Prof. Dr. Dr. Frauke Rostalski
Dr. Erik Weiss

Conformity Assessment in the Artificial Intelligence Act

By drafting the proposal for regulating Artificial Intelligence[1] (Artificial Intelligence Act) the European Commission introduced the first holistic legal framework for AI systems worldwide.[2] Especially, at the heart of the proposal – the regulation of so-called "high-risk" AI systems as defined in Art. 6 of the draft AI regulation – conformity assessments play a central role (cf. Art. 43 of the draft AI regulation). Their drafting has extensive consequences for the level of protection awarded to consumers. This report examines the systematics and the requirements of the conformity assessment in the AI regulation draft. Firstly, it will shine a light on the objectives of such procedures and their possible consequences for competition. Moreover, the question of liability of conformity assessment bodies will be addressed. Subsequently, this report will turn to the specific challenges for conformity assessments that accompany AI-systems. In addition to a clarification of term, a number of particularities of such applications arise at this point which must be accurately represented in a normative examination. This relates to both empirical factors which can be traced back to the nature of application of AI and challenges arising due to insufficient regulatory standards. Based on this, a number of general requirements regarding conformity assessment procedures of AI-applications are established. In the following, the "risk-based" category approach of the draft AI Regulation will be examined, questioning to what extent this approach affects the rights of consumers in the context of conformity assessments. For this, the report draws on the distinction made in the regulation proposal between prohibited AI-Systems und those with high, low and minimal risk. An extensive analysis of the conformity assessment system as envisaged in the draft AI regulation will be conducted. Finally, the insights of this analysis will be converted into guidelines to improve consumer protection within the conformity assessment systems of the draft AI Regulation.

Thus, the report is structured as follows:

[1] "Artificial Intelligence" is used as a term to indicate that it is impossible to equate it to human intelligence – the term intelligence in the context of "Artificial Intelligence" must be understood in the technical sense: an overview of this is provided in *Myers*, Psychologie, 3. Aufl. 2014, p. 400 et seqq.; for the use of the term intelligence in AI-research: *Kirste/Schürholz,* in Wittpahl, Künstliche Intelligenz, 2018, p. 21 (21). Cf. also *Rostalski/Weiss*, ZfDR 2021, 329 Fn. 2.

[2] Cf. Regarding the regulation proposal *Rostalski/Weiss*, ZfDR 2021, 329 et seqq.

Prof. Dr. Dr. Frauke Rostalski
Dr. Erik Weiss

Prof. Dr. Dr. Frauke Rostalski
Dr. Erik Weiss

A. Introduction

In the following, the general functions and implications of conformity assessments will be outlined, classified and evaluated (I.), and the specific challenges of corresponding examination procedures regarding AI-systems will be determined (II.). As a synthesis of both analyses general requirements for the "design" of conformity assessments of AI systems will be identified (III.)

I. General functions and implications of conformity assessments

A conformity test assesses whether a product complies with previously determined requirements, such as norms and standards.[3] If this is the case, conformity exists. Within a conformity assessment, different stages or "levels" can be distinguished according to the body that conducts said assessment. This may encompass the assessment being conducted by the manufacturer of the product or the system ("first party" conformity assessment), or an assessment by the buyer of the product ("second party" conformity assessment). Finally, the conformity assessment may be conducted by an independent third party ("third party" conformity assessment).[4] In general, the aim of such assessments is to ensure the quality of the product, as specified by legal and ethical requirements.

1. Classifying and distinguishing different conformity assessment procedures

Different forms of conformity assessments[5] may be considered, differing in terms of the effort required, the person or body conducting the assessment and whether it is done on a voluntary basis. Assessment procedures regarding products, services, as well as systems, processes and persons are conceivable. In this context, a voluntary commitment aims for specific organisations and companies to declare that they will uphold the rules that apply to them. Independent third parties are not included in this particular form of conformity assessment. As it is voluntary, it is not subject to government regulation. By itself, this voluntary commitment has no binding legal effect – such an effect does arise, however, from the legal standards establishing specific requirement that are applicable to the respective organisation or company. A seal of quality is used to identify the products properties for external users, which may relate to the products quality, but also specific safety features. However, there is no legal procedure for the conception and issuance of quality seals. Therefore, in

[3] *Heesen/Müller-Quade/Wrobel et al.*, in: Whitpaper „Zertifizierung von KI-Systemen der Plattform 'Lernende Systeme'", p. 7 (available at: https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG1_3_Whitepaper_Zertifizierung_KI_Systemen.pdf, Last accessed on 05.08.2022).
[4] For differentiation see *Heesen/Müller-Quade/Wrobel et al.*, in: Whitpaper „Zertifizierung von KI-Systemen der Plattform 'Lernende Systeme'", p. 7 (available at: https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG1_3_Whitepaper_Zertifizierung_KI_Systemen.pdf, last accessed on: 05.08.2022).
[5] Cf. for the following differentiation *Heesen/Müller-Quade/Wrobel et al.*, in: Whitpaper „Zertifizierung von KI-Systemen der Plattform ‚Lernende Systeme"', p. 6 et seq. (available at: https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG1_3_Whitepaper_Zertifizierung_KI_Systemen.pdf, last accessed on: 05.08.2022).

practice there is often an association of several manufacturers within an industry who define the criteria of a quality seal for themselves. Whether the requirements of the quality seal are met is usually checked by the manufacturer themselves. Thus, no third party or body is involved in this kind of conformity assessment.

The situation is different regarding product certification procedures that may be applicable to AI systems. The certificate confirms that specific nationally and internationally defined norms, standards and guidelines of legal and ethical nature are met, at least for a specific period of time. Certification may rely either on a voluntary or a legal basis. In practice, the former is more common. In this context the conformity assessment is conducted by an independent as well as sufficiently qualified third party – the so-called certification body.

However, conformity assessment procedures regarding AI systems are currently almost non-existent due to the general lack of sufficiently concrete, generally recognized norms and standards.[6] The same is true for approval procedures providing legally required assessments of products or systems with regard to their compatibility with applicable (national or European) legal standards. These approval examinations are carried out by the relevant federal office or they are delegated to other institution by the federal authority.

2. Importance for the protection of consumer rights

Conformity assessments play a prominent role in relation to consumer rights. This applies generally, but particularly regarding AI systems which are used in daily life by consumers. Depending on where and how they are used, i.e. the area and the modality of application, technological products based on AI-technology can display a high level of interference, or "criticality" regarding the legal sphere of individuals involved. Simultaneously, the independent verification of compliance with safety standards, especially verification by the consumer themselves, proves virtually impracticable, as the high complexity of the systems poses an obstacle. At first glance, and against the relevance of conformity assessments, it could be argued that they are not acts of regulation: From a legal standpoint, the conformity assessment merely ensures that the level of care that is demanded by the state either way is complied with by the people responsible for the product. As laws, as well as technical norms and standards, apply anyways and as they are enforceable by the state using coercion (if necessary), the question arises whether conformity assessments can add anything of value. The answer to that question is not least important, because obligatory conformity assessments shift the burden onto manufacturers of AI products. This must be justified – especially with regard to the protection of competitive interests.

[6] First research projects in this area already exist, such as the "Certified AI" Project, which we are co-leading (for more details, see the website https://www.zertifizierte-ki.de/, last accessed on: 05.08.2022) or the project „ExamAI – KI-Testing & Auditing" (cf. the internet presence: https://testing-ai.gi.de/, last accessed on: 05.08.2022).

Conformity assessments have already proven their worth in other product areas.[7] The reason for this lies in the fact, that certain products themselves pose such a significant risk to consumer rights that the law itself cannot provide a sufficient guarantee for the compliance with safety standards. Where the realisation of a specific risk leads to significant negative consequences, additional guarantees for protection are required to ensure that the rights of potentially affected individuals are sufficiently safeguarded. At this point, conformity assessments provide *one essential* element to ensure legal compliance.[8] They are supposed to provide the consumer using the product in question with additional assurances, that said product complies with the applicable regulations regarding manufacturing or usage, that were set in place for consumer protection. Conformity assessments thus create not only certainty regarding the validity of legal standards, but also that those standards were complied with. This level of certainty is further increased when a third person or body is interposed between consumer and manufacturer in the context of the conformity assessment. While mere regulation through laws, norms and standards relies on the legal conformity of the respective manufacturer, this mechanism of legal validity and legal realization is further stabilized in the context of conformity testing by involving a third party. This person or body provides additional protection for the consumer by scrutinizing the execution of the law applicable to the specific product and thereby helps strengthen the validity of the law itself.

For the rights of consumers, this additional mechanism of protection provided through conformity assessments of AI-systems plays *the central role*. However, it should not be overlooked that the consumer's sphere of interest has another dimension, which is represented only in the context of conformity assessments. Namely, conformity assessments review standards going beyond the applicable law. Until now, these standards only reflect the moral values of the involved individuals. In this respect, conformity assessments to a certain extent assume a regulatory role by reviewing features of AI systems which are not legally required, but which reflect a specific moral norm shared by the majority of our society. Conformity assessments therefore guarantee a standard that goes beyond the applicable law. This should be considered an improvement of the consumer's legal position as well, because it aligns with generally held values. Incidentally, this may have a spill over effect on the law: As a consolidated standard of morality these standards may be transferred into valid law.

Against this backdrop, it can be summarized that conformity assessment procedures contribute to the compliance with the law through additional guaranteed protection as well as

[7] See relevant practice examples in more detail under A. I. 4.

[8] In addition to voluntary self-commitment, which is not regulated by the state, legal conformity may be achieved through legally prescribed certification, an approval procedure by a federal office or body, or through enacting directives or prohibitions cf. *Heesen/Müller-Quade/Wrobel et al.*, in: Whitpaper „Zertifizierung von KI-Systemen der Plattform 'Lernende Systeme"', p. 7 et seq (available at: https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG1_3_Whitepaper_Zertifizierung_KI_Systemen.pdf, last accessed on: 05.08.2022).

to the protection of the user through extensive ethical requirements. Especially concerning particularly intrusive applications, only conformity assessments can guarantee the required level of safety which renders the use of such applications socially acceptable. Thus, conformity assessments thereby create something more than mere trust. The aspect of guaranteeing certainty of expectation, along with the additional safeguarding of legal conformity and meeting any criteria that goes beyond this, does not have priority. Instead at its core lies the guarantee of conformity with specific standards, which is decisively facilitated through the critical third-party review which forms the basis of trust. This results in an improvement of protection of consumer rights and interests – and even in a level of protection without which the use of specific AI applications would not be acceptable in the first place.

3. Competitive disadvantages for small and medium-sized companies?

Conformity assessments – similar to general regulation and measures of implementation – are suspected to endanger free competition, especially if the individuals subjected to the assessment, are less able to compete.[9] This is based on the central notion, that the implementation of assessment procedures involve additional efforts for the manufacturer or other responsible persons regarding the norm-conformity and the compliance with other ethical standards. It binds resources, which – according to critics – could be used elsewhere, to be able to meet the challenges of free competition. The burden falls particularly on small and medium-sized companies, for which the additional (also: economic) strain may have grave effects, compared to large companies, which have the financial means as well as the required staff to run conformity assessments without considerable effort.

However, it may make a difference whether the conformity assessment is obligatory or voluntary. Whilst approval procedures are usually state-mandated,[10] a certification procedure on a voluntary basis may be envisaged.[11] The question of whether the conformity assessment lies within the margin of discretion of the manufacturer of the AI application, is dependent on their criticality. The higher the estimated risk emanating from the AI system, the higher the likelihood for a mandatory conformity assessment.[12] Meanwhile, even voluntary procedures to assess legal conformity may constitute an additional burden for

[9] Cf. *Heesen/Müller-Quade/Wrobel et al.*, in: Whitpaper „Zertifizierung von KI-Systemen der Plattform 'Lernende Systeme'", p. 6 (available at: https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG1_3_Whitepaper_Zertifizierung_KI_Systemen.pdf, last accessed on: 05.08.2022).
[10] *Heesen/Müller-Quade/Wrobel et al.*, in: Whitpaper „Zertifizierung von KI-Systemen der Plattform 'Lernende Systeme'", p. 8 (available at: https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG1_3_Whitepaper_Zertifizierung_KI_Systemen.pdf, last accessed on: 05.08.2022).
[11] *Heesen/Müller-Quade/Wrobel et al.*, in: Whitpaper „Zertifizierung von KI-Systemen der Plattform "Lernende Systeme'", p. 7 (available at: https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG1_3_Whitepaper_Zertifizierung_KI_Systemen.pdf, last accessed on: 05.08.2022).
[12] Cf. A proposal of a criticality-pyramid of the Data Ethic Commission, Resort of the Data Ethic Commission of the Federal Ministry of Interior and Community, 2019, p. 177 et seqq. (available at: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=7, last accessed on 05.08.2022).

those who are presented with that offer. Voluntary procedures may acquire coercive force, if a number of market competitors use this offer. This may lead to a corresponding expectation on the consumer's side, with the result that the lack of the (voluntary) conformity assessment is considered a relevant flaw, which disadvantages the respective manufacturer.[13]

Therefore, both mandatory as well as voluntary conformity assessments must be scrutinised regarding the question whether they take the legal position of small and medium-sized companies sufficiently into account. For the latter, the respective assessment procedures entail a more significant strain than for large companies. This can be adequately taken into account through various measures. Initially, one might want to think of adapting the assessment criteria to the level of performance of the respective company. This is especially relevant regarding the expenditure of resources the company has to make in order to conduct the assessment. However, it must be ensured, that it is not accompanied by a relevant loss of safety guarantees. From a legal standpoint, the size of the company by itself should not determine to what extent the compatibility with applicable law and potentially ethical standards is scrutinized in the context of compatibility assessments. To account for this, a solution to this conflict arising for small and medium-sized companies may entail state support regarding the resources required for the assessment procedure. Such support may be envisaged by allowing professional consultation in relation to possible assessment procedures, as well as granting financial means to allow an expansion of the assessment infrastructure. In any case, this allows the competitive disadvantage by virtue of extensive conformity assessments that burdens smaller and medium-sized companies compared to bigger ones to be balanced out.

