness" of the legal requirements that argue for the establishment of a "company employee data protection". Rather, there is also a very concrete need for corresponding regulations by the company partners. This has also already been

[1037] Cf. also *Körner*, NJW 2018, 2825.

[1038] Cf. *Weichert*, NZA 2020, 1597 (1599).

[1039] This is also the case in the draft of a law to adapt data protection law to Regulation (EU) 2016/679 and to implement Directive (EU) 2016/680 (Data Protection Adaptation and Implementation Act EU - DSAnpUG-EU), BT-Drucks. 18/11325, p. 98.

[1040] Cf. in this respect only BAG, NZA 2003, 1193; cf. on the whole also *Holthausen*, RdA 2021, 19 (27 f.).

[1041] What role the collective bargaining parties could play in this context will not be examined in more detail here; but cf. for example *Haußmann/Thieme*, NZA 2019, 1612 (1619) with a discussion of the possibility of "regulating sector-typical system uses in outline"; cf. also *Henssler,* NZA Supplement 2020, 3 (7); BMAS, Ergebnisbericht des Zukunftsdialogs "Neue Arbeit - Neue Sicherheit", 2019, p. 32.

addressed in the consideration of the GDPR: Since the consent of the data subject is of central importance for the assessment of the lawfulness of the processing, but this already reaches its limits in purely practical terms when confronted with the distinctive features of AI, there is a considerable interest in concluding a works agreement.[1042] This interest is also likely to be mutual, because both the works council and the employer must be interested in avoiding uncertainties such as those often associated with the "voluntary nature" of consent.[1043]

However, agreements going beyond data protection, which are indeed encountered again and again in practice, should also serve a mutual interest of employer and employees.[1044] In particular, these can also defuse the dispute over the application of Section 87(1) No. 6 BetrVG, insofar as they provide for a simplified procedure in cases of minor changes to technical equipment, each of which in itself would set the co-determination procedure in motion again.[1045] In fact, a wide variety of agreements can already be found in practice.[1046] For these, the fact that analysis procedures constitute the (intellectual) property of the AI providers and are covered by business secrecy is a limiting factor.[1047] However, this does not change the usefulness of such agreements. For example, the purpose of the AI model or machine learning application used can be specifically regulated in a company agreement; possible sources of error can be identified; regulations can work towards the greatest possible transparency between the developer and user companies; and the effects of the use of AI can be addressed in a company agreement.[1048] In view of the speed of development, it is imperative that the corresponding co-determination processes be as open as possible.[1049]

[1042] Cf. *Holthausen*, RdA 2021, 19 (28) with the assessment that consent "in view of its individual, voluntary and always revocable character is a bulky, unwieldy and possibly even unsuitable authorisation for data processing". and that "effective consent management [...] is a truly Herculean task, especially in the case of mass offences and thus also in the case of Big Data, people analytics and the use of AI, and that it is extremely susceptible to failure with regard to the effectiveness requirements of consent".

[1043] Cf. also *Holthausen*, RdA 2021, 19 (28).

[1044] Cf. also *Holthausen*, RdA 2021, 19 (28), who points out the double advantage that on the one hand, the parties to the works agreement can "make more precise regulations in relation to the occasion and purpose than is possible with a mere recourse to the statutory authorisation criteria of the GDPR or national regulations ", and on the other hand works agreements also serve to exercise the co-determination rights of the works council. The author illustrates the design options using the example of a works agreement on a so-called "pulse survey"; ibid (29f.). *Ludwig/Hinze*, NZA 2021, 1444, promote the conclusion of so-called "digitalisation agreements", whereby they have in mind a *quid pro quo* (framework regulations in the interest of the workforce against a "reduction" of the right of co-determination under section 87 (1) no. 6 BetrVG).

[1045] Cf. *Haußmann/Thieme*, NZA 2019, 1612 (1619).

[1046] See *Thieltges*, Machine Learning Anwendungen in der Betrieblichen Praxis - Praktische Empfehlungen zur betrieblichen Mitbestimmung, 2020 with a comprehensive evaluation of relevant agreements.

[1047] Cf. also *Thieltges,* ZfP 2020, 3 (30).

[1048] Cf. again *Thieltges*, Machine Learning Anwendungen in der Betrieblichen Praxis - Praktische Empfehlungen zur betrieblichen Mitbestimmung, 2020, p. 28 ff.

[1049] Cf. on this in turn the concept paper of the German Trade Union Confederation(DGB) "Künstliche Intelligenz (KI) für Gute Arbeit" of March 2022, p. 14, which calls for considering the establishment of "living agreements"; on thoughts on reform cf. also *Klebe*, AuR 2020, 196.

5. Summary

In summary, it can be said that the idea of co-determination has not only lost none of its importance, but that securing sufficient co-determination in the era of AI and Big Data seems more urgent than ever. The DGB has recently presented a "Draft Bill for a Modern Works Constitution Act", which was prepared by a group of experts.[1050] It contains a number of concrete reform considerations that can form a basis for the urgent discussion on a reform of the Works Constitution Act.

[1050] The draft bill is accessible at: https://www.dgb.de/themen/++co++02729430-b4bf-11ec-9dbe-001a4a160123; its main features are presented in *Klapp/Klebe*, NZA 2022, 689: critical *Annuß*, NZA 2022, 694.

H. Conclusion

1. Caution is required with regard to the very term "artificial intelligence". "Artificial" intelligence is something quite different from human intelligence.[1051] Perhaps the best way to grasp AI applications is to see them as "complementary tools" with their own strengths and weaknesses.

2. Modern AI must always be thought of in conjunction with Big Data. Modern computer technology makes it possible to analyse huge amounts of data in the blink of an eye. AI aims to recognise patterns. However, it should be noted that algorithms "work with correlations and probabilities that do not necessarily follow a causality and do not necessarily lead to results that are 'correct' according to human insight". Accordingly, "erroneous, unfair or discriminatory conclusions can be drawn from the systematisation of accurate individual data [...], which - if they become the basis for decision-making - significantly affect the liberties of the person concerned".

3. Given the error-prone nature of AI, even despite the prohibition of automated decisions in Article 22 GDPR, there is every reason to reserve decisions to humans. Within the scope of application of Article 22 GDPR, effective control of the conditions of this provision must be ensured. From a practical point of view, the risk that humans might be increasingly inclined to let machines make decisions should be addressed.[1052]

4. The use of AI can also bring about a lot of good in working life, as can be seen in the example of occupational health and safety. At the same time, AI poses enormous challenges for many areas of labour law, such as data protection law. Caution is called for in many cases. For example, the initial euphoria about allegedly "non-discriminatory" decisions by machines seems unfounded. On the contrary, given the way AI works, there is every reason to be cautious: uncovering correlations does not mean making "intelligent" decisions.

Initial efforts towards an international regulation of AI within the framework of the Council of Europe are encouraging. The European Union is leading the way globally with the draft of an "AI law" and is, all in all, moving in the right direction. A sensible regulation of AI must focus on prevention. At the same time, however,

[1051] Cf. on this *Walsh*, Was Künstliche Intelligenz wirklich besonders macht, FAZ 7 June 2022, p. 20.
[1052] See also *Schneider*, BC 2022, 225 (231).

sufficient individual legal protection must be ensured, and this should be expanded to include elements of collective legal protection. Co-determination is more important than ever.

5. At the international level, the work of the Council of Europe should be highlighted in this context. The Committee of Ministers in particular has addressed AI from various angles. The Committee‘s statement of February 2019, for instance, includes a reference to the risk of "micro-targeting of individuals based on profiling", but also a warning about the ability of AI systems "not only to predict decisions, but also to influence emotions and thoughts and to change an expected course of action". Its recommendation of 8 April 2020 addresses the "[systematic] aggregation and analysis of data" and laments that "tracking at scale [may] have serious adverse effects on the exercise of human rights". Efforts to regulate AI internationally within the Council of Europe also inspire hope.

6. The European Union is leading the world with the draft of an "AI law" and is, all in all, moving in the right direction. However, the present draft is also subject to a number of concerns: The chosen legal basis raises fears that national regulations to protect workers could be "blocked". The chosen instrument, a regulation, has also raised doubts in this respect. However, there are also concerns about the relationship of the planned regulation to the GDPR, the risk-based approach chosen by the Commission and an approach that relies to a great extent on the development of standards and the self-assessment of providers, but does not involve the potentially affected parties, largely denies them claims and even minimises the role of the user.

7. With regard to Germany, the findings of the Enquete Commission in particular still seem worthy of consideration. This applies, for example, to the use of automated decision-making systems in the personnel sector, but especially to the necessary modernisation of co-determination in the light of AI. Within the framework of the works constitution, the German legislature has also taken the first tentative steps towards regulating AI.

8. The increasing use of AI will necessarily impact the question of whether a person qualifies as an employee. This applies in particular with regard to the possibilities it opens up for control and manipulation, which beg the question of whether, in a concrete case, an employee might be subject to external control but not to instruction. In this context, the Commission's proposal for a directive to improve working conditions in platform work is also important, as it contains remarkable provisions on so-called "algorithmic management".[1053]

[1053] Cf. *Krause*, NZA 2022, 521 (529ff.).

9. The granting of legal capacity to AI systems should be rejected in view of the current state of technical development.

10. Pursuant to Section 106(1) GewO and Section 315(1) BGB, the exercise of the right to issue instructions requires "equitable discretion". Since AI systems cannot exercise such discretion, the exercise of the right to issue instructions cannot be left to them.

11. Anti-discrimination law faces serious challenges in view of the increasing use of AI. This is particularly evident in the problem of so-called indirect discrimination, since the reference to an "apparently neutral provision, criterion or practice" (Section 3(2) AGG) describes the very way in which AI functions, which is supposed to detect characteristics that indicate the qualities sought by the employer. The danger that these (at the same time) refer to grounds of discrimination would then be almost palpable. In view of this and other problems, it is advisable to be open to the idea of a fundamental "restructuring" of anti-discrimination law, strengthening the concept of prevention, granting (further) facilitation of evidence for those affected and expanding the possibility of collective legal protection.

12. The current data protection law is also undergoing a shake-up under the influence of AI. To cite just one example, various principles for the processing of personal data, such as purpose limitation or data minimisation, are increasingly in conflict with the way AI works, a fact that calls the basic assumptions of current data protection law into question. Accordingly, many are calling for a move away from the current reference to persons and towards preventive regulation, which would be based on the means of analysis instead of the personal datum. Other key words in the reform discussion are: "data protection through technology design", an "accompanying legality control" that takes into account the constant change of algorithms and the conception of data protection (inter alia) as "collective goods protection".

13. The light and dark sides of AI are reflected in the field of occupational safety and health. AI can make a decisive contribution to the protection of workers. At the same time, however, it poses numerous hazards, some of which are new and all of which must be taken extremely seriously. It should also be borne in mind that AI often interacts with sensor technology and robotisation. The Commission has explicitly acknowledged the arising challenges in its proposal for a directive to improve working conditions in platform work. However, the issues go far beyond the area of employment on digital platforms.

14. The current liability law is also challenged by AI. AI liability law is currently in flux, with contours of non-contractual and contractual liability emerging. The scope of a possible strict liability will also be important, not least from the employee perspective.

15. The current co-determination regulations are based on the idea of the employee's "membership" in the enterprise as a "social association" and are committed in particular to the protection of personality and human dignity. Neither the one nor the other has lost any of its relevance under the conditions of the digitalisation of working life. In fact, the protection of personality and human dignity seems more urgent than ever in the face of increased threats. At the same time, co-determination must take new challenges into account. These include the possibility of using automated decision-making systems. In this respect, workplace co-determination is urgently needed to ensure that decisions are always made and answered for by people. Beyond the - manageable - new regulations of the *Betriebsrätemodernisierungsgesetz* (Works Council Modernisation Act), it is necessary to achieve a genuine modernisation of the works constitution, one which ensures sufficient control of AI, secures the expertise of the stakeholders in the company processes and thereby also increases the acceptance of AI solutions, the use of which has been thoroughly and carefully considered in advance.

16. The current focus on the development and widespread use of "human-like" AI could prove to be misguided. Instead, increased effort should be directed toward putting AI at the service of people.[1054]

In all of this, AI can and must be designed. This raises many questions, such as the extent to which black box algorithms are even tolerable, for example when it comes to decisions about job applications.[1055] But the basic question that cannot be avoided is: What kind of AI do we want?

[1054] Cf. *Brynjolfsson*, The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence, Dædalus 2022, 272. https://doi.org/10.48550/arXiv.2201.04200: “an excessive focus on developing and deploying HLAI [human-like artificial intelligence] can lead us into a trap. As machines become better substitutes for human labor, workers lose economic and political bargaining power and become increasingly dependent on those who control the technology. In contrast, when AI is focused on augmenting humans rather than mimicking them, humans retain the power to insist on a share of the value created. What is more, augmentation creates new capabilities and new products and services, ultimately generating far more value than merely human-like AI.”

[1055] Impressively, *Fletcher/Larson*, Optimizing Machines Is Perilous. Consider 'Creatively Adequate' AI – The future of artificial intelligence needs less data and can tolerate ambiguity, 25 Jan 2022: https://www.wired.com: "The push for optimisation has created design features that are either opaque (riddled with "black box" algorithms that no computer scientist can fathom) or infantilizing [...]. These features should all be walked back. Black box algorithms should be eliminated entirely; if we don't know what a computer is doing, it doesn't either".

Bibliography

Adams-Prassl, Jeremias: Regulating algorithms at work: Lessons for a „European approach to artificial intelligence“, ELLJ 2022, 30.

Adams-Prassl, Jeremias: What if Your Boss Was an Algorithm? The Rise of Artificial Intelligence at Work, Comparative Labor Law & Policy Journal 2019, 123.

Adams-Prassl, Jeremias/Binns, Reuben/Kelly-Lyth, Aislinn: Directly Discriminatory Algorithms, Modern Law Review 2022.

Ad hoc Committee on Artificial Intelligence (CAHAI): Feasibility Study, 2020.

Ajunwa, Ifeoma: Algorithms at Work: Productivity Monitoring Applications and Wearable Technology as the New Data-Centric Research Agenda for Employment and Labor Law, September 10, 2018, St. Louis U. L.J. 2019, 21: https://ssrn.com/abstract=3247286.

Ajunwa, Ifeoma: An Auditing Imperative for Automated Hiring Systems, Harvard Journal of Law & Technology 2021, 1.

Ajunwa, Ifeoma: The Paradox of Automation as Anti-Bias Intervention, Cardozo Law Review 2020, 1671.

Allen, Robin/Master, Dee: Technology Managing People - the legal implications, A report for the Trades Union Congress by the AI Law Consultancy, 2021.

Alkhatib, Ali/Bernstein, Michael: Street-Level Algorithms: A Theory at the Gaps Between Policy and Decisions, CHI 2019 Paper. https://doi.org/10.1145/3290605.3300760.

Aloisi, Antonio/De Stefano, Valerio: Introducing the Algorithmic Boss, April 20, 2021: https://www.ie.edu/insights/articles/introducing-the-algorithmic-boss.

Annuß, Georg: Betriebliche Mitbestimmung für das 21. Jahrhundert?, NZA 2022, 694.

Amolf: Responsive soft robots inspired by sputtering ketchup bottle, July 8, 2022. https://techxplore.com/news/2022-07-responsive-soft-robots-sputtering-ketchup.html?utm_campaign.

Armour, John/Eidenmüller, Horst: Self-driving Corporations, European Corporate Governance Institute - Law Working Paper No. 475/2019.

Arnold, Christian/Günther, Jens: Arbeitsrecht 4.0, 1st ed. 2018.

Bales, Richard A./V. W. Stone, Katerine: An Invisible Web at Work: Artificial Intelligence and Electronic Surveillance at the Workplace, Berkeley Journal of Employment & Labor Law, 2020, 1.

Ball, Kristie: Electronic Monitoring and Surveillance in the Workplace – Literature review and policy recommendations, 2021.

Banteka, Nadia: Artificially Intelligent Persons, Houston Law Review 2020: https://ssrn.com/abstract=3552269.

Barczak, Tristan: Algorithmus als Arkanum - Zu Staatsgeheimnissen im Digitalzeitalter und normativen Fundamenten einer Digitalordnung, DÖV 2020, 997.

Barocas, Solon/Hardt, Moritz/Narayanan, Arvind: Fairness and Machine Learning Limitations and Opportunities, 2021: https://fairmlbook.org/pdf/fairmlbook.pdf.

Barocas, Solon/Self, Andrew D.: Big Data's Disparate Impact, California Law Review 2016, 671.

Bartolo, Louisa/Thomas, Rachel: Qualitative humanities research is crucial to AI. https://www.fast.ai/ 2022/06/01/qualitative/.

Baumgartner, Ulrich/Gausling, Tina: Datenschutz durch Technikgestaltung und datenschutzfreundliche Voreinstellungen, ZD 2017, 308.

Bayreuther, Frank: Arbeitnehmereigenschaft und die Leistung fremdbestimmter Arbeit am Beispiel des Crowdworkers, RdA 2020, 241.

Bellon, Tina: Uber revamps driver pay algorithm in large U.S. pilot to attract drivers, Feb 26, 2022: https://www.reuters.com.

Bérastégui, Pierre: Exposure to psychosocial risk factors in the gig economy: a systematic review, ETUI Report, European Trade Union Institute, 2021: https://www.etui.org/sites/.

Berg, Janine/Furrer, Marianne/Harmon, Ellie/Rani, Uma/Silberman, M. Six: Digital labour platforms and the future of work Towards decent work in the online world, ILO, 2018.

Bernhardt, Annette/Kresge, Lisa/Suleiman, Reem: Data and Algorithms at Work - The Case for for Worker Technology Rights, November 2021.

Bertolini, Andrea: Artificial Intelligence and Civil Liability - Study requested by the JURI Committee, 2020.

Beyerer, J./Müller-Quade J. et al.: Protecting AI systems, preventing misuse - measures and scenarios in five application areas, white paper, n.d.

Bhuiyan, J.: Instacart shoppers say they face unforgiving metrics: 'It's a very easy job to lose'. Los Angeles Times. August 27, 2019.

Binns, Reuben/Veale, Michael: Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR, International Data Privacy Law 2021, 319.

Birhanem, Adeba/Ruane, Elayne/Laurent, Thomas/Brown, Matthew S./Flowers, Jonathan/Ventresque, Anthony/Dancy, Christopher L.: The Forgotten Margins of AI Ethics, 13. https://arxiv.org/pdf/2205.04221. pdf.

Bodie, Matthew T./Cherry, Miriam A./McCormick, Marcia L./Tang, Jintong: The Law and Policy of People Analytics, Saint Louis U. Legal Studies Research Paper No. 2016-6.

Botta, Jonas: Delegierte Selbstbestimmung?, MMR 2021, 946.

Boucher, Philip: Artificial Intelligence: How does it work, why does it matter, and what can we do about it?, 2020.

Brink, Stefan/Groß, Isabel Jana: Die DS-GVO wirkt ... und muss verbessern werden, RuP 2019, 105.

Brink, Stefan/Schwartmann, Rolf/Oetjen, Jan/Voss, Axel: In der Anwendung der DSGVO läuft einiges schief – so war sie nicht gemeint, FAZ 18 July 2022, p. 18.

Brink, Stefan/Wolff, Amadeus (eds.), BeckOK Datenschutzrecht, 39th Edition 2022.

Brown, Sara: Why it's time for „data-centric artificial intelligence", June 7, 2022. https://mitsloan.mit.edu/ ideas-made-to-matter/why-its-time-data-centric-artificial-intelligence?utm_campaign=Artificial%2BIntelligence%2BWeekly&utm_medium=email&utm_source=Artificial_Intelligence_Weekly_279.

Brownsword, Roger: Law, Technology and Society - Reimagining the Regulatory Environment, 2019.

Brynjolfsson, Erik: The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence, Dædalus 2022, 272. https://doi.org/10.48550/arXiv.2201.04200.

Buchholtz, Gabriele/Scheffel-Kain, Martin: Algorithmen und Proxy Discrimination in der Verwaltung: Vorschläge zur Wahrung digitaler Gleichheit, NVwZ 2022, 612.

Buiten, Miriam/de Streel, Alexandre/Peitz, Martin: EU Liability Rules for the Age of Artificial Intelligence, 2021.

Burgess, Matt: How GDPR Is Failing – The world-leading data law changed how companies work. But four years on, there's a lag on cleaning up Big Tech, May 23, 2022. https://www.wired.com/story/gdpr-2022/.

Butollo, Florian/Jürgens, Ulrich/Krzywdzinski, Martin: From lean production to Industrie 4.0. More autonomy for employees?, WZB Discussion Paper, No. SP III 2018-303, 2018.

Byrum, Joseph: Build a Diverse Team to Solve the AI Riddle, May 18, 2020: https://sloanreview.mit.edu/article/build-a-diverse-team-to-solve-the-ai-riddle.

Calo, Ryan/Rosenblat, Alex: The Taking Economy: Uber, Information, and Power Columbia Law Review 2017,1623.

Casalone, Carlo et al.: Human-centric AI: From Principles to Actionable and Shared Policies, September 2021.

Castelluccia, Claude/Le Métayer, Daniel: Understanding algorithmic decision-making: opportunities and challenges, 2019, 78.

Chalutz Ben-Gal, Hila: Human Resources Based Organizational Data Mining (HRODM): Themes, Trends, Focus, Future (2020).