4. Selected practical examples of conformity assessment procedures

Conformity assessment procedures play a practical and quite important role regarding numerous other products. One may think of e.g. Section 4 of the Ninth Ordinance to the Product Safety Act (Machinery Ordinance) which sets down the rules for conformity assessment procedures for machines. Similarly, Art. 52 of the Medical Device Regulation as well as the Annex IX-XI provide standards for conformity assessment procedures regarding the products listed in this regulation. Generally we can conclude from those practical regulation examples, that the scope of the respective requirements depends on the criticality of the particular products. It further plays a role in the eyes of the legislator to what extent the manufacturer themselves made efforts regarding the conformity of their product. Depending on whether e.g., existing harmonised technical standards were complied with, more

[13] According to *Campos Nave/Zeller*, BB (2012), 131 (134), companies that lack compliance assessments are at a competitive disadvantage, as it serves to induce the consumer's trust and sets it apart from the competition. Also cf. *Heesen/Müller-Quade/Wrobel et al.*, in: Whitpaper „Zertifizierung von KI-Systemen der Plattform "Lernende Systeme"", p. 35 (available at: https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG1_3_Whitepaper_Zertifizierung_KI_Systemen.pdf, last accessed on: 05.08.2022).

or less extensive duties in the context of conformity assessment procedures may arise for the person responsible.[14]

5. Liability of conformity assessment authorities

For the interests of all individuals involved – not least the consumers – it plays a decisive role, whether and to what extent conformity assessment authorities are liable when a mistake is made. Parallels can be found in German law, for instance regarding the question whether liability arises for the assessor for breaching their duty to duly examine a car in the context of the periodic technical inspection according to Section 29 of the German Road Traffic Licensing (StVZO).[15] In the past, the OLG (high court) Hamm held a public office liability claim against the state.[16] This requires a case of abuse of office. Only when this is indeed the case, the public employer is fully responsible for his conduct with respect to the affected individual.[17] However, the requirements for an assumption of abuse of office in German law are difficult to meet. It is only assumed if the acting public official intentionally harms another person in a way that is contrary to the public notion of morality ("Gute Sitten"), and thereby fulfils the requirements of Section 826 of the German Civil Code (BGB). Furthermore, an abuse of office may also be considered in case of negligent conduct; but this is dependent on the specifics of the individual case.[18]

The aforementioned general principles can be applied to conformity assessments relating to AI systems. Assessing the conformity of AI applications with – as possibly specified by harmonised norms and standards – requirements, functions as a further protection for people who come in contact with them. If the conformity assessments are conducted by a holder of sovereign power or someone whose activity can be attributed to a public body (e.g., because they are authorized by issuance of an administrative act), liability for abuse of office may be considered. The duty to refrain from abuse of one's office is owed to any individual who may be harmed by this abuse. Thus, in line with what has been stated above, conformity tests serve, among other things, to protect persons in general from injury including unlawful discrimination. It therefore seems obvious to apply the general rules of liability used in cases of abuse of office, insofar as a sufficient connection with an official body can be established. In that case, the liability of the employer who was responsible for the public official comes alongside a possible liability of other individuals, such as the manufacturer or operator of the AI system in question. Against this background, a difficulty may lie in clarifying whether the expert is performing the conformity assessment as a task of sovereign authority. This can principally be considered, if the individuals who based on their particular professional expertise in their respective technical field, support necessary surveillance measures of public authorities or carry out preparatory tasks that are required

[14] Cf. C. I. 3. c. for reasons for the development of different duties in the context of conformity assessments.
[15] Equivalent to the MOT (Ministry of Transport) test.
[16] OLG Hamm, DAR 2010, 138 et seqq.
[17] BGH NJW 1973, 458; NJW 2004, 3484 et seq.
[18] BGH NJW 1973, 458.

for authoritative decision-making.[19] Such capacity has been assumed in jurisprudence before, e.g. for an officially recognised expert for motor traffic regarding the approval of an operating license as per Section 21 StVZO and during a periodic technical inspection pursuant to Section 29 StVZO. Moreover, a MOT-expert, who carried out a preliminary inspection of systems requiring monitoring pursuant to Section 24 GewO, was classified accordingly. Regarding AI conformity assessments, it will ultimately depend on their design and the legal consequences that are attached to the assessment.

II. Challenges of conformity assessments of AI systems

AI systems pose challenges for conformity assessments that are yet unknown in other product areas, at least challenges that have not yet arisen in that shape. Respective assessment procedures must take these particularities into account.

1. Definition: AI system

In order to start approaching this difficulty, it is necessary to start with a clarification of the concept of the AI system, as the central term of this analysis.[20] Depending on individual preconceptions, the term AI is assigned to vastly different technologies. A widely accepted definition does not exist – not even in the field of computer science.[21] Machine learning (ML) techniques are often being referred to, when talking about AI systems.[22] Applications based on this technique learn to recognise patterns and rules on so-called training data, generalise them and finally apply them to unknown issues.[23] The European Commission's draft, on the other hand, is based on a significantly broader understanding of the term "AI systems". Art. 3 (1) of the draft AI Regulation defines an "artificial intelligence system" (AI system) as a software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with. Annex I lists the following techniques and systems: a) machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; c) Statistical approaches, Bayesian estimation, search and optimization methods. Thus, applications based on ML techniques represent only *one* group of AI systems that fall within the substantive scope of the draft. Due to the broad scope of the techniques listed in Annex I b) and c),

[19] *Dörr*, in: BeckOGK-BGB, Status: 1.5.2022, § 839 Rn. 55.
[20] Further explanations are based on *Rostalski/Weiss*, ZfDR 2021, 329 (331 et seq.).
[21] *Geminn*, ZD 2021, 354 (354); *Hacker*, NJW 2020, 2142 (2142); *Herberger*, NJW 2018, 2825 (2826).
[22] *Hacker*, NJW 2020, 2142 (2142); *v. Westphalen*, ZIP 2019, 889 (889); *Zech*, ZfPW 2019, 198 (200).
[23] *Leupold/Wiesner*, in: Leupold/Wiebe/Glossner, IT-Recht, 4. Aufl. 2021, Teil 9.6.4. Rn. 2 m.w.N

deterministic application as well as common expert systems[24] may be classified as AI systems as per Art. 3 (1) of the draft AI Regulation.[25] Against this background, it has been proposed (albeit controversially) to equate the term "AI system" as used in the regulation with "software" in order to simplify it.[26] The concerns voiced about the breadth of this definition do not hold water regarding the fear that the regulatory proposal would extend to any conceivable software. Insofar as the draft makes binding specifications regarding individual AI systems – i.e. with regard to prohibited AI systems, high-risk AI systems and low-risk AI systems[27] – the AI systems that are covered are specified in more detail through their context of use and are thereby limited. Moreover, the provided definition is merely a provisional one. The draft is based on a dynamic conception: Pursuant to Art. 4, 73 of the draft AI Regulation, the EU Commission is authorized to issue delegation acts to update Annex I and thus adapt the regulation according to newest developments. Regardless of whether one adopts a broad or restrictive understanding of the term, it remains crucial to consider the importance of a definition and its associated consequences. The legal evaluation and classification of applications is fundamentally dependent on the adopted definition.[28] Furthermore, it is difficult to place and to critically assess standpoints regarding AI systems, without the disclosure of the underlying conceptual understanding. This hinders discourse. Therefore, the tendency within the political debate to use the term "AI" as undefined cipher for a wide variety of technologies is sparsely conducive.[29]

The Data Ethics Commission has also advocated for a specification of the term[30] and defined AI in its final report as a collective term for those technologies and their applications that use digital methods based on potentially very large and heterogenous data sets in a complex mechanical process that mimics human intelligence to determine results that may potentially be applied automatically.[31] Simultaneously, it has called for extensive regulation that covers "all types of algorithmic systems"[32] as a generic term. The definition for AI systems as proposed in the draft AI regulation takes this into account as far as the breadth of the definition also covers "simple" control systems that are based on algorithms (=

[24] Cf. for the term *Styczynski/Rudion/Naumann*, Einführung in Expertensysteme, 2017, p. 10 et seqq. and *F. Puppe*, Einführung in Expertensysteme, 1991.
[25] *Spindler*, CR 2021, 361 (363), who points out, that this is a much wider definition than the one used by the *High Level Expert Group on AI*.
[26] As per *Bomhard/Merkle*, RDi 2021, 276 (277).
[27] Cf. B. for details on this distinction.
[28] *Steege*, SVR 2021, 1 (2).
[29] *Herberger*, NJW 2018, 2825 (2826).
[30] *Data Ethics Commission,* Empfehlungen der Datenethikkommission für die Strategie Künstliche Intelligenz der Bundesregierung vom 9.10.2018, p. 1 (available at: https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/it-digitalpolitik/datenethikkommission/empfehlungen-datenethikkommission.pdf?__blob=publicationFile&v=2, last accessed on: 05.08.2022).
[31] *Data Ethics Commission,* Gutachten der Datenethikkommission der Bundesregierung, 2019, p. 34, available at: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=7, last accessed on 05.08.2022).
[32] *Data Ethics Commission,* Gutachten der Datenethikkommission der Bundesregierung, 2019, p. 34 (available at: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=7, last accessed on: 05.08.2022).

rules)[33] that are 'manually' set by experts. The expansive scope of application that includes all types of algorithmic systems should be welcomed, since not only the use of AI systems in the narrow sense, but other algorithmic systems as well, raise fundamental ethical and legal questions.[34] Thus, the broad definition as per Art. 3 (1) in conjunction with Annex I of the draft AI Regulation is used here as well. It is very likely that this definition is legally binding in future.

2. AI-specific risks for consumer rights

AI systems have become the focus of national and European regulatory efforts in recent years – and with good reason. The reason for this is the special functioning of the respective technology which entails certain risks for consumer rights. This requires appropriate regulatory responses which must also be taken into account in the context of conformity assessments. The aforementioned special risks caused by AI systems are initially apparent in the area of data protection. Particularly machine learning methods generally rely on a large set of data.[35] This raises questions about legally permissible use of data that must be answered by society. It is especially important to prevent individuals from being degraded to mere objects of technical operations in the course of data processing.[36] Beyond that, AI systems are associated with a specific risk of legally prohibited discrimination against individuals or entire groups.[37] The use of new technologies cannot be allowed to entrench existing social marginalization or create new ones. One can already think of, e.g., AI systems based on ML-techniques to be used in the recruitment of job applicants. It is certainly conceivable, that as a result of the machine learning from previous recruitment cycles it assesses the quality of the application according to the applicant's gender.[38] This constitutes an act of impermissible discrimination as long as it is not based on an objective reason such as the specific nature of the job[39], and must thus be countered. Avoiding such illegal discrimination as well as the assurance of adequate data protection proves to be difficult in the context of AI systems (especially those based primarily on ML techniques), due to another special feature of this technology. What is referred to here, is the considerable lack of transparency regarding the operation of the respective applications, for both the manufacturer

[33] *Data Ethics Commission,* Empfehlungen der Datenethikkommission für die Strategie Künstliche Intelligenz der Bundesregierung vom 9.10.2018, p. 1 (available at: https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/it-digitalpolitik/datenethikkommission/empfehlungen-datenethikkommission.pdf?__blob=publicationFile&v=2, last accessed on: 05.08.2022).
[34] *Data Ethics Commission,* Gutachten der Datenethikkommission der Bundesregierung, 2019, p. 34 (accessible under: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=7, accessed 05.08.2022).
[35] *Lenzen*, Künstliche Intelligenz, 2020, p. 51.
[36] *Steege*, MMR 2019, 715 (719) with respect to predictive policing.
[37] Cf. *Lenzen*, Künstliche Intelligenz, 2020, p. 51 et seqq. For an overview of this problem.
[38] Real-life example of an application assessment software by Amazon s. https://www.zeit.de/arbeit/2018-10/bewerbungsroboter-kuenstliche-intelligenz-amazon-frauen-diskriminierung (last accessed on: 07.08.2022).
[39] Cf. General Act on Equal Treatment (AGG) Section 8 for acceptable reasons for unequal treatment due to practical job requirements.

and the consumer. The explicability of results obtained using a system based on ML techniques is limited at best.[40] Thus, it is often difficult to detect when a system is influenced by an inadmissible "bias".[41] The lack of transparency typical for certain AI systems thereby poses an independent risk to consumer rights. This particularly affects the freedom of self-determination of the individual. Where it is difficult or even impossible for users to understand how the AI system operates, it can have a negative impact on their perceptions as themselves as masters of their own decisions. Therefore, transparency and the ability to comprehend how the product works present a form of empowerment over technology, which can be considered as more or less significant depending on the use and application modalities of the AI system.

Here it should be noted, that especially non-transparent and complex AI systems often exhibit a particular stability and thus functional reliability.[42] Consequently, the aim to ensure the greatest possible transparency and the desire for an AI system to function without error can be contradictory. This leads to further risks for the rights of users.