Chander, Anupam: The Racist Algorithm?, Michigan Law Review 2017, 1023.

Chesterman, Simon: Artificial Intelligence and the Limits of Legal Personality, in: International and Comparative Law Quarterly 2020, 819.

Choi, Charles Q.: New Test Compares AI Reasoning With Human Thinking – The novel technique can help researchers see if AIs reason as hoped and are trustworthy, 27 April 2022. https://spectrum.ieee.org/ trustworthy-ai.

Choi, Yejin: The Curious Case of Commonsense Intelligence, Dædalus 2022, 139. https://doi.org/10.1162/ DAED_a_01906.

Circiumaru, Alexandru: Three proposals to strengthen the EU Artificial Intelligence Act - Recommendations to improve the regulation of AI - in Europe and worldwide: https://www.adalovelaceinstitute.org/blog/three-proposals-strengthen-eu-artificial-intelligence-act/.

Clinton, Paul: Smarter Video Telematics Wave Arrives, Automotive Fleet, March 19, 2019: https://www. automotive-fleet.com/327438/wave-of-smarter-video-telematics-solutions-arrives.

Collins, Laurence/Fineman, David R./Tsuchida, Akio: People analytics: Recalculating the route, Global Human Capital Trends, Deloitte Insights, February 28, 2017.

Columbus, Louis: How AI is shaping the future of work, June 9, 2022. https://venturebeat. com/2022/06/09/how-ai-is-shaping-the-future-of-work/.

Commissioner for Human Rights: Unboxing Artificial Intelligence: 10 steps to protect Human Rights; Council of Europe 2019.

Conrad, Conrad S.: Artificial Intelligence: New Consent Solutions to Data Protection, InTer 2021, 147.

Council of Europe: Algorithms and Human Rights - Study on the human rights dimensions of automated data processing techniques and possible regulatory implications, Council of Europe study DGI(2017) 12 prepared by the committee of experts on internet intermediaries (MSI-NET), 2018.

Council of Europe: Discrimination, artificial intelligence, and algorithmic decision-making, 2018.

Council of Europe: Towards Regulation of AI Systems - Global perspectives on the development of a legal framework on Artificial Intelligence (AI) systems based on the Council of Europe's standards on human rights, democracy and the rule of law, Compilation of contributions DGI (2020).

Däubler, Wolfgang: Digitalisierung und Arbeitsrecht – Künstliche Intelligenz – Homeoffice – Arbeit 4.0, 7. Aufl. 2020.

Däubler, Wolfgang/Beck, Thorsten (eds.): Allgemeines Gleichbehandlungsgesetz mit Entgelttransparenzgesetz, Berliner LADG, 5. Aufl., 2021.

Däubler, Wolfgang/Klebe, Thomas/Wedde, Peter: Betriebsverfassungsgesetz, 18. Aufl. 2022.

Däubler, Wolfgang/Wedde, Peter/Weichert, Thilo/Sommer, Imke (eds.): EU-DSGVO und BDSG, 2. Aufl. 2020.

Dayton, Leigh: Call for human rights protections on emerging brain-computer interface technologies - Industry self-regulation is not enough, say AI researchers: nature index 16 March 2021.

de Hert, Paul/Lazcoz, Guillermo: Radical rewriting of Article 22 GDPR on machine decisions in the AI era, European Law Blog 13 Oct 2021: https://europeanlawblog.eu.

De Stefano, Valerio/Durri, Ilda/Stylogiannis, Charalampos/Wouters, Mathias: Platform work and the employment relationship, ILO Working Paper 27, March 2021.

De Stefano, Valerio/Wouters, Mathias: AI and digital tools in workplace management and evaluation – An assessment of the EU's legal framework, May 2022.

Dehmel, Susanne: Rück- und Ausblick zur DS-GVO, ZD 2020, 62.

Dellermann, Dominik/Ebel, Philipp/Sollner, Matthias/Leimeister, Jan Marco: Hybrid Intelligence, 2018.

Dewey, John: The quest for certainty, 1929.

Diamantis, Mihailis: Employed Algorithms: A Labor Model of Corporate Liability for AI, October 19, 2021. 72 Duke L.J.: https://ssrn.com/abstract=3945882.

Djeffal, Christian: The Normative Potential of the European Rule on Automated Decisions: A New Reading for Art. 22 GDPR, ZaöRV 2020, 847.

Doellgast, Virginia/O'Brady Sean: Making call centre jobs better: The relationship between management practices and worker stress A Report for the CWA, 2020.

Dzida, Boris/Groh, Naemi: Diskriminierung nach dem AGG beim Einsatz von Algorithmen im Bewerbungsverfahren, NJW 2018, 1917.

Dzieza, Josh: "Robots Aren't Taking Our Jobs - They're Becoming Our Bosses." The Verge, February 27, 2020.

Ebers, Martin/ Heinze, Christian/Krügel, Tina/Steinrötter, Björn (eds.): Künstliche Intelligenz und Robotik, 2020.

Ebert, Philip/Freibichler, Wolfgang: Nudge management: applying behavioural science to increase knowledge worker productivity, Journal of Organization Design 2017, 6:4.

Ecoffet, Adrien/Huizinga, Joost/Lehman, Joel/Stanley, Kenneth O./Clune, Jeff: First return, then explore, Nature 2021, 580.

Edwards, Lilian: Regulating AI in Europe: four problems and four solution, March 2022.

Eiben, Ágoston E./Ellers, Jacintha/ Meynen, Gerben/Nyholm, Sven: Robot Evolution - Ethical Concerns, frontiers in Robotics and AI, 03 November201: doi: 10.3389/frobt.2021.744590.

Elton, Daniel C.: Self-explaining AI as an Alternative to Interpretable, in: Goertzel, Ben/Panov, Aleksandr/Potapovm Alexey, Yampolskiy, Roman (eds.): Artificial General Intelligence, 13th International Conference, AGI 2020, St. Petersburg, Russia, September 16-19, 2020, Proceedings, 2020, p. 95.

Enders, Peter: The use of artificial intelligence in legal decision-making, JA 2018, 721.

Epp, Clayton/Lippold, Michael/Mandryk, Regan L.: Identifying emotional states using keystroke dynamics, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2011, p. 715.

Esser, Josef: Vorverständnis und Methodenwahl in der Rechtsfindung: Rationalitätsgarantien der richterlichen Entscheidungspraxis, 1970.

European Agency for Safety and Health at Work: Actions by labour and social security inspectorates for the improvement of occupational safety and health in platform work, Policy Case Study 2022.

European Agency for Safety and Health at Work: Digital platform work and occupational safety and health: a review, 2021.

European Agency for Safety and Health at Work: Foresight on new and emerging occupational safety and health risks associated with digitalisation by 2025, European Risk Observatory Report, 2018.

European Agency for Safety and Health at Work: Impact of Artificial Intelligence on Occupational Safety and, Health, Policy Brief, 2021.

European Agency for Safety and Health at Work: Occupational Safety and Health in Digital Platform Wprk: Lessions from Regulations, Policies, Actions and Initiatives, Policy Brief 2021.

European Agency for Safety and Health at Work: OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, Discussion Paper, 2019.

European Parliament: AIDA Working Paper on AI and Bias, November 2021.

Evas, Tatjana: Civil liability regime for artificial intelligence - European added value assessment, Study European Research Service, 2020.

Expert Group on Liability and New Technologies - New Technologies Formation: Liability for Artificial Intelligence and other emerging digital technologies, 2019.

Finck, Michèle: Blockchain and the General Data Protection Regulation, 2019.

Fineman, D. R.: People analytics: Recalculating the route, Deloitte Insights, 2017.

Fjeld, Jessica/Achten, Nele/Hilligoss, Hannah/Nagy, Adam Christopher/Srikumar, Madhulika: Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI, Berkman Klein Center for Internet & Society at Harvard University, Research Publication No. 2020-1.

Fletcher, Angus/Larson, Erik J.: Optimizing Machines Is Perilous. Consider 'Creatively Adequate' AI - The future of artificial intelligence needs less data and can tolerate ambiguity, Jan 25, 2022: https://www.wired.com.

Floridi, Luciano/Holweg, Matthias/Taddeo, Mariarosaria/Amaya Silva, Javier/Mökander, Jakob/Wen, Yuni: capAI - A Procedure for Conducting Conformity Assessment of AI Systems in Line with the EU Artificial Intelligence Act, March 23, 2022: https://ssrn.com/abstract=4064091.

Forgó, Nikolaus/Helfrich, Marcus/Schneider, Jochen: Betrieblicher Datenschutz, 3rd ed. 2019.

Forgó, Nikolaus/Krügel, Tina: Der Personenbezug von Geodaten - Cui bono, wenn alles bestimmbar ist?, MMR 2010, 17.

Frank, Justus/Heine, Maurice: NZA 2021, 1448.

Franzen, Martin: Beschäftigtendatenschutz aus Luxemburg?, EuZA 2022, 261.

Freed, Sam: AI and Human Thought and Emotion, 2019.

Freed, Sam: AGI Needs the Humanities, in: Goertzel, Ben/Panov, Aleksandr/Potapovm Alexey, Yampolskiy, Roman (eds.): Artificial General Intelligence, 13th International Conference, AGI 2020, St. Petersburg, Russia, September 16-19, 2020, Proceedings, 2020, p. 107.

Freed, Sam: Report on "AI and Human Thought and Emotion" in: Goertzel, Ben/Panov, Aleksandr/Potapovm Alexey, Yampolskiy, Roman (eds.): Artificial General Intelligence, 13th International Conference, AGI 2020, St. Petersburg, Russia, September 16-19, 2020, Proceedings, 2020, p. 116.

Freund, Stefan: Die Abwägung im Gesellschaftsrecht, NZG 2020, 1328.

Freyler, Carmen: Robot-Recruiting, Künstliche Intelligenz und das Antidiskriminierungsrecht, NZA 2020, 284.

Friedmann, Cindy: Ethical concerns with replacing human relations with humanoid robots: an ubuntu perspective. https://doi.org/10.1007/s43681-022-00186-0.

Friedman, Atya / Nissenbaum, Helen: Bias in Computer Systems, ACM Transactions on Information Systems 1996, 330.

Fuller, Joseph B./Raman, Manjari/Bailey, Allison/Vaduganathan, Nithya et al.: Building the on-demand workforce, Harvard Business School, November 2020.

Gäbert, Jens: Reichweite der Mitbestimmungsrechte im Gesundheitsschutz gem. § 87 Abs. 1 Nr. 7 BetrVG bei Planung und Gestaltung von Arbeitsbedingungen, AuR 2021, 9.

Gasparotti, Alessandro/Harta, Lukas: European Strategy on Artificial Intelligence An Assessment of the EU Commission's Draft White Paper on AI, 2020,cepAdhoc v. 11.12.2020.

Gasparotti, Alessandro: Ethics Guidelines on Artificial Intelligence A comparison of EU and OECD guidelines, cepInput 07, 2019.

Geminn, Christian L.: Fairness und Transparenz im Datenschutzrecht, ZD-Aktuell 2021, 05557

Geminn, Christian L.: Menschenwürde und menschenähnliche Maschinen und Systeme, DÖV 2020, 172.

Genser, Jared/Herrmann, Stephanie/Yuste, Rafael: International Human Rights Protection Gaps in the Age of Neurotechnology, 2022.

Ghassemi, Marzyeh/Oakden-Rayner, Luke/Beam, Andrew L: The false hope of current approaches to explainable artificial intelligence in health care, Viewpoint November 01, 2021, e745: DOI:https://doi.org/10.1016/S2589-7500(21)00208-9.

Gierschmann, Sibylle: Gestaltungsmöglichkeiten durch systematisches und risikobasiertes Vorgehen - Was ist schon anonymous?, ZD 2021, 482.

Gilbert, Abigail/Thomas, Anna/Pissarides, Christopher/Al-Izzi, Hana/Miller, Catherine/Burnell, Burneöl: The Amazonian Era – How algorithmic systems are eroding good work, Institute for the Future of Work, 2021.

Gill, Karamjit S.: Nowotny, Helga (2021). In AI we trust: power, illusion and control of predictive algorithms, Polity, Cambridge, UK, 2021, AI & SOCIETY (2022) 37:411.

Global Commission on the Future of Work: Work for a brighter future, 2019.

Gola, Peter/Heckmann, Dirk: Bundesdatenschutzgesetz, 13th ed. 2019.

Göpfert, Burkard/Brune, Jan-Philipp: Moderne Führungsinstrumente auf dem arbeitsrechtlichen Prüfstand, NZA-Beil. 2018, 87.

Göpfert, Burkard/Seier, Jochen: Die "Transformations-Einigungsstelle": Inhalt und Grenzen eines "Qualifizierungs-Sozialplans", NZA 2019, 588.

Götz, Thomas: Big Data im Personalmanagement - Datenschutzrecht und betriebliche Mitbestimmung, 2020.

Grabitz, Eberhard/Hilf, Meinhard/Nettesheim, Martin: Das Recht der Europäischen Union 2021.

Graevenitz von, Albrecht: "Zwei mal Zwei ist Grün" - Mensch und KI im Vergleich, ZRP 2018, 238.

Greiner, Stefan/Kalle, Ansgar: Betriebliche Mitbestimmung nach § 87 Abs. 1 Nr. 1 und 6 BetrVG – Konvergenzen und Widersprüche im Digitalisierungskontext, RdA 2021, 76.

Grimmelmann, James/Westreich, Daniel: Incomprehensible Discrimination, California Law Review Online 2017, 164.

Grünberger, Michael: Reformbedarf im AGG: Beweislastverteilung beim Einsatz von KI, ZRP 2021, 231.

Grützmacher, Malte: Die deliktische Haftung für autonome Systeme - Industrie 4.0 als Herausforderung für das bestehende Recht?, CR 2016, 695.

Grützmacher, Malte: Die zivilrechtliche Haftung für KI nach dem Entwurf der geplanten KI-VO - Potential zivilrechtliche Auswirkungen des geplanten KI-Sicherheitsrechts: ein neues Schutzgesetz i.S.v. § 823 Abs. 2 BGB am Horizont, CR 2021, 433.

Gurley, Lauren Kaori: Internal Documents Show Amazon's Dystopian System for Tracking Workers Every Minute of Their Shifts.https://www.vice.com/en/article/5dgn73/internal-documents-show-amazonsdystopian-system-for-tracking-workers-every-minute-of-their-shifts.

Gurovich, Yaron/Hanani, Yair/Bar, Omri/Nadav, Guy/Fleischer, Nicole/Gelbman, Dekel/Basel-Salmon, Lina/Krawitz, Peter M./Kamphausen, Susanne B./Zenker, Martin: Identifying facial phenotypes of genetic disorders using deep learning; in: Nature medicine, 2019, p. 60.

Haagen, Christian: Verantwortung für Künstliche Intelligenz - Ethische Aspekte und zivilrechtliche Anforderungen bei der Herstellung von KI-Systemen, 2021.

Hacker, Philipp/Passoth, Jan-Hendrik: Varieties of AI Explanations under the Law. From the GDPR to the AIA, and Beyond, in: Holzinger, Andreas/Goebel, Randy/Fong, Ruth/Moon, Taesup/Müller, Klaus-Robert/Samek, Wojciech (eds.), Lecture Notes on Artificial Intelligence 13200: AI - beyond explainable AI, Springer, 2022 :http://dx.doi.org/10.2139/ssrn.3911324.

Hacker, Philipp: Europäische und nationale Regulierung von Künstlicher Intelligenz, NJW 2020, 2142.

Hacker, Philipp: Behavioural and knowledge attribution in the use of artificial intelligence, RW 2018, 243.

Häferer, Katka/Koops, Christian: Crowdworker als Arbeitnehmer, NJW 2021, 1787.

Hallinan, Rara/Leenes, Ronald/De Hert, Paul (eds.): Data Protection and Privacy Data Protection and Artificial Intelligence, 2021.

Hamon, Ronan/Junklewitz, Henrik/Sanchez, Ignacio/Malgieri, Gianclaudio/De Hert, Paul: Bridging the Gap Between AI and Explainability in the GDPR: Towards Trustworthiness-by-Design in Automated Decision-Making, in IEEE Computational Intelligence Magazine, Feb 2022, 72: doi: 10.1109/MCI.2021.3129960.

Hanley, Daniel A./Hubbard, Sally: Eyes Everywhere: Amazon's Surveillance Infrastructure and Revitalizing Worker Power, September 1, 2020: https://www.openmarketsinstitute.org.

Hanoch, Yaniv/Arvizzigno, Francesco/Hernandez García, Daniel/Denham, Sue/Belpaeme, Tony/Gummerum, Michaela: The Robot Made Me Do It: Human-Robot Interaction and Risk-Taking Behavior, Cyberpsychology, Behavior, and Social Networking 2021: doi: 10.1089/cyber.2020.0148**.**

Hardré, Patricia: When, How, and Why Do We Trust Technology Too Much?, in: Tettegah, Sharon Y./ Espelage, Dorothy L. (eds.), Emotions, Technology, and Behaviors - A Volume in Emotions and Technology, 2016, 85.

Haußmann, Katrin/Thieme, Luca Maria: Reformbedarf und Handlungsoptionen in der IT-Mitbestimmung, NZA 2019, 1612.

Haynes, John-Dylan/Sakai, Katsuyuki/Rees, Geraint/Gilbert, Sam/Frith, Chris/Passingham, Richard E.: Reading Hidden Intentions in the Human Brain, Current Biology 2007, 323.

Heberlein, Horst: Two years of application of the GDPR, ZD 2020, 487.

Heckelmann, Martin: Crowdworking - eine arbeitsrechtliche Bestandsaufnahme, NZA 2022, 73.

Heiderhoff, Bettina/Gramsch, Kilian: Klassische Haftungsregimes und autonome Systeme - genügt "functional equivalence" oder bedarf es eigenständiger Maßstäbe?, ZIP 2020, 1937.

Heimstädt, Maximilian/Dobusch, Leonhard: Streik-Vorhersage mit Twitter-Daten, FAZ v. 11.4.2022, p. 16.

Heiss, Stefan: Europäische Haftungsregeln für Künstliche Intelligenz, EuZW 2021, 93.

Henssler, Martin: General Report on the Three Panels of the 4th German Labour Law Day, NZA Supplement 2020, 3.

Herberger, Maximilian/Martinek, Michael/Rüßmann, Helmut/Weth, Stephan/Würdinger (eds.), Markus: jurisPK-BGB Bd. 2.

Herberger, Maximilian: "Artificial Intelligence" and Law, NJW 2018, 2825.

Herbosch, Maarten: The Diligent Use of AI Systems: A Risk Worth Taking?, EuCML 2022, 14.

Hern, Alex: Amazon's Alexa could turn dead loved ones' voices into digital assistant – Technology promises ability to 'make the memories last' by mimicking the voice of anyone it hears, 23 June 2022. https://www.theguardian.com/technology/2022/jun/23/amazon-alexa-could-turn-dead-loved-ones-digitalassistant?tpcc=nleyeonai.

Herzog, Roman/Herdegen, Matthias/Scholz, Rupert/Klein, Hans H.: Grundgesetz Kommentar, 92nd Ergänzungslieferung August 2020.

Hießl, Christina: Case law on algorithmic management at the workplace: Cross-European comparative analysis and tentative conclusions (September 1, 2021): https://ssrn.com/abstract=3982735.

Hießl, Christina: Case law on the classification of platform workers: Cross-European comparative analysis and tentative conclusions: https://ssrn.com/abstract=3839603.

High-Level Expert Group on Artificial Intelligence: Policy and Investment Recommendations for Trustworthy AI, 2019.

Hilb, Michael: Toward artificial governance? The role of artificial intelligence in shaping the future of corporate governance, Journal of Management and Governance, 2020, 851.

Hoeren, Thomas/Sieber, Ulrich/Holznagel, Bernd (eds.): Handbuch Multimedia-Recht, 57th ed. 2022.

Hofmann, Franz: The influence of digitalisation and artificial intelligence on liability law, CR 2020, 282.

Holthausen, Joachim: Big Data, People Analytics, KI und Gestaltung von Betriebsvereinbarungen - Grund-, arbeits- und datenschutzrechtliche An- und Herausforderungen, RdA 2021, 19.

Honer, Mathias: Nudging: No Challenge for the Dogmatics of Fundamental Rights, DÖV 2019, 940.

Höpfner, Clemens/Daum, Jan Alexander: Der "Robo-Boss" - Künstliche Intelligenz im Arbeitsverhältnis, ZfA 2021, 467.

Horner, Susanne/Kaulartz, Markus: Haftung 4.0, CR 2016, 7.

Horstmeier, Gerrit: Ein digitales Upgrade für das Betriebsverfassungsrecht?, BB 2022, 116.

Huff, Julian/Götz, Thomas: Evidence instead of gut feeling? - Possibilities and legal limits of Big Data in HR. NZA Supplement 2019, 73.

Hutson, Matthew: Can Computers Learn Common Sense? A.I. researchers are making progress on a longterm goal: giving their programs the kind of knowledge we take for granted, April 5, 2022. https://www. newyorker.com/tech/annals-of-technology/can-computers-learn-common-sense.

Ienca, Marcello /Andorno, Roberto: Towards new human rights in the age of neuroscience and neurotechnology, Liefe Sciences, Society and Policy, 2017: https://doi.org/10.1186/s40504-017-0050-1 2017.