Such risks also arise regarding the vulnerability of AI systems against external interference.[43] One may think of external attacks on the systems which aim to manipulate it to cause damage to third parties. Depending on how significantly the AI systems affect the rights of consumers, the demands for ensuring security against external interference grow louder. In this context, it is also important to consider the potential negative results of a successful attack on the respective AI system. If an AI application is used in a large capacity, manipulations of the system can have particularly significant effects on the rights and interests of the individuals involved.

3. Dynamic development of self-learning systems

The aforementioned AI-specific risks must be given special consideration in the context of a conformity assessment. This can be rather complicated due to another special feature of AI systems that are based on ML techniques. This refers to the *dynamic development*, as is typical for such AI applications. Arguably, the key advantage of using such AI systems is that, to a certain extent, they are able to adapt to new situations "autonomously". This means, that conclusions are drawn from the data that was processed by the system. These conclusions influence future solutions of a given problem. This can result in the deviation from previous working methods which cannot always be foreseen by humans. From this

[40] Cf. *Lenzen*, Künstliche Intelligenz, 2020, p. 54 et seq. for an overview of the so-called „Black-Box-Problem" as well as the research area of „Explainable AI" as a solution.
[41] *Wischmeyer*, Regulierung intelligenter Systeme, in AöR 143 (2018), 1 (28).
[42] *Cremers et al.*, Vertrauenswürdiger Einsatz von Künstlicher Intelligenz (Whitepaper des Verbundprojekts „Zertifizierte KI"), p. 17 (available at: https://www.iais.fraunhofer.de/content/dam/iais/KINRW/Whitepaper_KI-Zertifizierung.pdf , last accessed on: 07.08.2022); cf. *Ebers*, in: Rechtshandbuch Künstliche Intelligenz und Robotik, 2020, p. 90 et seqq.
[43] Cf. *Cremers et al.*, Vertrauenswürdiger Einsatz von Künstlicher Intelligenz (Whitepaper des Verbundprojekts „Zertifizierte KI"), p. 18 (available at: https://www.iais.fraunhofer.de/content/dam/iais/KINRW/Whitepaper_KI-Zertifizierung.pdf, last accessed on: 07.08.2022) for this AI-area.

kind of "self-dynamisms" of AI systems arises the desire to adapt assessment procedures to this special feature. While in general, static products must only pass one conformity assessment to ensure an appropriate safety standard, this is not sufficient for self-learning AI systems. Rather, a continuous review is needed to exclude such risks that only arise by virtue of self-development of the application through self-learning. Conformity assessments must therefore be equally dynamic – and thus able to adapt to their object of assessment and its specific characteristics.[44]

4. Lack of specified ethical and legal requirements

AI-conformity assessments are focused on ensuring the compliance with applicable law and potentially additional ethical standards. Where conformity is to be assessed, it presupposes the existence of an object of reference – specifically: applicable law or established ethical standards against which the AI application can be critically reviewed. However, the latter is currently not available at the extent which would be generally required to ensure the trustworthy use of technologies using artificial intelligence. Indeed, whilst data protection regulation has already reached very high standards within national law as a result of the General Data Protection Regulation, which also addresses important questions related to AI applications,[45] this does not apply to all AI-specific risks. Especially concerning the area of requirements regarding transparency and freedom from bias of AI systems that need to be met from a legal and ethical perspective, there remains a considerable need to catch up. The EU Commission's AI regulation draft represents a significant step in that direction. It followed a series of national and European papers that formulated the first general guidelines for the legal and general, societal treatment of AI systems. The recommendations by the High-Level Expert Group on Artificial Intelligence, as well as those by the German Data Ethics Commission are characterized by a high level of abstraction.[46] This proves to be rather beneficial for the initial societal orientation. However, based on this it is impossible to conduct a specific conformity assessment of an individual AI application. Against this backdrop, it is to be welcomed that the AI regulation draft makes initial specific regulatory proposals, even though they still lack adequate specification in large parts.[47] Additionally,

---

[44] Cf.*Heesen/Müller-Quade/Wrobel et al.*, in: Whitpaper „Zertifizierung von KI-Systemen der Plattform „Lernende Systeme", p. 30 et seq. (available at: https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG1_3_Whitepaper_Zertifizierung_KI_Systemen.pdf, last accessed on: 05.08.2022) for this requirement.
As well as regarding the difficulty to transfer this goal into practical operation, further details under C., I., 5. And C., II.

[45] Cf. *Spiecker gen. Döhmann/Bretthauer,* Dokumentation zum Datenschutz, 86. Ergänzungslieferung 2022, Hambacher Erklärung zur Künstlichen Intelligenz. Sieben datenschutzrechtliche Anforderungen, G 2.4.81.

[46] *Rostalski/Weiss*, ZfDR 2021, 329 (333 et seq.); *Hochrangige Expertengruppe für künstliche Intelligenz*, Ethik-Leitlinien für eine vertrauenswürdige KI, 2019 (available at: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai, last accessed on: 07.08.2022); *Datenethikkommission*, Empfehlungen der Datenethikkommission für die Strategie Künstliche Intelligenz der Bundesregierung vom 9.10.2018, p. 3 (available at: https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/it-digitalpolitik/datenethikkommission/empfehlungen-datenethikkommission.pdf?__blob=publicationFile&v=2, last accessed on: 07.08.2022).

[47] Cf. B., IV., 2. for more details.

the procedure of the EU for the issuance of a regulation has not yet been completed. This is then partly followed by the implementation of the regulation into national law leading to further specification – e.g., regarding possible sanctioning standards.[48] Parallelly, a controversial ethics debate on how to deal with AI systems has arisen. In order to develop ethical standards that can be taken into account in the context of the assessment procedure, further societal negotiations are required which however are evidently not near completion. Such regulatory uncertainties encumber conformity assessments as well. The latter are directly dependent on suitable specifications regarding the use of AI systems. Furthermore, they must be accessible to reliable review. Thus, the existing legal gaps as well as the lack of common ethical standards are posing a further challenge to the implementation of AI conformity assessments.

5. Lack of AI-centric assessment procedures – A challenge

Conformity assessments in relation to AI systems represent to a significant degree the "re-mapping" of previously unknown areas of human action. Experiences acquired in the context of other products can only be transferred to a limited extent at best. The reason being the aforementioned special features of this technology, which must be taken into account in their corresponding assessment. However, because these challenges have only been addressed peripherally in other contexts, pioneering work must be done. Nevertheless, this work is not carried out in a vacuum. Experiences from other areas may be valuable – at least to a limited extent. For example, continuous monitoring procedures are already employed in the area of cloud-auditing. Though, it cannot yet be applied to AI systems for technical reasons. However, it can provide a starting point for further development efforts, which may allow similar continuous assessments of such systems in the future. Additionally, knowledge gained in the past regarding the structure required for conformity assessment procedures must be taken into account as well. This can also benefit the development of assessment procedures for AI systems.

III. General requirements for conformity assessments of AI systems

Conformity assessment procedures regarding AI systems have to address the challenges posed by technology as outlined above. In addition to conceptual clarity regarding the test object (AI system), consideration must be given to the numerous specifics that arise with this technology. Thus, in the context of conformity assessments, special attention must be paid to data protection, protection against discrimination and the compliance with transparency requirements. A particular challenge is posed by the dynamic nature of AI applications. This must also be appropriately reflected within the conformity assessments of such systems. Last, but not least, it must be determined under what conditions conformity assessments of AI systems have to be mandatory for the responsible individuals. How this may look will be outlined in more detail later.

[48] Cf. Art. 71 of the draft AI Regulation.

B. The so-called "risk-based regulatory approach" of the draft AI Regulation from the perspective of the consumers

Before proposals for the concrete structure of conformity assessments of AI systems can be articulated, the current proposal of the EU Commission will be examined in greater detail with regard to this issue. The stipulation for conformity assessment procedures are embedded in the "risk-based regulatory approach" of the draft AI Regulation. Said approach will be outlined and subjected to a critical analysis as relevant for the consumers. This will also be of importance for the subsequently discussed requirements for conformity assessment procedures.

I. Scope of application of the draft AI Regulation

First of all, it is necessary to identify the scope of the regulation. This is necessary as, among other things, the scope determines the extent to which the requirements of the regulation can be circumvented in a way that would hinder consumer protection. The scope of application is stipulated in Art. 2 (1) of the draft AI Regulation, while Art. 2 (2) till (4) contain certain restrictions, or exceptions, e.g. with regard to AI systems that are developed or used solely for military purposes (Art. 2 (3) of the draft AI Regulation). According to Art. 2 (1) of the draft AI Regulation it applies to:

a) providers placing on the market or putting into service AI systems in the Union, irrespective of whether those providers are established within the Union or in a third country;
b) users of AI systems located within the Union;
c) providers and users of AI systems that are located in a third country, where the output produced by the system is used in the Union.

Particular attention should be paid in this context to Art. 2 (1) c) of the draft AI Regulation. As cloud usage expands, it is likely to become increasingly more difficult to pinpoint the specific location of an AI system; moreover, it should be easy enough to relocate it to a third country.[49] The European Commission anticipated this development and therefore allowed the application pursuant to c). Thus, in recital 11 it is explicitly stated that "to prevent the circumvention of this Regulation and to ensure an effective protection of natural persons located in the Union, this Regulation should also apply to providers and users of AI systems that are established in a third country, to the extent the output produced by those systems is used in the Union." Taking into account Art 2 (1) c) of the draft AI Regulation, it should be difficult to evade the application of the AI-regulation where the AI system has

[49] *Bomhard/Merkle*, RDi 2021, 276 (278).

any impact - which have to be specified in more detail[50] - in the European Union. Therefore, the scope of application – especially in terms of consumer protection – is extensive and thus effective.

II. The so-called "risk-based categorisation approach"

The extent to which legal requirements are imposed on AI systems is determined within the draft AI Regulation by a so-called risk-based approach. The more intrusive ("critical") an AI application turns out to be, the more extensive are the requirements it must meet. In this way, the appropriate protection of European fundamental rights and values is meant to be warranted in individual cases. Against this background, the draft distinguishes between four risk levels: AI systems with a prohibited risk level (cf. Art. 5 of the draft AI Regulation), with a high-risk level (cf. Art. 6 of the draft AI Regulation), with a low risk level (Art. 52 of the draft AI Regulation) and with a minimal risk level (Art. 69 of the draft AI Regulation). The allocation of individual AI systems to the respective risk levels, as well as the specific structure of the categories is based on assessments made by the European Commission. The degree of consumer protection depends decisively on this allocation system and structure. In the following, this report will examine the question of whether the current version of the so-called "risk-based categorisation approach" adequately takes consumer interests into account. This is due to the fact, that the conformity assessment system as stipulated in the regulation proposal is only capable of providing consumer protection to an extent, as is addressed in the regulatory concept and its requirements.

III. Sufficient consumer protection in the context of prohibited AI-practices as stipulated in Art. 5 of the draft AI Regulation?

The first question that must be addressed is, whether consumer protection is sufficiently taken into account in the context of prohibited AI-practices as stipulated in Art. 5 of the draft AI Regulation. AI systems that carry an unacceptable risk are prohibited from the outset under Art. 5 of the draft AI Regulation.[51] This includes systems that according to the Commission, are particularly harmful to European fundamental rights and values.[52] Of considerable note are systems that aim to distort a person's behaviour (Art. 5 (1) a) and b) of the draft AI Regulation, as well as so-called "social scoring" systems (Art. 5 (1) c) of the draft AI Regulation), as they are associated with serious risks for the consumer.[53]

[50] Cf. *Bomhard/Merkle*, RDi 2021, 276 (278) regarding the question whether it should cover merely direct output or indirect output as well.
[51] The following explanations regarding AI systems concerning Art. 5 of the draft AI Regulation are based on *Rostalski/Weiss*, ZfDR 2021, 329 (337 et seqq.).
[52] COM(2021) 206 final, p. 1 et seqq., under 1.1., as well as p. 21, recitals 15-18.
[53] Cf. for an extensive analysis *Rostalski/Weiss*, ZfDR 2021, 329 (342 et seqq.) regarding the "prohibition" of biometric real-time-remote identification systems as per Art. 5(1)d) of the draft AI Regulation.

1. AI-systems of behaviour manipulation (Art. 5 (1) a) and b) of the draft AI Regulation)

Certain AI systems that are intended to be used for the purpose of behavioural manipulation are prohibited under Art. 5 (1) a) and b) of the draft AI Regulation. Such applications can have a significant impact on behaviour by affecting a person "subliminally"[54] - beyond consciousness -, or by exploiting people's weaknesses or need for protection (e.g., as a result of a disability). For this ban, the draft also requires that the intended behavioural influence causes or is likely to cause physical or psychological harm. The prohibition is intended to protect the right of EU citizens to self-determination as a foundation of the European value system.[55] Accordingly, it aligns with a central concern of consumer protection.

The prohibition of AI systems that are used for behavioural manipulation is the only prohibition in the draft that extends to both government and private sectors. However, it is striking that the, for this provision significant, term "subliminal techniques beyond a person's consciousness", is not explained in more detail either within the draft itself or in the materials. The wording indicates that influences through the use of so-called "dark patterns" seem to be included.[56] These are digital design patterns based particularly on psychological insights to induce people to behave in certain ways that run counter to their actual intentions (i.e. their own will) or behaviour that they would not otherwise have adopted.[57] Examples of this are displaying a countdown, which suggest that there is an (apparently) limited time offer, or the reference that a product is (supposedly) in low stock, as well as references to the (alleged) actions of other users.[58] Additionally, the ban addresses the practice of "nudging" which is quite similar to "dark patterns".[59] The term "nudging" describes the intentional manipulation of individuals by deliberately triggering subconscious processes that lead to changes in behaviour.[60] Unlike "dark patterns", "nudging" does not serve the sole purposes of the user. Rather, it is intended to bring about behaviour that corresponds with the actual interests of the targeted individual (as presumed by a third party) or at least with the interests of other members of society.[61] Finally, the use of personalised advertising is also likely to constitute an act of manipulation within the meaning of the provision.[62]

---

[54] The draft proposal uses the formulation „subliminal techniques beyond a person's consciousness" – this is redundant: everything beyond a person's consciousness occurs subliminal.
[55] "Practices that have a significant potential to manipulate persons […]" are to be captured, COM(2021) 206 final, p. 12 et seq. The German Basic Law protects the right to self-determination as part of the right to personal freedom, *Di Fabio*, in: Maunz/Dürig, Grundgesetz, 96. EL November 2021, Art. 2 Abs. 1 Rn. 147.
[56] Similarly *Ebert/Spiecker genannt Döhmann*, NVwZ 2021, 1188 (1189).
[57] Cf. for a definition *Martini/Drews/Seeliger/Weinzierl*, ZfDR 2021, 47 (49 with further references).
[58] *Martini/Drews/Seeliger/Weinzierl*, ZfDR 2021, 47 (49).
[59] Similarly *Valta/Vasel*, ZRP 2021, 142 (143).
[60] *Hufen*, JuS 2020, 193 (193 et seq.).
[61] Cf. for the definition and distinction *Martini/Drews/Seeliger/Weinzierl*, ZfDR 2021, 47 (51 with further references), who suggest the use of the terms „Dark Patterns" and "Dark Nudging" synonymously.
[62] *Bomhard/Merkle*, RDi 2021, 276 (279).

However, both the prohibition as per Art. 5 (1) a) of the draft AI Regulation and the one for the protection of persons with disabilities or e.g., children[63] according to Art 5(1) b) of the draft AI Regulation are subject to restrictions. Both prohibitions require intent ("in order to […] distort a person's behaviour […])[64] regarding the "material" distortion of a person's behaviour. Such intent cannot be presumed "if the distortion of human behaviour results from factors external to the AI system which are outside of the control of the provider or the user."[65] The (supposed) scope of this restriction has invoked criticism; it waters down the prohibition too much.[66] However, this is not to be agreed with. Rather, the current provision proves to be too restrictive insofar as it requires intent regarding the use of the AI system. Consequently, where AI systems are used without the intention to "materially" influence behaviour within the meaning of the provision, they are not prohibited even if they do in fact cause such manipulation. But when it comes to the comprehensive protection of autonomy, it proves to be dysfunctional. Therefore, instead of an intention requirement, a restricting objective criterion is preferable. One could e.g., think of a "suitability for material distortion of behaviour" requirement. Moreover, such a structure would take the principle of proportionality into account. This would ensure that the ban extends only to those AI systems that pose a particular threat to the autonomy of EU citizens. Thus, certain methods of personalised advertising that do not materially affect behaviour do not fall foul of the provision.[67] This is not precluded by virtue of the open nature of the concept of "materiality". The necessity of specification[68] is generally a characteristic of indeterminate legal terms and thus should not be criticised. This difficulty could be addressed through e.g., the inclusion of exemplary cases of AI systems that are subject to the ban in the materials. This would also serve as meaningful guidance for legal practitioners, who could make use of the technique of grouping comparable cases together.

As shown above, the requirement of causing physical or psychological harm through materially influencing behaviour, or more specifically the suitability for causing such harm constitutes a further restriction of the ban in question. However, it is not apparent why behavioural manipulation should only constitute a form of undesirable infringement of the right to self-determination if it is suitable to do so. Additionally, the restriction leads to an objectively inappropriate shift of the legal interest that is protected by the ban: From protecting self-determination to protecting health. Against this background, it is also not consistent to merely demand an expansion of the catalogue that would cover economic or

[63] They are explicitly listed in the draft as a group of persons covered "due to their age", cf. COM(2021) 206 final, p. 21, recital 16.
[64] The recitals explicitly refer to the "intention to materially distort the behaviour of a person", cf. COM(2021) 206 final, p. 21, recital 16.
[65] COM(2021) 206 final, p. 21, recital 16.
[66] In this sense, probably *Ebert/Spiecker genannt Döhmann*, NVwZ 2021, 1188 (1189), according to which both prohibitions are severely weakened by this restriction
[67] The concern voiced by *Bomhard/Merkle*, RDi 2021, 276 (279) that some business models of personalized advertising would „face extinction" due to this ban must therefore be relativized.
[68] *Bomhard/Merkle*, RDi 2021, 276 (279).

material impairments on their part.[69] Indeed, this as well as the *mandatory* requirement of the, at least, potential impairment of health, is a well-intentioned attempt to define the limits of a no longer tolerable impairment of human autonomy. One argument in favour of such effort is that the principle of proportionality demands the specific limitation of undesirable manipulation. In this respect, it seems preferable – not least for reasons of effectiveness – to use an abstract criterion such as the objective capability for material behavioural manipulation. After what has already been laid out, this would, of course, need to be specified. In particular, the question would need to be answered as to what kind of manipulations we do not want to accept (anymore) within our community. A consensus could be found, e.g., regarding manipulations that are capable of damaging health. This "case group" could be included as a "standard example" of "material distortion of behaviour". It certainly makes sense to prohibit forms of manipulation that can cause physical or mental harm. This way, the boundary separating prohibited from acceptable behavioural manipulation would be drawn closer, and simultaneously gain a degree of flexibility[70] without having to exchange the legally protected right as protected by the prohibition. Indeed, it is still the individual right to self-determination. AI systems may still be caught by the prohibition if they can lead to comparatively serious manipulations, even if they are not capable of adversely affecting health in the individual case.

A provision that adequately takes into account the rights of consumers could be formulated as follows:

a) *"The placing on the market, putting into service or use of an AI system that deploys subliminal*[71] *techniques beyond a person's consciousness and which is capable of materially distort a person's behaviour. The capability to materially distort a person's behaviour can usually be assumed, if the distortion can cause that person or another person physical or psychological harm;*
b) *The placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific group of person due to their age, physical or mental disability and is capable to materially distort the behaviour of a person pertaining to that group. The capability to materially distort the behaviour of a person pertaining to that group can usually be assumed if the distortion causes that person or another person physical or psychological harm."*

2. So-called "social scoring" AI systems (Art. 5 (1) no. 1 c) of the draft AI Regulation)

Under certain conditions, Art. 5 (1) no. 1 c) of the draft AI Regulation also prohibits so-called "social scoring" by means of AI systems. The provision covers the evaluation or classification of the trustworthiness of natural persons over a certain period of time based on their social behaviour or known or predicted personal or personality characteristics. As

[69] *Ebert/Spiecker genannt Döhmann*, NVwZ 2021, 1188 (1189); similarly *Valta/Vasel*, ZRP 2021, 142 (143).
[70] The indicative effect of the fulfillment of the standard example could be disproved in individual cases.
[71] See Fn. 56 above on the redundancy of this feature – in the best case, it would be deleted in its entirety in a new version.

a further condition, the AI-based assessment either has to lead to a detrimental or unfavourable treatment in social contexts unrelated to the original data collection (Art. 5 (1) c) i) of the draft AI Regulation), or the detrimental or unfavourable treatment must be unjustified or disproportionate to the social behaviour or its gravity (Art. 5 (1) c) ii) of the draft AI Regulation). For it to fall under the scope of the prohibition, the assessment has to be carried out "by public authorities or on their behalf". "Social scoring" by means of AI systems in the purely private sector is therefore not affected by the regulation.

The background to the ban is likely to be the development of a state-run "social scoring" system in China, which, albeit frightening by our standards, is nevertheless progressing rapidly.[72] This system is based on the concept of points awarded for behaviour that is desirable from the government's point of view, while less desirable behaviour leads to points being deducted. A "low" score can result in considerable restrictions in everyday life, such as access to social services or the allocation of loans as well as training or jobs.[73] Such a regime would be incompatible with a value system based on individual self-determination. If individuals had to fear that their entire behaviour would be evaluated according to the interests of the state, they would no longer be able to decide freely for or against a certain behaviour. Impending interventions in their everyday life would massively restrict them in their responsible decisions, or even virtually exclude what we understand by a "free" decision. Against this background, the ban is certainly to be welcomed – not least from a consumer's perspective. Its limitation to the public sector, however, has rightly been criticised.[74] That is because in Europe, the dystopia of such systems seems much more probable in the private sector, e.g. when it comes to the conclusion of contracts or the access to services.[75] A research project by Schufa[76], in which data from social media was used in the context of credit rating checks, is a prime example. The project was ultimately discontinued due to intense public protests,[77] which also demonstrates the considerable lack of acceptance towards such AI applications in the broader population. In the further legislative process, social scoring applications in the private sector should therefore also be prohibited. The occasional criticism[78] directed at the alleged "considerable degree of legal uncertainty" associated with the balancing of interests to be carried out within the framework of Article 5 (1) c) ii) of the draft AI Regulation must be contradicted. As can be concluded *e contrario* from Art. 5 (1) c) i) of the draft AI Regulation, this second alternative concerns cases where the detrimental or unfavourable treatment occurs precisely in the context for which the data was originally generated or collected. An evaluation based on the use of data for a specific purpose cannot be illegitimate per se, which makes it necessary to precisely define the conduct that should be prohibited. Exactly that is the purpose of requiring a detrimental or

[72] *Valta/Vasel*, ZRP 2021, 142 (143); *Geminn*, ZD 2021, 354 (356).
[73] *Wagner*, ZD 2020, 140 (141).
[74] *Ebert/Spiecker genannt Döhmann*, NVwZ 2021, 1188 (1189); *Valta/Vasel*, ZRP 2021, 142 (143).
[75] *Valta/Vasel*, ZRP 2021, 142 (143); *Ebert/Spiecker genannt Döhmann*, NVwZ 2021, 1188 (1189).
[76] Schutzgemeinschaft für allgemeine Kreditsicherung, a German private credit bureau.
[77] *Valta/Vasel*, ZRP 2021, 142 (143).
[78] *Spindler*, CR 2021, 361 (365).

unfavourable treatment that is "unjustified or disproportionate" to the social behaviour or its gravity. In order to make the criterion (even) more accessible, including examples in the legislative materials seems to be advisable for the further legislative process, so that practitioners can use them for orientation.

3. Interim result

In principle, the Commission's efforts to identify AI systems that go beyond what is tolerable under our system of values are to be welcomed – also in terms of appropriate consumer protection. Beyond the weaknesses already pointed out, however, it is also generally doubtful whether the list of prohibited practices is sufficiently comprehensive.[79] One need only think, for example of the use of AI systems in the courts, all of which are classified, without differentiation, as (merely) high-risk AI systems under Art. 6 (2) in conjunction with Annex III no. 8 of the draft AI Regulation. Whether, for instance, the use of a system such as "COMPAS", which some US criminal courts use for prognoses regarding a person's risk of relapse,[80] would be compatible with our European system of values, seems highly doubtful.[81] This example illustrates the need to initiate a critical social discourse on the limits of tolerable AI systems. It is to be hoped that the list of Article 5 of the draft AI Regulation will be reviewed in the further legislative process to establish if there is any need for amendments, and adapted accordingly.

IV. Sufficient consumer protection through the requirements for so-called "high-risk AI systems" (Art. 6 of the draft AI Regulation) pursuant to Art. 8 et seqq. of the draft AI Regulation?

High-risk AI systems within the meaning of Article 6 of the draft AI Regulation are, in a way, the central category and thus the "centerpiece" of the draft. Applications that "have a significant harmful impact on the health, safety and fundamental rights of persons in the

[79] This is criticized by MEP Alexandra Geese, who argues that the AI-assisted determination of gender, sexual orientation, ethnicity, state of health and disability should also constitute a prohibited use, see https://alexandrageese.eu/offener-brief-an-die-kommission-keine-ki-zur-geschlechtserkennung-in-der-eu/, last accessed on: 07.08.2022.

[80] Cf. for more details on this system and its use in decisions on pre-trial detention, the severity of the penalty and the suspension of the sentence *Nink*, Justiz und Algorithmen, 2021, p. 376 et seqq.

[81] Sceptical of the compatibility with Art. 6 (1) ECHR is *Valerius*, Legal Tech im Strafverfahren, ZStW 133 (2021), 152 (166 et seqq.), who also discusses the compatibility of sentencing algorithms with the right to a fair hearing under Art. 103 (1) of the German Basic Law (164 et seq.). On the compatibility of the use of such AI systems with the German Basic Law, see also *Nink*, Justiz und Algorithmen, 2021, p. 422 et seqq., who, among other things, raises the issue of their compatibility with judicial independence under Article 97 (1) of the German Basic Law; cf. furthermore the opinion of the Bundesrat on the draft AI Regulation, in which it suggests "to examine whether it should be expressly made clear in Article 5 (1) b) of the proposed Regulation that the judicial decision may not be transferred to an AI system and that no AI systems may be used which, by their design, harbour the risk that the judicial decision-making process is influenced to the effect that, in the case of several justifiable legal opinions, a preselection is made and the result is based on it." On the AI-based creation of sentencing databases to overcome factually unjustified regional differences in sentencing, see *Rostalski/Völkening*, KriPoZ 2019, 265 et seqq.; *Rostalski/Völkening*, ZfDR 2021, 27 et seqq.