Ienca, Marcello: Brain Machine Interfaces, Artificial Intelligence and Neurorights: https://brain.ieee.org/.

Institute of Electrical and Electronics Engineers: Ethically Aligned Design, - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent System, 1st ed, 2019.

International Labour Organization: Synthesis Report of the National Dialogues on the Future of Work, 2017.

Ivankovics, Peter: Gentle control or manipulative influence? On the possibility of encroachment on fundamental rights through nudging, JuWissBlog No. 57/2018.

Jansen, Anne/van der Beek, Dolf/Cremers, Anita/Neerincx, Mark/van Middelaar, Johan: Emergent Risks to Workplace Safety, Working in the Same Spot as a Cobot, Report for the Ministry of Social Affairs and Employment, 2018.

Jernigan, Carter/Mistree, Behram F. T.: Gaydar: Facebook friendships expose sexual orientation; in: First Monday 2009, No. 10.

Jesuthasan, Ravin /Boudreau, John: Work Without Jobs - We need a new operating system built on deconstructed jobs and organisational agility, Jan 5, 2021, MIT Sloan Management Review.

Jesuthasan, Ravin /Boudreau, John: Are You Ready to Lead Work Without Jobs? We're moving toward a system of work, design that will profoundly change the roles of organizational leaders, 08 April 2021, MIT Sloan Management Review.

Jesuthasan, Ravin /Boudreau, John: Reinventing Jobs - A 4-Step Approach for Applying Automation to Work, 2018.

Jesuthasan, Ravin /Boudreau, John: Work without Jobs: How to Reboot Your Organization's Work Operating System (Management on the Cutting Edge), 2022.

Johnson, Kristin N.: Automating the Risk of Bias, The George Washington Law Review 2019, 1214.

Johnson, Melanie: 7 effective uses of AI in recruitment. https://www.unleash.ai/artificial-intelligence/ 7-effective-uses-of-ai-in-recruitment/.

Joos, Daniel/Meding, Kristofer: Technical Organisational Measures (TOMs) in "smart" employer decisions, CR 2020, 834.

Joos, Daniel: Einsatz von künstlicher Intelligenz im Personalwesen unter Beachtung der DS-GVO und des BDSG, NZA 2020, 1216.

Jüngling, Alexander: Die Digitalstrategie der EU-Kommission: Regulierung von Künstlicher Intelligenz, MMR 2020, 440.

Junker, Abbo: Doppelt gemoppelt hält nicht besser – Der Kommissionsvorschlag zur Plattformarbeit, EuZA 2022, 141.

Junyang Lin/Men, Rui/Yang, An: M6: A Chinese Multimodal Pretrainer, 2021.

Kahn, Jeremy: Researchers are peering inside computer brains. What they've found will surprise you, Fortune March 4, 2021.

Kaiser, Stepan/Kraus, Hans: Big Data im Personalmanagement: Erste Anwendungen und ein Blick in die Zukunft, ZfO 2014, 379.

Kaminski, Margot E./Malgieri, Gianclaudio: Algorithmic Impact Assessments under the GDPR: Producing Multi-layered Explanations (September 18, 2019), International Data Privacy Law, 2020, forthcoming, U of Colorado Law Legal Studies Research Paper No. 19–28. https://ssrn.com/abstract=3456224.

Kang, Jerry/Bennett, Mark W./Carbado, Devon W./Casey, Pamela/Dasgupta, Nilanjana/Faigman, David L./Godsil, Rachel D./Greenwald, Anthony/Levinson, Justin D./Mnookin, Jennifer L.: Implicit Bias in the Courtroom, UCLA Law Review 2012: https://ssrn.com/abstract=2026540.

Kang, Jerry: What Judges Can Do About Implicit Bias, Court Review 2021: https://ssrn.com/abstract=4033906.

Kaulartz, Markus/Braegelmann, Tom (eds.): Rechtshandbuch Artificial Intelligence und Machine Learning, 2020.

Kaushik, Divyansh/Lipton, Zachary C./London, Alex John: Resolving the Human Subjects Status of Machine Learning's Crowdworkers. https://arxiv.org/pdf/2206.04039.pdf.

Kelber, Ulrich: "Alle meine Daten" - der Abschlussbericht der Datenethikkommission, ZD 2020, 7.

Kellogg, Katherine/Vantine, Melissa A./Christin, Angèle: Algorithms at Work: The New Contested Terrain of Control, Academy of Management Annals 2020, 366.

Kelly-Lyth, Aislinn: The AI Act and Algorithmic Management, Comparative Labor Law & Policy Journal, Dispatch No. 39, 1.

Kiel, Heinrich/Lunk, Stefan/Oetker, Harmut (eds.): Münchener Handbuch zum Arbeitsrecht, 4th ed. 2018.

Kim, Pauline: Data-Driven Discrimination at Work, William & Mary Law Review 2017, 857.

Kim, Sehoon: Working With Robots: Human Resource Development Considerations in Human-Robot Interaction, Human Resource Development Review 2022, 48.

Kirchhof, Gregor: Nudging – zu den rechtlichen Grenzen informalen Verwaltens, ZRP 2015, 136.

Klapp, Micha/Klebe, Thomas: Die Zukunft der Betriebsverfassung – ein Gesetzentwurf für das 21. Jahrhundert, NZA 2022, 689.

Klebe, Thomas: Betriebsrat 4.0 - Digital und global?, NZA Supp. 2017, 77.

Klebe, Thomas: Betriebsverfassung 2030: Zukunftsanforderungen und Weiterentwicklung, AuR 2020, 196.

Klebe, Thomas: Künstliche Intelligenz - eine Herausforderung für die Mitbestimmung, Soziales Recht 2019, 128.

Klebe, Thomas/Klengel, Ernesto: Mitbestimmungsrechte im Spiegel neuerer Rechtsprechung zum Datenschutz, NZA 2021, 1144.

Klebe, Thomas/Schmidt, Marlene/Klengel, Ernesto: Betriebsverfassung und Datenschutzrecht, in: Gräfl, Edith/Lunk, Stefan/Oetker, Hartmut/Treibinger, Yvonne (eds.), 100 Jahre Betriebsverfassungsrecht, 2020, **p**. 303.

Klingbeil, Thilo/Kohm, Simon: Datenschutzfreundliche Technikgestaltung und ihre vertraglichen Implikationen, MMR 2021, 3.

Knight, Will: This Warehouse Robot Reads Human Body Language – Machines that understand what their human teammates are doing could boost productivity without taking jobs, June 28, 2022. https://www.wired.com/story/warehouse-robot-reads-body-language/.

Knitter, Philipp: Digitale Weisungen – Arbeitgeberentscheidungen aufgrund algorithmischer Berechnung, Diss. Berlin, 2022.

Koch, Petra: Selbstständigkeit in der virtualisierten Arbeitswelt, Diss. Kassel, 2010.

Kocher, Eva: Digitale Plattformarbeit - die Verantwortung von Marktorganisatoren, ZEuP 2021, 606.

Koene, Ansgar/ Clifton, Chris/Webb, Helena/Patel, Menisha/Machad, Caio/LaViolette, Jack/Richardson, Rashida/Reisman, Dillon: A governance framework for algorithmic accountability and transparency, 2019.

Kohte, Wolfhard: Arbeitsschutz in der digitalen Arbeitswelt, NZA 2015, 1417.

Kolain, Michael/Grafenauer, Christian/Ebers, Martin: Anonymity Assessment - A Universal Tool for Measuring Anonymity of Data Sets Under the GDPR with a Special Focus on Smart Robotics, November 24, 2021, Rutgers University Computer & Technology Law Journal 2022: https://ssrn.com/abstract=3971139, 29.

Kollmar, Frederike/El-Auwad, Maya: Grenzen der Einwilligung bei hochkomplexen und technisierten Datenverarbeitungen, K & R 2021, 73.

Körner, Marita: Beschäftigtendatenschutz in Betriebsvereinbarungen unter der Geltung der DS-GVO, NJW 2018, 2825.

Kosinski, Michal/Stillwell, David/Graepel, Thore: Private traits and attributes are predictable from digital records of human behaviour; in: Proceedings of the National Academy of Sciences, 2013, 5802.

Kosinski, Michal/Wang, Yilun: Deep neural networks are more accurate than humans at detecting sexual orientation from facial images; in: Journal of Personality and Social Psychology, 2018, 246.

Kosinski, Michal/Stillwell, David/Graepel, Thore: Private traits and attributes are predictable from digital records of human behaviour; in: Proceedings of the National Academy of Sciences, 2013, 5802.

Kostopoulos, Lydia: Decoupling Human Characteristics from Algorithmic Capabilities, 2021.

Krafft, Tobias D./Zweig, Katharina A.: Transparenz und Nachvollziehbarkeit algorithmenbasierter Entscheidungsprozesse - Ein Regulierungsvorschlag aus socioinformatischer Perspektive, 2019.

Krause, Rüdiger: Auf dem Weg zur unionsrechtlichen Regelung von Plattformtätigkeiten, NZA 2022, 521.

Krause, Rüdiger: Berufliche Weiterbildung in der Transformation der Arbeitswelt. NZA 2022, 737.

Krause, Rüdiger: Sozialverträgliche Arbeitnehmerüberwachung — Technikbasierte Beschäftigtenkontrolle als Gegenstand betrieblicher Mitbestimmung im digitalen Zeitalter, in: Gräfl, Edith/Lunk, Stefan/Oetker, Hartmut/Treibinger, Yvonne (eds.): 100 Jahre Betriebsverfassungsrecht, 2020, S. 353.

Krenn, Mario/Pollice, Robert/Guo, Si Yue et al.: On scientific understanding with artificial intelligence. https://doi.org/10.48550/arXiv.2204.01467.

Kresge, Lisa: Data and Algorithms in the Workplace: A Primer on New Technologies, UC Berkeley Labor Center Working Paper, Technology and Work Program, November 2020.

Krishna, Satyapriya /Han, Tessa /Gu, Alex/Pombra, Javin/Jabbari, Shahin/Wu, Steven/Lakkaraju, Himabindu: The Disagreement Problem in Explainable Machine Learning: A Practitioner's Perspective: https://arxiv.org/abs/2202.01602.

Kritikos, Mihalis: What if blockchain could guarantee ethical AI?, European Parliamentary Research Service, 2020.

Krülls, Sebastian: Zur Notwendigkeit einer Reform des § 87 Abs. 1 Nr. 6 BetrVG, RdA 2021, 279.

Kugelmann, Dieter: Künstliche Intelligenz aus Sicht der Datenschutzaufsicht - Steuerung statt Verhinderung, Datenschutz und Datensicherheit (DuD) 2021.

Kumkar, Lea Katharina: Legal Transactions Involving Automated and Autonomous Systems, K & R 2020, 801.

Kuner, Christopher/Bygrave, Lee A./Docksey, Christopher/Drechsler, Laura: The EU General Data Protection Regulation - A Commentary, 2020.

Kuntz, Thilo: Künstliche Intelligenz, Wissenszurechnung und Wissensverantwortung, ZfPW 2022, 177.

Lambrecht, Anja/Sen, Ananya/Tucker, Catherine E./Wiertz, Caroline: Algorithmic Recommendations and Earned Media: Investigating Product Echo Chambers on YouTube, October 27, 2021: https://ssrn.com/abstract=3951425.

Lanzing, Marjolein: „Strongly Recommended“ Revisiting Decisional Privacy to Judge Hypernudging in Self-Tracking Technologies, Philosophy & Technology 2019, 549.

Larenz, Karl/Canaris, Claus-Wilhelm: Lehrbuch des Schuldrechts: Besonderer Teil, 13th ed. 1994.

Lecher, Colin: "How Amazon Automatically Tracks and Fires Warehouse Workers for 'Productivity", The Verge, August 25, 2019.

Leibold, Kevin: Reichweite, Umfang und Bedeutung des Auskunftsrechts nach Art. 15 DS-GVO - Entscheidungsübersicht, ZD-Aktuell 2021, 05313.

Levy, Karen/Barocas, Solon: Refractive Surveillance: Monitoring Customers to Manage Workers, International Journal of Communication 2018, 1166.

Lewinski von, Kai/de Barros Fritz, Raphael: Arbeitgeberhaftung nach dem AGG infolge des Einsatzes von Algorithmen bei Personalentscheidungen, NZA 2018, 620.

Li, Jingwei/Bzdok, Danilo/Chen, Jianzhong et al.: Cross-ethnicity/race generalization failure of behavioral prediction from resting-state functional connectivity, Science Advances 2022, 144. DOI: 10.1126/sciadv. abj18.

Lin Junyang et al., M6: A Chinese Multimodal Pretrainer, 2021: https://arxiv.org/abs/2103.00823.

Lindner, Ralf /Goos, Kerstin/Güth, Sandra/Som, Oliver/ Schröde, Sandra: "Responsible Research and Innovation" als Ansatz für die Forschungs-, Technologie- und Innovationspolitik -Hintergründe und Entwicklungen, Office of Technology Assessment at the German Bundestag, Background Paper No. 22, 2016.

Linnenkohl, Karl/Kilz, Gerhard/Rauschenberg, Hans-Jürgen/Reh, Dirk: Der Begriff des Arbeitnehmers und die "informationelle Abhängigkeit", ArbuR 1991, 203.

Löber, Lena Isabell: Auf dem Weg zu einem Beschäftigtendatenschutzgesetz für das digitale Zeitalter?, ZD-Aktuell 2022, 01120.

Lohmann, Melinda F./Preßler, Theresa: Die Rechtsfigur des Erfüllungsgehilfen im digitalen Zeitalter, RDi 2021, 538.

Ludwig, Daniel/Hinze, Jacob: Digitalisierung und IT-Mitbestimmung - Wie die Betriebsparteien den Wandel gemeinsam gestalten können, NZA 2021, 1444.

Malorny, Friederike: Datenschutz als Grenze KI-basierter Auswahlentscheidungen im Arbeitsrecht, RdA 2022, 170.

Marabelli, Marco/Vaast, Emmanuelle/Carlile, Paul R.: Making Lemonade: Dealing with Analytics Surveillance in the Workplace, Academy of Management Annual Meeting, 2020.

Martini, Mario/Botta, Jonas: Iron Man am Arbeitsplatz? - Exoskeletons between efficiency striving, data and health protection, NZA 2018, 625.

Martini, Mario: Blackbox Algorithmus - Grundfragen einer Regulierung Künstlicher Intelligenz, 2019.

Matejek, Michael/Mäusezahl, Steffen: Ordinary vs. sensitive personal data, ZD 2019, 551.

Mathur, Arunesh/Acar, Gunes/Friedman, Michael/Lucherini, Elena/Mayer, Jonathan/Marshini, Chetty/Narayanan, Arvind: Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites, Proceedings of the ACM on Human-Computer Interaction, 2019, 81.

Mattiuzzo, C./Vock, S./Mössner, T./Voß, S.: Safe machine with - or despite - artificial intelligence, ARP 2021, 188.

Matz, Sandra C.; Netzer, Oded: Using Big Data as a window into consumers' psychology; in: Current Opinion in Behavioral Sciences, 2017, 7.

Mayer, Jonathan/Mutchler, Patrick/Mitchell, John C.: Evaluating the privacy properties of telephone metadata; in: Proceedings of the National Academy of Sciences, 2016, 553.

Mayson, Sandra G.: Bias In, Bias Out, Yale L.J. 2019, 2218.

McCrea, Bridget: Labor Management Systems (LMS): The New Age of Employee Engagement, Logistics Management, June 3, 2020: https://www.logisticsmgmt.com.

McCrudden, Christopher: Nudging and human dignity, 06 January 2015: https://verfassungsblog.de/nudging-human-dignity-2/.

McDaniel, John L. M./Pease, Ken G.: Predictive Policing and Artificial Intelligence, 2021.

McGuire, Gez: There's no going back: how AI is transforming recruitment, Personnel Today 20 January, 2021.

Mehrabi, Ninareh/Morstatter, Fred/Saxena, Nripsuta/Lerman, Kristina/Galstyan, Aram: A Survey on Bias and Fairness in Machine Learning, 25 Jan 2022: https://arxiv.org/pdf/1908.09635.pdf.

Mendelsohn, Juliane: Die "normative Macht" der Plattformen - Gegenstand der zukünftigen Digitalregulierung?, MMR 2021, 857.

Moore, Phoebe V.: Data subjects, digital surveillance, AI and the future of work, European Parliamentary Research Service, 2020.

Moore, Phoebe V.: OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, European Agency for Safety and Health at Work Discussion Paper, 2019.

Moore, Phoebe V.: The Quantied Self in Precarity - Work, Technology and What Counts, 2018.

Moore, Phoebe V.: The Threat of Physical and Psychosocial Violence and Harassment in: Digitalized Work, ACTRAV Bureau for Workers' Activities, ILO, 2018.

Morrison, Sara: Dark patterns, the tricks websites use to make you say yes, explained- How design can manipulate and coerce you into doing what websites want, April 1, 2021, www.vox.com.

Moser, Christine/den Hond, Frank/Lindebaum, Dirk: Morality in the Age of Artificially Intelligent Algorithms, 7 Apr 2021: https://doi.org/10.5465/amle.2020.0287.

Moser, Christine/den Hond, Frank/Lindebaum, Dirk: What Humans Lose When We Let AI Decide - Why you should start worrying about artificial intelligence now, MIT Sloan, Feb 07, 2022: https://sloanreview.mit.edu/article/what-humans-lose-when-we-let-ai-decide/.

Möllers, Thomas M. J.: Juristische Methodenlehre, 3. Aufl. 2020.

Möslein, Florian: Die normative Kraft des Ethischen, RDi 2020, 34.

Müller-Glöge, Rudi/Preis, Ulrich/Schmidt, Ingrid (eds.): Erfurter Kommentar zum Arbeitsrecht, 22nd ed. 2022.

Müller-Hengstenberg, Claus Dieter/Kirn, Stefan: Haftung des Betreibers von autonomen Softwareagents, MMR 2021, 376.

Müller-Hengstenberg, Claus Dieter/Kirn, Stefan: Causality and responsibility for damage caused by autonomous smart systems, CR 2018, 682.

Münch v., Ingo/Kunig, Philip: Grundgesetz-Kommentar, 7th ed. 2021.

Narayanan, Arvind: How to recognise AI snake oil (set of slides): https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf.

National Institute of Standards and Technology: Towards a Standard for Identifying and Managing Bias in Artificial Intelligence, NIST Special Publication 1270, March 2022.

Negri, Sergio Avila: Robot as Legal Person: Electronic Personhood in Robotics and Artificial Intelligence, frontiers in Robots and AI, HYPOTHESIS AND THEORY: 10.3389/frobt.2021.789327.

Newlands, Gemma: Algorithmic Surveillance in the Gig Economy: The Organization of Work through Lefebvrian Conceived Space, Organization Studies 2020.

Nieto-Reyes, Alicia/Duquem Rafael/ Montañam José/Lage, Carmen: Classification of Alzheimer's Patients through Ubiquitous Computing, Sensors 2017, 1679.

Nink, David: Justiz und Algorithmen - Über die Schwächen menschlicher Entscheidungsfindung und die Möglichkeiten neuer Technologien in der Rechtsprechung, 2021.

Nowotny, Helga: In AI we trust: power, illusion and control of predictive algorithms, 2021.

Ockenfels-Martinez, Martha/ Boparai, Sukhdip Purewal: The Public Health Crisis Hidden in Amazon Warehouses, Oakland, CA. Human Impact Partners and Warehouse Workers Resource Center, 2021: https:// humanimpact.org/hipprojects/amazon.

O'Gieblin, Meghan: Prediction Engines Are Like Karma: You Get What You Stream, June 18, 2022. https:// www.wired.com/story/prediction-engines-are-like-karma-you-get-what-you-stream/.

Orwat, Carsten: Discrimination Risks through the Use of Algorithms, 2020.

Paal, Boris/Pauly, Daniel A.: Datenschutzgrundverordnung - Bundesdatenschutzgesetz. 3rd ed. 2021.

Pandey, Akshat /Caliskan, Aylin: Disparate Impact of Artificial Intelligence Bias in Ridehailing Economy's Price Discrimination Algorithms: https://dl.acm.org/doi/pdf/10.1145/3461702.3462561.

Papakonstantinou, Vagelis/de Hert, Paul: Refusing to award legal personality to AI: Why the European Parliament got it wrong - European Law Blog: https://europeanlawblog.eu/2020/.

Partnership on AI: Framework for Promoting Workforce Well-being in the AI-Integrated Workplace, 2020.

Papakonstantinou, Vagelis/de Hert, Paul: Refusing to award legal personality to AI: Why the European Parliament got it wrong – European Law Blog. 25 Nov. 2020. https://europeanlawblog.eu/2020/.

Paulus David/Matzke, Robin: Smart, Contracts und das BGB - Viel Lärm um nIchts?, ZfPW 2018, 431.

Pavlus, John: The Easy Questions That Stump Computers – What happens when you stack logs in a fireplace and drop a match? Some of the smartest machines have no idea, May 2, 2020. https://www. theatlantic.com/technology/archive/2020/05/computers-common-sense/611050/.

Pignot, Edouard: Who is pulling the strings in the platform economy? Accounting for the dark and unexpected sides of algorithmic control, 2021.

Prassl, Jeremias: The Concept of the Employer, 2015.