Union"[82] are meant to fall under this category. Art. 8 et seqq. of the draft AI Regulation set out a number of requirements with regard to such AI systems, such as the establishment of risk management systems (Art. 9 of the draft AI Regulation), the observance of specific criteria regarding data governance (Art. 10 of the draft AI Regulation) or compliance with comprehensive documentation (Art. 11, 12 of the draft AI Regulation) and transparency obligations (Art. 13 of the draft AI Regulation).

1. Definition and differentiation of the „high-risk AI systems" pursuant to Art. 6 (1) and (2) of the draft AI Regulation

The draft distinguishes between two types of high-risk AI systems in Art. 6 (1) and (2) of the AI Regulation.

a. High-risk AI systems in the sense of Art. 6 (1) of the draft AI regulation

Classification in the high-risk category according to Art. 6 (1) of the draft AI Regulation is based on two conditions. According to Article 6 (1) a) of the draft AI Regulation, AI systems must be safety components of products covered by specific Union harmonisation legislation listed in Annex II. A safety component of a product is defined by Art. 3 no. 14 of the draft AI Regulation as a component of a product which fulfils a safety function for that product or the failure or malfunctioning of which endangers the health and safety of persons or property. However, according to Art. 6 (1) a) of the draft AI Regulation, AI systems are also covered if they themselves constitute a product covered by the harmonisation legislation listed in Annex II. According to Article 6 (1) b) of the draft AI Regulation, both cases also require that either the products or the AI systems themselves are required to undergo a third-party conformity assessment in accordance with the relevant harmonisation legislation. Harmonisation legislation includes various European directives and regulations based on the "New Legislative Framework" regulating European conformity testing.[83] Noteworthy examples are the Regulation on machinery products (Annex II, Section A, No. 1 of the draft AI regulation), the Regulation on medical devices (Annex II, Section A, No. 11 of the draft AI Regulation) as well as the Regulation on in vitro diagnostic medical devices (Annex II, Section A, No. 12 of the draft AI Regulation). AI applications in the medical sector are likely to be predominantly subject to the sectoral legal provisions of the latter two regulations and to be regularly classified as high-risk AI systems as a consequence.

b. High-risk AI systems in the sense of Art. 6 (2) of the draft AI regulation

Pursuant to Art. 6 (2) of the draft AI Regulation, independent, so-called "stand-alone", AI systems are also classified as high-risk if "in the light of their intended purpose, they pose a high risk of harm to the health and safety or the fundamental rights of persons"[84]. The

[82] COM(2021) 206 final, p. 24, recital 27.
[83] *Spindler,* CR 2021, 365 (366).
[84] COM(2021) 206 final, p. 26, recital 32.

AI systems covered are listed in Annex III of the draft. First, the abstract field of application is defined, the relevant areas of which are then specified in more detail in the form of subgroups. The sectors covered – more narrowly defined in detail – are "biometric identification and categorisation of natural persons" (Annex III no. 1 of the draft AI Regulation), "management and operation of critical infrastructures" (Annex III no. 2 of the draft AI Regulation), "education and vocational training" (Annex III no. 3 of the draft AI Regulation), "employment, workers management and access to self-employment" (Annex III no. 4 of the draft AI Regulation), "access to and enjoyment of certain essential private and public services and benefits" (Annex III no. 5 of the draft AI Regulation), "law enforcement" (Annex III no. 6 of the draft AI Regulation), "migration, asylum and border control management" (Annex III no. 7 of the draft AI Regulation) and "administration of justice and democratic processes" (Annex III no. 8 of the draft AI Regulation). In order to react dynamically to future changes, the Commission is empowered by Art. 7, 73 of the draft AI Regulation to extend the list of high-risk applications in Annex III. Two cumulative conditions apply: The respective AI system must be used in one of the areas listed in Annex III (Art. 7 (1) a) of the draft AI Regulation). In addition, it must 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 in Annex III (Art. 7 (1) b) of the draft AI Regulation). Article 7 (2) of the draft AI Regulation contains an extensive list of criteria to be taken into account when determining whether the risk is at least comparable with the one defined in Article 7 (1) b) of the draft AI Regulation.

The regulatory technique via Annex III allows for a flexible expansion of high-risk AI systems, which is certainly forward-looking. However, the rigid reference to sectors poses the risk that covering newly emerging hazardous AI systems will be unfeasible.[85] If an AI system cannot be assigned to any of the sectors listed in Annex III, it falls through the grid. It cannot be subsequently classified as a high-risk application either, as Article 7 (1) a) of the draft AI Regulation requires a sectoral reference for an extension of the list. As a result, the intended future-proof design of the regulation is at least partially undermined. In his opinion on the European Commission's AI White Paper, the European Data Protection Supervisor had already called for the sectoral classification not to be a rigid requirement, but merely one of several criteria for determining high-risk applications.[86] We suggest that this demand will be made again emphatically in the course of the further legislative process, and ultimately also implemented.[87] Otherwise, there is a risk of leaving a gap in the regulatory

[85] *Hoffmann*, K&R 2021, 369 (371).
[86] *European Data Protection Supervisor*, EDPS Opinion on the European Commission's White Paper "On Artificial Intelligence – A European approach to excellence and trust 2020", p. 11 et seq. recital 29 (available at: https://edps.europa.eu/sites/default/files/publication/20-06-19_opinion_ai_white_paper_en.pdf, last accessed on: 07.08.2022).
[87] *Hoffmann*, K&R 2021, 369 (371).

concept causing an enormous risk for abuse as well as possible dangers for the rights of consumers.

2. The requirements of Art. 8 et seqq. of the draft AI Regulation

The requirements for high-risk AI systems are codified in detail in Art. 8 et seqq. of the draft AI Regulation and will be briefly presented in the following.[88]

a. Establishment of a risk management system (Art. 9 of the draft AI Regulation)

Article 9 of the draft AI regulation firstly requires the establishment of a so-called "risk management system". It must be in place during the entire life cycle of an AI system, and regularly and systematically updated (Art. 9 (2) sentence 1 of the draft AI Regulation). It includes the elements already known from other contexts, in particular the identification and analysis of possible risks (Art. 9 (2) sentence 2 a) to c) of the draft AI Regulation) as well as the determination and, if necessary, the adoption of appropriate risk management measures (Art. 9 (2) sentence 2 d), (3), (4) of the draft AI Regulation). There are no excessive requirements placed on these measures. Rather, they must be designed – in accordance with the principle of proportionality – in such a way "that any residual risk associated with each hazard as well as the overall residual risk of the high-risk AI systems is judged acceptable" (Art. 9 (4) sentence 1 Draft AI Regulation). Finally, the testing of AI systems, at the latest before they are placed on the market or put into service, is mandatory (Art. 9 (5)-(7) of the draft AI Regulation). These requirements are evidently aimed at identifying any risks of an AI system at an early stage, and minimizing them to a tolerable level. This objective as such is to be welcomed. At the same time, the requirements raise two practical problems in particular. On the one hand, it is difficult to determine which risks in a particular case are still "foreseeable", especially with regard to AI systems based on machine learning. It seems likely that the standards applied here will ultimately not be too strict, if one wants to keep these requirements actually feasible. On the other hand, a more detailed definition of the appropriate risk management measures is required, in particular when it comes to effective conformity assessments. According to Article 9 (3) sentence 2 of the draft AI Regulation, the measures shall "take into account the generally acknowledged state of the art, including as reflected in relevant harmonised standards or common specifications". At this point, as well as in general, the prominent role of harmonised technical standards becomes evident. According to the Commission, standardisation should "play a key role to provide technical solutions to providers to ensure compliance with this Regulation" (recital 61). This point is also reflected in its legal effects: Systems that are in conformity either with harmonised standards (as defined in Art. 40 of the draft AI Regulation) or with common specifications (as defined in Art. 41 of the draft AI Regulation) are presumed to be in conformity with

[88] For a more detailed introduction and analysis of individual requirements, cf. *Spindler*, CR 2021, 361 (366 et seqq.); *Ebers/Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (533 et seqq.); *Linardatos*, GPR 2022, 58 (63 et seqq.).

the respective requirements (cf. Art. 40, Art. 41 (3) of the draft AI Regulation). The Commission obviously relies on harmonised technical standards for the practical feasibility of the requirements of Art. 9 of the draft AI Regulation.

b. Requirements for data and data governance according to Art. 10 of the draft AI Regulation

Article 10 of the draft AI Regulation sets out requirements for the data used by AI systems, and for data governance. In particular, the data itself must be "relevant, representative, free of errors and complete" (Art. 10 (3) sentence 1 of the draft AI Regulation). Data governance itself requires, among other things, the formulation of relevant assumptions, a prior assessment of the availability, quantity and suitability of the data sets that are needed, and an examination in view of possible biases (cf. Art. 10 (2) of the draft AI Regulation). With these requirements, the Commission aims to ensure that the use of AI systems does not become the source of discrimination (recital 44). Once again, the intention behind these requirements is to be welcomed. AI systems are only as "intelligent" as the respective training data allow. If a data set is distorted, the system will reproduce the distortion. For this reason, in a conformity assessment problems arise regarding the practical feasibility and consequently the evaluation, if a system meets the respective requirements. The central question of *when* the data used meet the requirements is left open. In practice, the problem often arises that it is not recognizable in advance that a data set is not completely non-discriminatory. One example would be the case of so-called "proxy" discrimination. Even though a protected characteristic, such as skin colour or gender, is not addressed by the algorithm itself, discrimination against the protected characteristic can nevertheless occur. Amazon's applicant assessment software, for instance, discriminated against women even though gender as such was not recorded. Based on individual information provided in the CV, the system nevertheless recognised whether the information was to be associated with a man or a woman. In previous recruitment practice, women were less likely to be recruited. The system thus indirectly included gender in its assessment. The example illustrates that the absence of a protected characteristic in a data set cannot guarantee complete protection against discrimination. The question in particular of whether data is "free of errors" in the sense of the provision is also directly related to the particular understanding of fairness upon which it is based. In this respect, further concretisation is required if the requirements set out are to be practicable and verifiable.

c. Technical documentation in the sense of Art. 11 of the draft AI Regulation

According to Art. 11 of the draft AI Regulation, technical documentation of the AI system shall be drawn up before it is placed on the market or put into service and kept up-to date (Art. 11 (1) AI of the draft AI Regulation). The information to be documented is comprehensively listed and specified in more detail in Annex IV of the draft, cf. Art. 11 (2) of the draft AI Regulation. The central purpose of this documentation obligation is explicitly named in the provision itself: According to paragraph 1, subparagraph 2 it is meant to create

a basis for the necessary conformity assessments pursuant to Article 43 of the draft AI Regulation. Accordingly, this is an elementary requirement of the draft, without which any conformity assessments and thus the regulatory system as a whole would hardly be able to function.

d. Record-keeping obligations pursuant to Art. 12 of the draft AI Regulation

Closely related to the technical documentation obligations are the record-keeping obligations under Art. 12 of the draft AI regulation. According to Art. 12 (1) sentence 1 of the draft AI Regulation, AI systems must be "designed and developed with capabilities enabling the automatic recording of events while the high-risk AI system is operating". Paragraph 4 specifies in more detail, what must be documented. These record-keeping obligations are ultimately also intended to ensure that the respective AI system can be assessed in view of its conformity with the Regulation. They also address the so-called "black box"- problem of certain AI systems, which was already outlined before. To resolve this problem, the research field "Explainable AI" is working on the development of so-called "explanatory models".[89] This approach involves, among other things, research into how different inputs statistically affect the output of an AI system. Given that processes and events in the use of an AI system can only be effectively looked into if they have also been recorded, the recording obligations under Art. 12 of the draft AI Regulation may well also be an indispensable element of the draft. With regard to the practical feasibility of these requirements, reference is again made to harmonised technical standards: According to Art. 12 (1) sentence 2 of the draft AI regulation, logging shall "conform to recognised standards or common specifications".

e. Transparency and information requirements pursuant to Art. 13 of the draft AI Regulation

The transparency and information requirements of Art. 13 of the draft AI Regulation can be considered as a part of a number of documentation and recording obligations. According to Art. 13 (1) sentence 1 of the draft AI Regulation, AI systems must be designed and developed in such a way "to ensure that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately". An "appropriate type and degree" of transparency "shall be ensured" (Art. 13 (1) sentence 2 of the draft AI Regulation). In addition, instructions for use must be made available in accordance with Article 13 (2) of the draft AI Regulation. These should contain "concise, complete, correct and clear information that is relevant, accessible and comprehensible to users". The required information is specified in more detail in paragraph 3. Transparency with regard to work processes and the provision of sufficient information are necessary conditions for people to be able to use AI systems on their own terms. At the same time, these conditions ensure trust in the individual AI systems and are thus also commendable from this perspective. All the

[89] Cf. for a brief overview *Lenzen*, Künstliche Intelligenz, 2020, p. 54 et seq.

same, the question of practical feasibility also arises in this respect, especially with regard to the processes of AI systems posing the black-box problems already described. At least currently, it appears difficult to make these processes transparent in a way "to enable users to interpret the system's output and use it appropriately". If understanding how or by what means a certain result has come about is impossible, there is no suitable basis for an appropriate interpretation. At least, this applies when essential work steps can no longer be understood or reconstructed. It is more than possible, that this requirement will turn out to present a fundamental (and depending on the context justified) barrier so the use of even very well-functioning AI systems. The reason for this is that the specific working processes of particularly efficient systems are often, at least at present, very difficult to access and trace. It is therefore desirable, that these requirements will be further specified in the course of the ongoing legislative process. This especially applies to the "black box" problem, which should be addressed in an appropriate manner.

f. Human oversight pursuant to Art. 14 of the draft AI Regulation

Art. 14 of the draft AI regulation lays down the obligation to ensure that there is an effective human oversight of the AI system for the duration of its use (Art. 14 (1) draft AI Regulation). If possible, the necessary measures must already be integrated into the system or be designed in such a way that they can be implemented by the respective user (Art. 14 (3) draft AI Regulation). These requirements are specified in more detail in Article 14 (4) of the AI Regulation. Inter alia, the human supervisor must be enabled to "fully understand the capacities and limitations of the high-risk AI system and be able to duly monitor its operation [...]" (Art. 14 (4) a) of the draft AI Regulation), "remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system ('automation bias') [...]" (Art. 14 (4) b) of the draft AI Regulation), "be able to correctly interpret the high-risk AI system's output [...]" (Art. 14 (c) of the draft AI regulation), and to react immediately to malfunctions (cf. Art. 14 (4) d) and e) of the draft AI Regulation).