Preis, Ulrich: § 611 a BGB - Potenziale des Arbeitnehmerbegriffes, NZA 2018, 817.

Prince, Anya E.R./Schwarcz, Daniel: Proxy Discrimination in the Age of Artificial Intelligence and Big Data, Iowa Law Review 2020, 1257.

Purdy, Mark: How the Metaverse Could Change Work, April 05, 2022. https://hbr.org/2022/04/how-themetaverse-could-change-work.

Purdy, M., Zealley, J., Maseli, O.: The Risks of Using AI to Interpret Human Emotions. Harvard Business Review, 2019.

Purtova, Nadezhda: The Law of Everything. Broad Concept of Personal Data and Future of EU Data Protection Law, Law, Innovation and Technology 2018, 40.

Puschky, Ricarda: Federated Learning - eine datenschutzfreundliche Methode zum Training von KI-Modellen?, ZD-Aktuell 2022, 00019.

Quach, Katyanna: AI models still racist, even with more balanced training, 1 May 2022. https://www. theregister.com/2022/05/01/ai_models_racist/?tpcc= nleyeonai.

Racabi, Gali: What Can U.S. Labor Take from the Proposed E.U. Directive of Regulations of Platform Workers?, https://onlabor.org.

Rachlinski, Jeffrey John/Johnson, Sheri Lynn/Wistrich, Andrew J./Guthrie, Chris: Does Unconscious Racial Bias Affect Trial Judges?. Notre Dame Law Review, 2009, Vanderbilt Public Law Research Paper No. 09-11: https://ssrn.com/abstract=1374497.

Raghavan, Manish /Barocas, Solon/Kleinberg, Jon/Levy, Karen: Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices, 2020./.

Raue, Benjamin/von Ungern-Sternberg, Antje: Ethische und rechtliche Grundsätze der Datenverwendung, ZRP 2020, 49.

Reece, Nina: Workers say no to increased surveillance since COVID-19. https://www.tuc.org.uk/blogs/ workers-say-no-increased-surveillance-covid-19.

Reichold, Hermann: Betriebsverfassung als Sozialprivatrecht, 1995.

Reichwald, Julian/Pfisterer, Dennis: Autonomy and Intelligence in the Internet of Things, CR 2016, 208.

Reinartz, Oliver: Das Betriebsrätemodernisierungsgesetz, NZA-RR 2021, 457.

Renan Barzilay, Arianne: Data Analytics at Work: A View From Israel on Employee Privacy and Equality in the Age of Data-Driven Employment Management, Comparative Labor Law & Policy Journal 2019.

Rhue, Lauren: Affectively Mistaken? How Human Augmentation and Information Transparency Offset Algorithmic Failures in Emotion Recognition AI, November 22, 2019.

Richardi, Reinhard (ed.): Betriebsverfassungsgesetz mit Wahlordnung - Kommentar, 17th ed. 2022.

Riesenhuber, Karl: Arbeitnehmer(ähnlich) Schutz von Crowd-Dienstleistern?, ZfA 2021,5.

Riesenhuber, Karl: The Concept of the Employer, EuZA 2021, 133.

Risak, Martin/Dullinger, Thomas: The concept of 'worker' in EU law - Status quo and potential for change, ETUI Report 140, 2018.

Robinette, Paul/ Li, Wenchen/Allen, Robert/Howard, Ayanna M./Wagne , Alan R.: Overtrust of robots in emergency evacuation scenarios, in: Bartneck, Christoph (ed.), The Eleventh ACMIEEE International Conference on Human Robot Interaction, Piscataway, NJ, 2016, p. 101.

Röder, Gerhard/Gebert, Christian: Technologischer Wandel und Betriebsänderung - Bringen Industrie 4.0 und E-Mobilität den "Qualifizierungssozialplan"?, NZA 2017, 1289.

Rolfs, Christan/Giesen, Richard/Kreikebohm, Ralf/Meling, Miriam (eds): BeckOK Arbeitsrecht, 63rd Edition, 2022.

Rollberg, Christoph: Algorithmen in der Justiz - Rechtsfragen zum Einsatz von Legal Tech im Zivilprozess, 2020.

Roos, Philipp/Weitz, Caspar Alexander: High-Risk AI Systems in the Commission's Draft AI Regulation, MMR 2021, 844.

Rosenblatt, Alex/Stark, Luke: Algorithmic Labour and Information Asymmetries: A Case Study of Uber's Drivers, International Journal of Communication 10(2016), 3758.

Roßnagel, Alexander/Geminn, Christian L./Jandt, Silke/Richter, Philipp: Datenschutzrecht 2016 "Smart" genug für die Zukunft? Ubiquitous Computing und Big Data als Herausforderungen des Datenschutzrechts, 2016.

Roßnagel, Alexander/Geminn, Christian: Vertrauen in Anonymisierung, ZD 2021, 487.

Roßnagel, Alexander: Big Data - Small Privacy? - Conceptual Challenges for Data Protection Law, ZD 2013, 562.

Roßnagel, Alexander: Die Evaluation der Datenschutz-Grundverordnung, MMR 2020, 657.

Roßnagel, Alexander: Technology, Power and Law, MMR 2020, 222.

Rudkowski, Lena: "Predictive policing" at the workplace, NZA 2020, 72.

Ruschemeier, Hannah: 9th Speyerer Forum zur digitalen Lebenswelt: Regulierung Künstlicher Intelligenz in der Europäischen Union zwischen Recht und Ethik, NVwZ 2020, 446.

Russell, Stuart/Norvig, Peter: Artificial Intelligence - A Modern Approach, 4th ed., 2022.

Säcker, Franz Jürgen et al. (eds.): Münchener Kommentar zum BGB, 9th ed. 2022.

Sánchez-Monedero, Javier/Dencik, Lina/Edwards, Lilian: What Does It Mean to 'Solve' the Problem of Discrimination in Hiring?, 2019.

Sartor, Giovanni: The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, 2020.

Scheiber, Noam: How Uber Uses Psychological Tricks to Push Its Drivers' Buttons: https://www.nytimes.com.

Schild, Hans-Hermann: Der Beschäftigtendatenschutz auf dem Prüfstand des EuGH, ZD-Aktuell 2022, 01178.

Schleipfer, Stefan: Pseudonymity in various forms, ZD 2020, 284.

Schliesky, Utz: Digital Ethics and the Law, NJW 2019, 3692

Schmid, Alexander: Duty of "integrated product monitoring" for automated and networked systems, CR 2019, 141.

Schmidt, Alexander J.: Unionsrechtlicher Arbeitnehmerbegriff in der Plattformökonomie, NZA 2021, 1232.

Schneider, Thomas: Entscheidungsfindung, Entscheidungsformen: Formelwissen – Bauchgefühl – Künstliche Intelligenz, BC 2022, 225.

Schreiner, Paul: Co-determination right of section 87(1) no. 6 BetrVG is not subject to any materiality threshold, DB 2019, 554.

Schubert, Claudia: Crowdworker - Arbeitnehmer, arbeitnehmerähnliche Person oder Selbständiger - Zugleich eine Besprechung zum Urteil des LAG München v. 4.12.2019 - 8 Sa 146/19, RdA 2020, 248.

Schürmann, Kathrin: Datenschutz-Folgenabschätzung beim Einsatz Künstlicher Intelligenz, ZD 2022, 316.

Schulze, Marc-Oliver: Entwurf des Betriebsrätemodernisierungsgesetzes, ArbRAktuell 2021, 211.

Schwartmann, Rolf/Jaspers, Andreas/Thüsing, Gregor/Kugelmann, Dieter (eds.): DS-GVO/BDSG: Datenschutz-Grundverordnung, Bundesdatenschutzgesetz, 2. ed., 2020.

Schwarze, Roland: Die "arbeitsgleiche Durchführung" eines Tätigkeitsverhältnisses, RdA 2020, 38.

Sesing, Andreas/Tschech, Angela: AGG und KI-VO-Entwurf beim Einsatz von Künstlicher Intelligenz, MMR 2022, 24.

Sesing, Andreas: Grenzen systemischer Transparenz bei automatisierter Datenverarbeitung, MMR 2021, 288.

Silver, David/Singh, Satinder/Precup Doina/Sutton, Richard S.: Reward is enough, Artificial Intelligence, October 2021, 103535.

Simitis, Spiros/Hornung, Gerrit/Spiecker gen. Döhmann, Indra (ed.): Datenschutzrecht, 1st ed. 2019.

Singh, Namrata: Employee Surveillance Rises Alongside Work-from-Home Rates, December 18, 2021: https://icetonline.com/employee-surveillance-rises-alongside-work-from-home-rates/.

Sittard, Ulrich/Pant, Benjamin: Der Arbeitnehmerbegriff im Wandel aus Tradition und Moderne - zum Arbeitnehmerstatus eines Crowdworkers, jm 2021, 416.

Smith, Gary: High-tech redlining: AI is quietly upgrading institutional racism: How an outlawed form of institutionalized discrimination is being quietly upgraded for the 21st century: https://www.fastcompany.com/90269688/high-tech-redlining-ai-is-quietly-upgrading-institutional-racism.

Smuha, Nathalie/Ahmed-Rengers, Emma /Harkens, Adam/ Li, Wenlong/MacLaren, James/Pisellif, Ricardo/Yeung, Karen: How the EU can achieve trustworthy AI: A response to the European Commission's Proposal for an Artificial Intelligence Act, LEADS Law @University of Birmingham for a Legal, Ethical & Accountable Digial Siciety, 5 August 2021.

Söbbing, Thomas: Künstliche Intelligenz im HR-Recruiting-Prozess: Rechtliche Rahmenbedingungen und Möglichkeiten, InTer 2018, 64.

Söbbing, Thomas: Protection of algorithms - legal requirements and contractual design, ITRB 2019, 192

Solon, Olivia: "Big Tech Call Center Workers Face Pressure to Accept Home Surveillance." NBC News, August 8, 2021: https://www.nbcnews.com/tech/tech-news/big-tech-call-center-workers-face-pressure-accepthome-surveillance-n127622.

Sousa Antunes, Henrique: Civil Liability Applicable to Artificial Intelligence: A Preliminary Critique of the European Parliament Resolution of 2020 (December 5, 2020): https://ssrn.com/abstract=3743242.

Spencer, David/Cole, Matt /Joyce, Simon/ Whittaker, Xanthe/Stuart Mark: Digital Automation and the Future of Work, European Parliamentary Research Service, 2021.

Spiecker gen. Döhmann, Indra: Zur Zukunft systemischer Digitalisierung - Erste Gedanken zur Haftungs- und Verantwortungszuschreibung bei informationstechnischen Systemen, CR 2016, 698.

Spindler, Gerald/Schuster, Fabian: Recht der elektronischen Medien, 4th edition 2019.