These requirements are aimed at safeguarding the human freedom of self-determination, which is an indispensable foundation of our liberal value system. Users must be given the opportunity to appropriately control the system. A prerequisite for this, however, is a certain level of proficiency on the part of the overseeing person. The overseer must have the necessary knowledge to be able to exercise supervision adequately. Behind this requirement lies another challenge that should not be underestimated. One might justifiably be tempted to ask who or how many persons can fulfil this extremely demanding profile at all. The number of highly qualified and appropriately trained personnel presently appears to be rather modest. Therefore, it is to be feared that in practice, smaller companies in particular will be forced to resort to insufficiently qualified personnel. As a consequence, the laudable goal of human oversight will ultimately remain nothing more than a mere lip service. Such a scenario must be resolutely prevented. Therefore, at the very least, a targeted enhancement of the training of correspondingly skilled personnel is necessary.

g. Accuracy, robustness and cybersecurity pursuant to Art. 15 of the draft AI Regulation

Finally, according to Art. 15 of the draft AI Regulation, AI systems must be designed and developed in such a way that they "achieve […] an appropriate level of accuracy, robustness and cybersecurity, and perform consistently in those respects throughout their lifecycle" (Art. 15 (1) of the draft AI Regulation). The reliability and the security of AI systems are crucial factors for ensuring their trustworthy use. They, too, are indispensable preconditions. However, there is again the practically relevant question of what an "appropriate level" in the sense of the regulation looks like in individual cases. It seems likely that in the commission's view, the answer will ultimately lie once more in the development of correspondingly harmonised technical standards.

h. Interim result

All requirements for high-risk AI systems are in principle to be welcomed. Taken as a whole, they pave the way for the trustworthy use of corresponding applications. At the same time, however, their lack of clarity also poses great challenges for practitioners in terms of implementation and monitoring in individual cases. Standardisation should play a "key role" here. The following section will therefore critically examine whether the great importance of harmonised technical standards in the draft AI Regulation is a convincing solution.

3. The role of "harmonised standards" (Art. 40 of the draft AI Regulation) and "common specifications" (Art. 41 of the draft AI Regulation)

„Harmonised standards" are harmonised European standards as defined in Art. 2 (1) c) of the "Regulation on European standardisation"[90] (Art. 3 no. 27 of the draft AI Regulation). Harmonised standards are technical specifications adopted by an officially recognised standardisation body. They aim to be applied repeatedly or continuously. Whilst their adherence is not mandatory, they have been adopted on the basis of a Commission's mandate to apply Union harmonisation legislation. A "common specification", on the other hand, is a document, other than a standard, containing technical solutions providing a means to comply with certain requirements and obligations established under this Regulation, Art. 3 no. 28 of the draft AI Regulation. According to Art. 41 (1) of the draft AI Regulation, the European Commission may adopt such specifications, if the Commission considers the relevant harmonised standards to be insufficient or if there is a need to address specific

[90] Regulation (EU) no 1025/2012 of the European Parliament and of the Council of 25 October 2012 on European standardisation, amending Council Directives 89/686/EEC and 93/15/EEC and Directives 94/9/EC, 94/25/EC, 95/16/EC, 97/23/EC, 98/34/EC, 2004/22/EC, 2007/23/EC, 2009/23/EC and 2009/105/EC of the European Parliament and of the Council and repealing Council Decision 87/95/EEC and Decision No 1673/2006/EC of the European Parliament and of the Council (OJEU 2012 L 316, 12).

safety or fundamental right concerns.[91] AI Systems in conformity with either harmonised standards or with common specifications shall be presumed to be in conformity with those requirements covered by the respective standards or specifications, cf. Art. 40, Art. 41 (3) of the draft AI Regulation. This illustrates the "outstanding position of harmonised technical standards"[92]. According to the Commission, as noted, standardisation should "play a key role to provide technical solutions to providers to ensure compliance with this Regulation".[93] In view of the lack of specifications of the requirements for high-risk AI systems within the proposed regulation itself, these should therefore – according to the legislative will expressed in the legislative materials – be reserved for further definition as part of harmonised technical standards.

Such a far-reaching outsourcing of regulatory powers to private actors, however, raises considerable concerns with regard to democratic principle:[94] It is, in principle, the task of the respective democratically legitimised legislator to sufficiently concretise abstract requirements. Once a provision has been substantiated to a certain degree, it may – depending on the context of application – also be permissible and, within a certain framework, even reasonable to make further specifications through standardisation processes. Outsourcing the specification of abstract requirements in their entirety, however, is likely to exceed the scope of legally permissible delegation[95]; this applies all the more when taking into account the lack of judicial review as a corrective under the rule of law.[96] The requirements imposed on AI systems within the regulation therefore need to be sufficiently substantiated within the regulatory framework itself, be it in the text of the regulation or (at least) in the recitals. In particular, the issues raised above regarding the requirements for high-risk AI systems pursuant to Art. 8 et seqq. of the draft AI regulation would have to be addressed and dealt with by the European legislator itself. The adequately specified provisions could then be further substantiated and standardised as part of possible standardisation processes; only under the premise, of course, that a diverse and representative composition of the standardisation committees is guaranteed. Otherwise, there's a risk of helping particular interests to a certain legal standing in a way that is questionable under the rule of law. It therefore seems necessary to establish binding guidelines for the composition of standardisation committees

[91] DIN/DKE, in their joint position paper, call for the discarding of such specifications as part of the draft AI Regulation, see *DIN/DKE*, Position Paper on the EU "Artificial Intelligence Act", 2021, p. 4 (available at: https://www.din.de/resource/blob/800324/ c50ed443e81c47f8860b3f5c2b3b0742/21-06-din-dke-position-paper-artificial-intelligence-act-data.pdf, last accessed on: 07.08.2022).
[92] *Spindler*, CR 2021, 361 (369).
[93] COM(2021) 206 final, p. 32, recital 61.
[94] In agreement *Ebers/Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528 (532).
[95] Even if the ECJ has so far apparently avoided explicitly speaking of a "delegation" of legislative powers with regard to standardisation, there is no doubt about the legal effect of harmonised standards, cf. on this in detail *Ebers*, RDi 2021, 588 (593 et seq.).
[96] Cf. for more details in this regard *Ebers*, RDi 2021, 588 (595 et seq.).

and for the standardisation processes themselves, which, in addition to taking economic interests into account, ensure adequate inclusion of the consumer perspective.[97]

There is another reason to reject the concretization of the requirements of the draft AI regulation by means of standardisation: In this case, concretization does not only include technical regulation, which is the core area of standardization. Instead, as we could see, concretization means in particular the clarification of legal and ethical terms, regarding, for example, the sufficient degree of transparency and the understanding of fairness. These questions are essential for European citizens and their fundamental rights and values and should therefore be answered by the European legislative bodies themselves.

Strengthening the role of common technical specifications does not solve the lack of democratic legitimacy. It is true that the European Commission and therefore a central European legislative body is responsible for drafting the specifications. However, it is left to the Commission's discretion to establish technical standards ("the Commission may", cf. Art. 41 (1) of the draft AI Regulation). Thus, there are no binding and detailed provisions for when and how this has to be carried out. Anyways, a subsequent specification does not appear to be very expedient even with the establishment of binding framework conditions, since the Commission should – according to the position taken by this text – make a specification already in the regulatory framework itself.

V. Sufficient consumer protection in the context of AI systems with low (cf. Art. 52 of the draft AI Regulation) and minimal (cf. Art. 69 of the draft AI Regulation) risk?

1. Low-risk AI systems (cf. Art. 52 of the draft AI Regulation)

AI systems that "pose specific risks of impersonation or deception"[98] and are therefore not harmless from a consumer's perspective are subject to certain information and transparency obligations pursuant to Art. 52 of the draft AI Regulation.[99] While the Commission catalogizes them as low-risk AI systems,[100] other voices within the scientific community classify them as medium-risk AI systems.[101] Of course, overlaps are certainly possible. For this reason, a high- risk AI system can sometimes also be subject to the requirements of Article 52 of the draft AI Regulation. This follows from the provision of Article 52 (4) of the draft

[97] The fact that standardisation practice is in some cases anything but representative in the composition of the relevant committees is demonstrated by an evaluation of the composition of 62 of a total of 69 committees of the DIN, which is structured as follows (the values in each case represent the median of the proportion of seats): Business (75.5 %), science and research (6.5 %), users (6 %), public sector (4.5 %), occupational health and safety (1 %), consumer protection, trade unions, NGOs, environmental protection (0 %), *Huggins*, in: Verantwortung und Recht, Tagungsband der 62. Tagung Junges Öffentliches Recht, 2022, 315 (320); cf. further regarding the insufficient opportunities for participation of stakeholders, and the possible reasons, *Ebers*, RDi 2021, 588 (595).
[98] COM(2021) 206 final, p. 34, recital 70.
[99] The following explanation of AI systems within the meaning of Art. 52 of the draft AI Regulation is based on *Rostalski/Weiss*, ZfDR 2021, 329 (350 et seqq.).
[100] COM(2021) 206 final, p. 12.
[101] Cf. etwa *Bomhard/Merkle*, RDi 2021, 276 (282).

AI Regulation, according to which its paragraphs 1-3 shall not affect the obligations with regard to high- risk AI systems referred to in Title III.

a. AI systems intended to interact with natural persons

Art. 52 (1) of the draft AI Regulation firstly covers AI systems that are intended to interact with natural persons. This means that they must be designed and developed in such a way that natural persons are informed that they are dealing with an AI system. Exceptionally, this obligation may be waived if this is obvious due to the circumstances and context of use. The same goes for AI systems that are authorized by law to detect, prevent, investigate and prosecute criminal offences, unless these systems are available for the public to report a criminal offence.

These provisions allow a specific regulation with regard to so-called "chatbots" (applications capable of engaging in (online) conversation).[102] „Chatbots" can be used not least to distort the public opinion and to spread fake news.[103] Transparency with regard to the fact that one is dealing with a "chatbot" can make it easier for the human "conversation partner" to correctly grasp the setting of the conservation and the content of the statement made by the bot. A lack of such transparency means a considerable risk to the autonomy of the person concerned. Not least, this can have a negative impact on the democratic decision-making process. Against this background, the information obligation proves to be a sensible way of safeguarding consumer rights. It nevertheless seems preferable to substantiate in greater detail the possible exception made in obvious cases.[104] One possible regulatory technique could be to include examples in the recitals such as certain smart devices or apps.[105]

The blanket exemption from the transparency obligation of those AI systems that are authorized by law to detect, prevent, investigate and prosecute criminal offences (unless they are available for the public to report a criminal offence) causes doubts. It is based on the – principally correct – consideration that criminal prosecution is of particular social importance. Against this backdrop, it principally seems acceptable to set aside certain conflicting interests, such as the information that one's interlocutor is an AI system. This weighing of interests, however, does not prove to be justified without any exception. Instead, the area of application as well as the degree of suspicion against the person subjected to the respective measure can have an important influence on the balancing of individual against public interests. For example, the use of "chatbots" for crime prevention or investigation without a concrete suspicion proves to be particularly problematic with regard to the autonomy of the person affected. Thus, the European legislator should provide further specifications in this regard. While it is true that AI systems must be authorized to this effect

[102] *Kalbhenn*, ZUM 2021, 663 (669); *Engelmann/Brunotte/Lütkens*, RDi 2021, 317 (321).
[103] *Kalbhenn*, ZUM 2021, 663 (669).
[104] Critical in this regard *Engelmann/Brunotte/Lütkens*, RDi 2021, 317 (321).
[105] *Bomhard/Merkle*, RDi 2021, 276 (282), argue in favor of exempting the use of applications.

"by law", this caveat is not in itself a sufficient guarantee that the freedom of self-determination of a person coming into contact with such measures is adequately protected.

b. Emotion recognition systems or biometric categorization systems

Pursuant to Article 52 (2) of the draft AI Regulation there is also an obligation to inform affected subjects about emotion recognition systems or biometric categorization systems. According to Art. 3 no. 34 of the draft AI Regulation, an "emotion recognition system" is an AI system that serves the purpose of identifying or inferring emotions or intentions of natural persons on the basis of their biometric data. An "biometric categorization system" serves the purpose of assigning natural persons to specific categories, such as sex, age, hair colour, eye colour, tattoos, ethnic origin or sexual or political orientation on the basis of their biometric data, according to Art. 3 No. 35 of the draft AI Regulation. Again, AI systems that are authorized for the purposes of law enforcement are not covered by these requirements.

In this context, too, the information obligation serves to protect the autonomy of the person coming into contact with the system. In this way, they are enabled to decide freely and in knowledge of the relevant parameters whether they want to use such an AI system or not. However, it is difficult to see why the Commission's proposal places such systems in the category of merely low risk. In particular, the detection of emotional weaknesses by means of AI technology has a potential for abuse that should not be underestimated.[106] This has led not least the European Data Protection Supervisor and the European Data Protection Board to call, in a joint opinion, for a general ban on such systems, with narrowly defined exceptions with appropriate safeguards solely for health or research purposes.[107]

This demand can be endorsed. The current version of the draft AI Regulation does not, for example, prevent the use of emotion recognition software in schools, as long as it is disclosed to the parties involved. However, considerable risk to the individual self-determination of young people arise as a consequence, in particular where such software is used in assessing the individual performance of pupils. Even if the technology works flawlessly and does not demonstrate any gender- or ethnicity-based biases, such technology poses a great risk to the personal development of those concerned. Young people who would constantly have to reckon with an analysis of their emotions and the use of corresponding measures[108] could no longer express themselves freely. Their personal development would be impaired

[106] Cf. about this and more generally about the dangers posed by such systems *Stenner*, Berechnete Gefühle, 2.7.2021, available at: https://netzpolitik.org/2021/emotionale-ki-berechnete-gefuehle/, last accessed on: 07.08.2022.
[107] *EDPB/EDPS*, Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act), dated 18.06.2021, p. 12 (available at: https://edpb.europa.eu/system/files/2021-06/edpb-edps_joint_opinion_ai_regulation_en.pdf, last accessed on: 07.08.2022).
[108] Cf. on this potential context of application *Stenner*, Berechnete Gefühle, 2.7.2021, available at: https://netzpolitik.org/2021/emotionale-ki-berechnete-gefuehle/, last accessed on: 07.08.2022.

to an unacceptable extent. The reason is that being aware of the use of such technologies in school performance assessment will ultimately lead to conformist patterns of behaviour. Fatigue, boredom or lack of interest, for example, would be immediately recognized by appropriate software and could be included in the performance evaluation. Pertinent examples are known from abroad.[109] Individual pupils are consequently put under enormous pressure to perform. Individual freedom, not least of an inward-looking nature, are increasingly squeezed out of everyday school life. This especially applies if there is an ideal of a non-stop recording of pupils not only in terms of their outwardly presented performance, but also as learning persons with a corresponding inner readiness and motivation for permanent learning. Individuality, however, is a value in itself. It is fostered by freedom, especially in education and upbringing. Schools must not be allowed to degenerate into places where human beings and their individual way of existing are subjected to constant surveillance. Allowing the all-encompassing technical recording of the pupils and their inner personality, revealed by the emotions visible on their face, is a transgression that is unjustifiable in a liberal society under the rule of law. At the bottom the control of the individual as a mere object of state (here: educational) goals is incompatible with this type of societal constitution. These considerations also hold true in the context of the workplace. Employees do not owe their innermost feelings to their employer either. Emotions, and what they say about the thoughts of whoever experiences them, should therefore be excluded from the systematic recording made possible by AI systems. In the further legislative process, the classification of such AI systems should therefore be reviewed. A much more restrictive approach to this type of technology is needed in order to prevent possible abuse and to safeguard fundamental European values. Simply imposing an obligation to provide information is by no means sufficient for this purpose.

c. AI systems used to generate "deep fakes"

Art. 52 (3) of the draft AI Regulation contains a special provision for AI systems that generate or manipulate image, audio or video content that resembles existing persons, objects, places or other entities or events and falsely appear to a person to be authentic or truthful (so-called "deep fakes"). Users of such AI systems will be required to disclose that the respective content has been artificially generated or manipulated. An exception is made for the area of law enforcement and the exercise of freedom of expression, art and science, insofar as this is necessary and safeguards for the rights and freedoms of third parties are in place.

It is surprising that the draft AI Regulation categorizes AI systems used to generate deep fakes as only low-risk AI systems. Deep fakes can be abused in a great variety of ways. The possible dissemination of misinformation, with great outreach and effect, is particularly noteworthy. One need only think of deep fakes of politicians or other public figures. They

[109] Cf. on the use of such systems in China https://www.deutschlandfunk.de/alles-unter-kontrolle-chinas-intelligenter-schule-entgeht-102.html, last accessed on: 07.08.2022.

can be discredited and defamed in a particularly drastic way through the use of deep fakes.[110] As the software for creating deep fakes is freely available, their creation usually requires little time.[111] This increases the risk of disinformation campaigns by means of this technology, which harbour an enormous potential danger for individual autonomy, the liberal democratic order and other goods and interests that can be affected as a result of the disinformation. The media landscape can be flooded with deep fakes as part of targeted misinformation campaigns.[112] In addition, the widespread use of this technology can lead to people no longer considering videos or sound recordings a reliable source and generally distrusting information.[113] Such developments prove particularly risky in a "media society" based on images and pictorial information: a loss of trust in the media goes hand in hand with the imminent risk of a loss of trust in political and social actors. As trust is the "grease" of a free society, potential attacks on it must be taken particularly seriously, not least from a legal perspective.

The draft AI Regulation considers the protection of human autonomy to be particularly relevant. With that in mind, it is difficult to understand why AI systems designed to generate content so dangerous to autonomy are classified only as part of Art. 52 of the draft AI Regulation. In contrast, there are occasional calls to generally prohibit all AI systems generating deep fakes due to their disruptive potential, with exceptions only within narrow limits. The latter refer, for example, to the area of law enforcement or the exercise of artistic and scientific freedom.[114] Whether a general ban with such narrowly defined exceptions presents an adequate solution should be the subject of a critical social debate. A discussion is needed on the question in which specific fields of application the use of deep fakes should be permissible, and in which it should be banned altogether.[115] In any case, an absolute ban regardless of the respective context of use appears to be disproportionate.

2. AI systems of minimal risk (cf. Art. 69 of the draft AI Regulation)

If an AI system does not fall into one of the specifically listed categories, the draft AI Regulation classifies it as posing only a "minimal risk"[116]. These applications are not subject to any additional legal requirements. On the contrary, the proposal emphasizes the possibility for providers or their stakeholders to voluntarily adhere to so-called "codes of conduct" in the area of such AI systems, Art. 69 of the draft AI Regulation. These may refer to the

[110] *Ebert/Spiecker genannt Döhmann*, NVwZ 2021, 1188 (1191 et seq.); cf. for an overview of the issue of so-called "Deepfakes" and their legal assessment *Thiel*, ZRP 2021, 202 et seqq.
[111] *Thiel*, ZRP 2021, 202 (203).
[112] *Linardatos*, GPR 2022, 58 (68 et seq.); see also *Chesney/Citron* California Law Review 107 (2019), 1753 (1777 et seq.); *Wilkerson* Missouri Law Review 86 (2021), 407 (410 et seqq.).
[113] *Thiel*, ZRP 2021, 202 (203).
[114] As stated by *Linardatos*, GPR 2022, 58 (68 et seq.).
[115] See for potential ways to differentiate *van Huijstee et al.*, Tackling deepfakes in European policy, Study for the Panel for the Future of Science and Technology, European Parliamentary Research Service, 2021, p. 61 (available at: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/690039/EPRS_STU(2021)690039_EN.pdf, last accessed on: 07.08.2022).
[116] COM(2021) 206 final, p. 12.

compliance with requirements imposed on high-risk AI systems (Art. 69 (1) of the draft AI Regulation), but also to the achievement of other objectives, such as environmental sustainability, accessibility for persons with a disability, stakeholders' participation in the design and development of the AI systems and diversity of development teams (Art. 69 (2) of the draft AI Regulation). The establishment of corresponding codes of conduct is to be "encouraged and facilitated". According to the EU Commission, "the vast majority of AI systems currently used in the EU"[117] falls into this category.

Promoting and facilitating the establishment of codes of conduct with regard to AI systems is proving to be particularly significant. It is in the interests of consumers to offer providers of AI systems incentives for the most far-reaching voluntary commitments possible. However, it is not yet clear, how exactly the providers are to be motivated for this project. In particular, there are no concessions equivalent to those offered by Art. 40 in conjunction with Art. 28 (1), (4), 24 (3), 32 (3) of the General Data Protection Regulation for providers who have joined codes of conduct.[118] It therefore seems advisable to include concrete incentives for the establishment of such codes in the regulation in the course of the further legislative process.

C. Analysis of the conformity assessment procedures proposed in the draft AI Regulation

I. Conformity assessments in the sense of Art. 43 of the draft AI Regulation

Pursuant to Art. 43 of the draft AI Regulation, AI systems in the high-risk category are to undergo an audit to assess and certify their conformity with the requirements already mentioned in Art. 8 et seq. of the draft AI Regulation.

1. Introduction: Presentation of the envisaged different assessment procedures

In principle[119], the draft provides for two different assessment procedures: external control carried out by a notified body in the sense of Art. 33 (1) of the draft AI Regulation on the one hand; internal control by the provider itself on the other hand.

a. Internal control

The internal control is described more specifically in Annex VI. According to this, the provider is subject to three obligations: 1. He must check whether the existing quality management system complies with the requirements of Art. 17 of the draft AI Regulation. 2. He

[117] *European Commission*, New rules for Artificial Intelligence – Questions and Answers, 21.4.2021, available at: https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683, last accessed on: 07.08.2022.
[118] *Spindler*, CR 2021, 361 (371); *Geminn*, ZD 2021, 354 (359).
[119] Cf. for the exception in the area of high-risk AI systems in accordance with Art. 6 (2) in conjunction with Annex III No. 5 b) of the draft AI Regulation which are placed on the market or put into service by credit institutions within the meaning of Directive 2013/36/EU, the regulation in Art. 43 (2) sentence 2 of the draft AI Regulation.

must examine the information contained in the technical documentation in order to assess whether the AI system complies with the relevant essential requirements in Title III, Chapter 2 (Art. 8 et seq. of the draft AI Regulation). 3. Finally, he must also evaluate whether the design and development process of the AI system and its post-market observation in accordance with Art. 61 of the draft AI Regulation are in conformity with the technical documentation.

b. Involvement of a notified body

The procedure involving a notified body is described in more detail in Annex VII. It consists of several components. On the one hand, the respective approved quality management system in the sense of Art. 17 of the draft AI Regulation is examined by the notified body (cf. in detail Annex VII No. 3 of the draft AI Regulation) and monitored in regular audits – if necessary in combination with additional tests (cf. in detail Annex VII No. 5 of the draft AI Regulation). In addition, the notified body checks the technical documentation according to strict regulations (cf. in detail Annex VII No. 4 of the draft AI Regulation). For this purpose, the notified body is quite powerful. For example, the notified body must be granted unrestricted access to the training and test data sets used by the provider, also via application programming interfaces (API) and other means and instruments suitable for remote access, Annex VII No. 4.3 of the draft AI Regulation. The notified body is also entitled to request further evidence if necessary and to test the AI system independently, Annex VII No. 4.4 of the draft AI Regulation. Finally, the notified body must even be granted access to the source code of the AI system upon justified request, if this is necessary for the conformity assessment, Annex VII No. 4.5 of the draft AI Regulation.

c. Interim result

Both procedures are based on an examination and assessment of the quality management system and the technical documentation of the respective AI system. Together, these two components can enable a comprehensive risk analysis and an examination of compliance with the requirements of Art. 8 et seq. of the draft AI Regulation. Against this background, it is basically convincing to orientate audit procedures towards these two instruments.

2. Notified bodys in the sense of Art. 3 No. 22 of the draft AI Regulation

It must be clarified to whom the tasks of a notified body may be transferred and under what conditions. The notified body is legally defined in Art. 3 No. 22 of the draft AI Regulation as "a conformity assessment body designated in accordance with this Regulation and other relevant Union harmonisation legislation". According to Art. 3 No. 21 of the draft AI Regulation, a conformity assessment body is "a body that performs third-party conformity assessment activities, including testing, certification and inspection", whereby the conformity assessment activity according to Art. 3 No. 20 of the draft AI Regulation extends to verification of compliance with the requirements of Art. 8 et seq. of the draft AI Regulation.

a. Legal requirements and designation procedures

The legal requirements for notified bodies are explained in detail in Art. 33 (2) to (11) of the draft AI Regulation. A special focus is placed on the sufficient qualification as well as objectivity and independence of such bodies. Both criteria are addressed in several paragraphs.

In order to ensure sufficient qualification, the notified bodies must first of all "satisfy the organisational, quality management, resources and process requirements that are necessary to fulfil their tasks", Art. 33 (2) of the draft AI Regulation. In addition, according to Art. 33 (3) of the draft AI Regulation, the internal processes must be designed "such as to ensure that there is confidence in the performance by and in the results of the conformity assessment activities that the notified bodies conduct". The procedures for carrying out conformity assessments must allow for adequate consideration of the particularities of each individual case, cf. Art. 33 (7) of the draft AI Regulation. Finally, notified bodies "shall have procedures for the performance of activities which take due account of the size of an undertaking, the sector in which it operates, its structure, the degree of complexity of the AI system in question" (Art. 33 (9) of the draft AI Regulation), and at the same time have "sufficient internal competences to be able to effectively evaluate the tasks conducted by external parties on their behalf" (Art. 33 (10) of the draft AI Regulation).

The independence of the notified body must be maintained both in relation to the provider of an assessed AI system, and in relation to all other actors who have an economic interest in the assessed AI system, Art. 33 (4) of the draft AI Regulation. Documented structures and procedures regarding the organisation and functioning must guarantee the "independence, objectivity and impartiality" of the respective body, Art. 33 (5) of the draft AI Regulation.