Srinivasan, Ramya/Chander, Ajay: Biases in AI Systems - A Survey for Practitioners, acmqueue 2021, 47.

Stanford University Human-Centered Artificial Intelligence, Artificial Intelligence Index Report 2021.

Staudinger: Kommentar zum Bürgerlichen Gesetzbuch, Neubearbeitung 2022.

Steege, Hans: Algorithm-based discrimination through the use of artificial intelligence, MMR 2019, 715.

Stephen Cave/Kanta Dihal: The Whiteness of AI, Philosophy & Technology 2020, 685.

Straker, Christian /Niehoff, Maurice: ABIDA-Fokusgruppe - Diskriminierung durch Algorithmen und KI im eRecruiting, ZD-Aktuell 2018, 06252.

Sühr, Tom/Biega, Asia J./Zehlike, Meike/Gummadi, Krishna P./Chakraborty, Abhijnan: Two-Sided Fairness for Repeated Matchings in Two-Sided Markets: A Case Study of a Ride-Hailing Platform, Applied Data Science Track Paper 2019.

Sullivan, Charles A.: Employing AI, Seton Hall Public Law Research Paper 2018.

Sullivan, Diane: Interviews Don't Work so Why Not be a Hiring Nihilist? Because it's all a lottery anyway, Oct 28, 2020.

Sunstein, Cass R.: The Ethics of Nudging, Yale Journal on Regulation 32 (2015), 413.

Sweeney, Latanya: Simple Demographics Often Identify People Uniquely, Carnegie Mellon University, Data Privacy Working Paper 3.

Tabarrini, Camilla: Understanding the Big Mind, EuCML 2020, 135.

Takshi, Sahar: Unexpected Inequality: Disparate-Impact From Artificial Intelligence in Healthcare Decisions, Journal of Law and Health, 2021, 215.

Teubner, Gunther: Digital Legal Subjects? On the private law status of autonomous software agents, AcP 2018, 155.

The Norwegian Data Protection Authority: Artificial Intelligence and Privacy, Report January 2018.

Thieltges, Andree: Big Data, Machine Learning und Künstliche Intelligenz - Neue Herausforderungen für die betriebliche Mitbestimmung, ZfP 2020, 3.

Thieltges, Andree: Machine Learning Anwendungen in der Betrieblichen Praxis - Praktische Empfehlungen zur betrieblichen Mitbestimmung, 2020.

Thüsing, Gregor/Hütter-Brungs, Gisela: Crowdworking: Lenkung statt Weisung - Was macht den Arbeitnehmer zum Arbeitnehmer?, NZA-RR 2021, 231.

Thüsing, Gregor: Gleicher Lohn für gleichwertige Arbeit, NZA 2000, 570.

Tischbirek, Alexander: Ermessensdirigierende KI, Zeitschrift für Digitalisierung und Recht (ZfDR) 2021, 307.

Todoli-Signes, Adrián: Making algorithms safe for workers: occupational risks associated with work managed by artificial intelligence: https://doi.org/10.1177/10242589211035040.

Unger, Oliver: Grundfragen eines neuen europäischen Rechtsrahmens für KI, ZRP 2020, 234.

Valta, Matthias/Vasel, Johann Justus: Commission Proposal for a Regulation on Artificial Intelligence, ZRP 2021, 142.

Vamplew, Peter et al. : Scalar reward is not enough: A response to Silver,Singh, Precup and Sutton (2021): https://arxiv.org/pdf/2112.15422.pdf.

Vásquez, Sheila: Privacy by design: a joint challenge of IT, engineers and managers to effectively implement data protection law, DSRITB 2021, 149.

Veale, Michael/Zuiderveen Borgesius, Frederik: Demystifying the Draft EU Artificial Intelligence Act, CRi 2021, 97.

Veith, Charlotte: Künstliche Intelligenz, Haftung und Kartellrecht - Zivilrechtliche Verantwortlichkeit beim Einsatz von Künstlicher Intelligenz und Implikationen für das Kartellrecht, 2021.

Vincent, James, "Amazon Turns Warehouse Tasks into Video Games to Make Work 'Fun.'" The Verge, May 22, 2019. https://www.theverge.com/2019/5/22/18635272/amazon-warehouse-working-conditions-gamification-video-games.

Volkova, Svitlana/ Bachrach, Yoram: On Predicting Sociodemographic Traits and Emotions from Communications in Social Networks, in: Cyberpsychology, Behavior, and Social Networking, 2015, 726.

Wachter, Sandra: The Theory of Artificial Immutability: Protecting Algorithmic Groups under Anti-Discrimination Law (February 15, 2022). Tulane Law Review, Forthcoming. https://ssrn.com/abstract=4099100.

Walsh, Toby: Was Künstliche Intelligenz wirklich besonders macht, FAZ 7.6.2022, S. 20.

Waas, Bernd: Verbesserung der Arbeitsbedingungen von Plattformbeschäftigten, ZRP 2022 105.

Waas; Bernd/Heerma van Voss, Guus (eds.): Restatement of Labour Law in Europe, vol. 1, The Concept of Employee, 2017.

Wachter, Sandra/Mittelstadt, Brent/Russell, Chris: Bias Preservation in Machine Learning: The Legality of Fairness Metrics Under EU Non-Discrimination Law, https://ssrn.com/abstract=3792772.

Wachter, Sandra/Mittelstadt, Brent/Russell, Chris: Why Fairness cannot be automated Bridging the Gap between EU Non-Discrimination Law and AI, 2020.

Wachter, Sandra/Mittelstadt, Brent: A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI, Columbia Business Law Review 2019: https://ssrn.com/abstract=3248829.

Wagner, Gerhard: Haftung für Künstliche Intelligenz - Eine Gesetzesinitiative des Europäischen Parlaments, ZEuP 2021, 545.

Wagner, Gerhard: Verantwortlichkeit im Zeichen digitaler Techniken, VersR 2020, 717.

Wakefield, Jane: AI: Ghost workers demand to be seen and heard, 28 March 2021: www.bbc.com.

Wallace, Elizabeth: How Will AI Nudging Affect Our Privacy?, ODSC, March 18, 2019.

Wang, Yilun/Kosinski, Micha: Deep neural networks are more accurate than humans at detecting sexual orientation from facial images; in: Journal of Personality and Social Psychology, 2018, 246.

Wank, Rolf: Der Arbeitnehmerbegriff im neuen § 611 a BGB, AuR 2017, 140.

Weber, Klaus et al. (eds.): Rechtswörterbuch, 27th Edition 2021.

Wei, Yanhao/Yildirim, Pinar/Van den Bulte, Christophe: Credit Scoring with Social Network Data, Marketing Science 2016, 234.

Weichert, Thilo: Big Data und Datenschutz - Chancen und Risiken einer neuen Form der Datenanalyse, ZD 2013, 251.

Weichert, Thilo: Datenschutz-Grundverordnung - arbeitsrechtlich spezifiziert, NZA 2020, 1597.

Werkmeister, Christoph/Brandt, Elena: Datenschutzrechtliche Herausforderungen für Big Data, CR 2016, 233.

Wiebe, Gerhard: Produktsicherheitsrechtliche Betrachtung des Vorschlags für eime KI-Verordnung, BB 2022, 899.

Wiese, Günther/Kreutz, Peter/Oetker, Hartmut/Raab, Thomas/Weber, Christoph/Franzen, Martin/Gutzeit, Martin/Jacobs, Matthias: Gemeinschaftskommentar zum Betriebsverfassungsgesetz (GK-BetrVG), 12th ed. 2021.

Wiggers, Kyle: AI experts warn Facebook's anti-bias tool is 'completely insufficient', March 31, 2021: https://venturebeat.com

Willemsen, Josef/Mehrens, Christian: Arbeitnehmerüberlassung versus Dienstleistung, NZA 2019, 1473.

Winter, Christian/Battis, Verena/Halvani, Orena: Challenges for the Anonymisation of Data, ZD 2019, 489.

Wisskirchen, Gerlind/Haupt, Jan: Crowdworker: Arbeitnehmer oder Selbstständiger?, RdA 2021, 355.

Wu, Xiaolin; Zhang, Xi: Automated inference on criminality using face images, 2016.

Wybitul, Tim: Der neue Beschäftigtendatenschutz nach § 26 BDSG und Art. 88 DSGVO, NZA 2017, 413.

Wynsberghe, Aimee van: Artificial Intelligence: From ethics to policy, 2020.

Xenidis, Raphaële/Senden, Linda: EU non-discrimination law in the era of artificial intelligence: Mapping the challenges of algorithmic discrimination' in Bernitz a.o. (eds.), General Principles of EU law and the EU Digital Order, 2020, 151. SSRN: https://ssrn.com/abstract=3529524

Yeung, Karen: A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework, Council of Europe, DGI(2019)05, Committee of experts on human rights dimensions of automated data processing and different forms of artificial intelligence (MSI-AUT), 2018.

Yi, Zeng/Kang, Sun/Enmeng, Lu: Declaration on the ethics of brain-computer interfaces and augment intelligence, AI and Ethics 2021, 209.

Yuste, Rafael/Genser, Jared/Herrmann, Stephanie: It's Time for Neuro-Rights-New Human Rights for the Age of Neurotechnology: Horizons, 2021, 154.

Zech, Herbert: Entscheidungen digitaler autonomer Systeme: Empfehlen sich Regelungen zu Verantwortung und Haftung?, Gutachten A zum 73. Deutschen Juristentag, 2022.

Zech, Herbert: Künstliche Intelligenz und Haftungsfragen, ZfPW 2019, 198.

Zehlike, Meike/Hacker, Philipp/Wiedemann, Emil: Matching code and law: achieving algorithmic fairness with optimal, Data Mining and Knowlewdge Discovery, 2020, 163.

Zekos, Georgios I.: Political, Economic and Legal Effects of Artificial Intelligence - Governance, Digital Economy and Society, 2022, p. 483.

Zuiderveen Borgesius, Frederik J.: Strengthening legal protection against discrimination by algorithms and artificial intelligence, The International Journal of Human Rights 2020, 1572.

Zuiderveen Borgesius, Frederik J.: Discrimination, artificial intelligence, and algorithmic decision-making, Council of Europe, 2018

Zuiderveen Borgesius, Frederik J.: Singling out people without knowing their names - Behavioural targeting, pseudonymous data, and the new Data Protection Regulation, Computer Law & Security Review 2016, 256.

The authors themselves are responsible for the content.

This and all other HSI Working Papers can be found at
www.hugo-sinzheimer-institut.de/veroeffentlichungen

Hugo Sinzheimer Institut für Arbeits- und Sozialrecht (HSI)
der Hans-Böckler-Stiftung
Wilhelm-Leuschner-Straße 79
60329 Frankfurt am Main

www.hugo-sinzheimer-institut.de

Impressum: https://www.boeckler.de/de/impressum-2713.htm