Lastly, Article 33 (6) of the draft AI Regulation specifies the requirements for maintaining the confidentiality of the information provided, and Article 33 (8) of the draft AI Regulation stipulates a basic obligation for an appropriate liability insurance.

In order to be notified as a conformity assessment body, an application must be submitted to the notified authority of the Member State in which the applicant body is established, according to Art. 31 (1) of the draft AI Regulation. Pursuant to Art. 3 No. 19 of the draft AI Regulation, a notified authority is the national authority responsible for setting up and carrying out the necessary procedures for the assessment, designation and notification of conformity assessment bodies and for their monitoring. This authority is also subject to extensive requirements regarding its sufficient qualification (cf. Art. 30 (4), (7), (8) of the draft AI Regulation) as well as independence and objectivity (cf. Art. 30 (3), (5) of the draft AI Regulation). Pursuant to Art. 31 (2) of the draft AI Regulation, the application for notification shall be accompanied by a description of the conformity assessment activities, the

conformity assessment module or modules and the AI technologies for which the conformity assessment body claims to be competent, as well as by an accreditation certificate, if one exists, issued by a national accreditation body attesting that the conformity assessment body fulfils the requirements laid down in Art. 33 of the draft AI Regulation. If a corresponding accreditation certificate is not available, the notified authority must be provided with the documentary evidence necessary for the verification, recognition and regular monitoring of its compliance with the requirements of Art. 33 of the draft AI Regulation in accordance with Art. 31 of the draft AI Regulation. The notified authority shall notify the applicant conformity assessment body if it fully complies with the requirements of Art. 33 of the draft AI Regulation, Art. 32 (1) of the draft AI Regulation.

b. Evaluation of the concept

The abstract regulations regarding the requirements to be fulfilled by notified bodies as well as the respective designation procedures are to be welcomed from a consumer's point of view. As shown, the regulation includes both an adequate procedure and stipulations to guarantee the qualification, independence and objectivity of the bodies involved. If these requirements are met, the intended protection of European fundamental rights and values can be achieved in the best possible way. However, it remains to be observed to what extent these abstract requirements can be verified and implemented in practice in a workable and at the same time sufficiently reliable manner.

3. Assessment procedures with regard to "high-risk AI systems"

The concrete form of the audit required by Art. 43 of the draft AI Regulation depends on the specific nature of the respective high-risk application.

a. Conformity assessment of AI systems within the meaning of Art. 6 (1) of the draft AI Regulation

In the case of AI systems in the sense of Art. 6 (1) of the draft AI Regulation, which are subject to conformity tests by third parties in accordance with the relevant harmonisation legislation, the test is to be integrated into these conformity assessment procedures, Art. 43 (3) of the draft AI Regulation. However, it is necessary that the respective competent bodies fulfil the requirements of Art. 33 (4), (9) and (10) of the draft AI Regulation, which apply to notified bodies in the sense of the draft AI Regulation. By integrating the verification into the respective conformity assessments, the aim is to "minimise the burden on operators and avoid any possible duplication"[120].

[120] COM(2021) 206 final, p. 32, recital 63.

b. Conformity assessment of AI systems in the sense of Art. 6 (2) of the draft AI Regulation

With regard to high-risk systems in the sense of Art. 6 (2) of the draft AI Regulation, a distinction is made between two independent assessment procedures. On the one hand, the provider himself is allowed to carry out the internal control, Annex VI of the draft AI Regulation. On the other hand, the AI system can be assessed by a notified body, Annex VII of the draft AI Regulation. It depends on the sector of the respective high-risk application, which assessment procedure is to be carried out in each individual case.

In the case of an AI system from the sector "Biometric identification and categorisation of natural persons" in the sense of Annex III No. 1 of the draft AI Regulation, a distinction must be made as to whether harmonised standards according to Article 40 of the draft AI Regulation or, where applicable, common specifications according to Article 41 of the draft AI Regulation exist and have been applied. If such technical specifications have been fully applied by the provider of the AI system, he can choose between an internal control according to Annex VI and an assessment with the participation of a notified body according to Annex VII (Art. 43 (1) sentence 1 of the draft AI Regulation). However, if such requirements do not exist or if the provider has not applied them or has only applied them in part, the conformity assessment procedure must be carried out with the participation of a notified body, Art. 43 (1) sentence 2 of the draft AI Regulation.

If, on the contrary, the respective AI system originates from one of the sectors in the sense of Annex III No. 2 to 8, it only requires an internal control in accordance with Annex VI pursuant to Art. 43 (2) of the draft AI Regulation. However, this does not apply to AI systems for checking creditworthiness in the sense of Annex III No. 5 b) of the draft AI Regulation. In this respect, the verification of conformity with the AI Regulation is integrated into the conformity assessment according to Art. 97 to 101 of the European "Directive on access to the activity of credit institutions and the prudential supervision of credit institutions and investment firms"[121].

c. Critical analysis of the (different) modes of assessment procedures

The current draft AI Regulation can thus be summarised as follows with respect to conformity assessments of high-risk AI systems:

[121] Directive 2013/36/EU of the European Parliament and of the Council of 26 June 2013 on access to the activity of credit institutions and the prudential supervision of credit institutions and investment firms, amending Directive 2002/87/EC and repealing Directives 2006/48/EC and 2006/49/EC (Document 32013L0036).

aa. Principally mandatory involvement of external conformity assessment bodies for high-risk AI systems in the sense of Art. 6 (1) of the draft AI Regulation

In the case of AI systems as defined in Art. 6 (1) of the draft AI Regulation, which are already subject to conformity assessments by third parties according to the relevant harmonisation legislation, the assessment with regard to the requirements of the draft AI Regulation is integrated into these conformity assessment procedures. For product-related high-risk AI systems, an assessment by an external party is therefore mandatory in principle. An exception exists according to Art. 43 (3) subparagraph 3 of the draft AI Regulation only in the case that the legal acts listed in Annex II Section A allow the provider of the product to dispense with a conformity assessment by third parties, provided that this provider has applied all harmonised standards covering all relevant requirements; however, the provider may only make use of this option if he has also applied harmonised standards or, where applicable, common specifications according to Art. 41 of the draft AI Regulation, which cover the requirements of Art. 8 et seqq. of the draft AI Regulation. This exception is in line with the framework regarding stand-alone AI systems. It can be supported from this perspective, and will be discussed in more detail there.[122]

bb. Optional involvement of conformity assessment bodies in case of compliance with all harmonised technical standards for high-risk AI systems pursuant to Art. 6 (2) in conjunction with Annex III No. 1 of the draft AI Regulation (Art. 43 (1) sentence 2 of the draft AI Regulation)

As things stand at present, the involvement of a notified conformity assessment body in the area of stand-alone high-risk AI systems is only mandatory in the case of such systems as defined in Annex III No. 1 of the draft AI Regulation, i.e. systems in the sector of "biometric identification and categorisation of natural persons", and also only in the case of the lack of existence or the lack of compliance with harmonised technical standards (Article 43 (1) sentence 2 of the draft AI Regulation). Should a provider fully comply with technical standards, it is at its discretion to involve a certification body in the AI systems affected by this (Art. 43 (1) sentence 1 of the draft AI Regulation). The same applies, as already shown, according to Art. 43 (3) subparagraph 3 of the draft AI Regulation for product-related high-risk AI systems. This kind of regulatory concept is also applied in other contexts as, for example, in the provision of § 4 of the Neunte Verordnung zum Produktsicherheitsgesetz (Maschinenverordnung) (conformity assessment procedure for machinery). This stipulates that a machine which is listed in Annex IV of Directive 2006/42/EC and which is manufactured according to the harmonised standards mentioned in § 3 (5) of the Maschinenverordnung can be subjected to various optional conformity procedures by the manufacturer or his authorised representative. The prerequisite is, of course, that the standards take into account all relevant essential health and safety requirements. If this is also the case, the manufacturer has the choice between the conformity assessment procedures in Annexes

[122] See the following comments on this subject: C. I. 3. c. bb.

VIII-X of Directive 2006/42/EC. Annex VIII of Directive 2006/42/EC includes a conformity assessment without the involvement of an external body.[123] This option is not available for the manufacturer or his authorised representative if the machinery is listed in Annex IV of Directive 2006/42/EC while the harmonised standards in question have not been taken into account or have only been partially taken into account during its manufacture, or if these standards do not meet all the relevant essential health and safety requirements or if there are no harmonised standards for the machinery in question. Under this condition, the conformity assessment procedures must be carried out by an external body on the part of the manufacturer or his authorised representative.

This regulatory concept is convincing insofar as compliance with all harmonised technical standards as a concretisation of the abstract requirements by the provider offers sufficient guarantee for their observance in principle. It is true that an additional certification is always in the interest of consumers, since in this way a further guarantee for the actual compliance with the applicable standards is provided. However, it must be taken into account that the content of the assessment is limited to the aspects that are the subject of the relevant harmonisation standards. Insofar as the provider ensures compliance with them through its own testing, this interest is already considered. Admittedly, this leaves the (residual) risk that the provider does not comply with the rule by only insufficiently testing AI system. This, of course, can have negative effects on consumer rights. However, it must be taken into account that merely claiming compliance with harmonised technical standards entails a massive liability risk for the provider. The liability provisions provided for in the draft AI Regulation are of considerable weight. For example, Art. 71 of the draft AI Regulation provides that the Member States shall adopt sanctioning provisions – such as fines – which respond to infringements of the Regulation and are "effective, proportionate, and dissuasive". Depending on the respective infringement, sanctions of up to 30 million euros or 6% of the total worldwide annual turnover of the previous business year are conceivable. In addition, an independent civil liability regime for the use of artificial intelligence is to be expected, which will further tighten the legal situation for providers of such applications.[124] Compliance with the harmonised technical standards therefore avoids serious sanctions and is therefore associated with a high incentive for the responsible persons. However, on this basis the best possible way to protect consumer rights is to adequately reflect them in harmonized norms and standards. As already shown above, this presupposes at least sufficient

[123] Whether the option of internal control for high-risk machinery will be retained in the future seems at least questionable in view of a proposal for a Regulation from the European Commission of 21 April 2021. It states explicitly: „[This policy] removes the internal check option for the conformity assessment of the high-risk machines, and ensures full coherence with the AI Regulation proposal", COM(2021) 202 final, p. 5. The plea for deleting the option of internal controls while at the same time referring to the full alignment with the AI Regulation is surprising insofar as in the latter the option of internal controls for a large number of stand-alone high-risk AI systems within the meaning of Art. 6 (2) of the draft AI Regulation is opened. It remains to be seen, therefore, how the further legislative process on this matter will unfold.

[124] European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)); for this regulatory proposal, see the overview at *Zech*, NJW-Beil. 2022, 33 (36 et seqq.).

involvement of this interest group in the process of developing norms and standards. As explained above, it seems important to include specifications of particulary relevant requirements already in the European regulation. Furthermore, monitoring procedures offer additional protection for the preservation of specifications that also protect consumers.[125] We therefore classify them as an important component in the context of ensuring the conformity of AI systems.[126] Last but not least, the provisions made in Art. 65 of the draft AI Regulation serve to further safeguard consumer rights. The provision stipulates that the respective national market surveillance authority must verify compliance with the requirements and obligations set out in the regulation if it has sufficient grounds to assume that an AI system poses a risk to the health or safety or the protection of the fundamental rights of persons. The market surveillance authority has various obligations in this respect, such as informing competent authorities and bodies and requiring the person responsible for the AI system to bring it into conformity with the regulation, to withdraw it from the market or to recall it within a period of time appropriate to the nature of the risk. Corresponding steps are to be taken by the market surveillance authority itself if the responsible party does not comply. This intervention-intensive external control also constitutes an additional safeguarding of consumer rights, which justifies the regulatory regime provided so far with regard to the performance of conformity assessments.

cc. Relevance of assessment procedures for other stand-alone AI systems

The vast majority of stand-alone AI systems in the high-risk category only require internal control by the providers themselves. The Commission justifies this with "the early phase of the regulatory intervention and the fact the AI sector is very innovative and expertise for auditing is only now being accumulated"[127]. Against this background, the impression may be given that conformity assessment bodies are not of increased importance according to the regulatory proposal. However, it must be countered that the differentiation of conformity assessment procedures for AI systems with reference to a product that is subject to harmonisation standards and for stand-alone AI systems is already doubtful in itself.[128] Thus, in both cases a high-risk application is involved, which, according to the Commission's assessment, poses significant risks to European fundamental rights and values. The Commission itself discloses that this envisaged differentiation is based less on factual differences than on predicted difficulties in implementation and, moreover, is merely of provisional character. On the other hand, "it is appropriate", according to the Commission's assessment, "to limit, at least in an initial phase of application of this Regulation, the scope of application of third-party conformity assessment for high-risk AI systems other than those related to products"[129]. The statements make it clear that the (optional) competence

[125] See the comments under B. IV. 3.
[126] See the comments under C. II.
[127] COM(2021) 206 final, p. 14.
[128] The following statements are based on *Rostalski/Weiss*, ZfDR 2021, 329 (347 et seqq.).
[129] COM(2021) 206 final, p. 33, recital 64.

and thus the relevance of notified conformity assessment bodies in the area of stand-alone AI systems is to be expanded in the future.

There are good reasons for such an expansion. However, it is questionable whether a time-stretched adaptation of the assessment procedures is actually necessary. It is to be expected that the full application of the regulation will still be several years away – the European legislative process alone will take at least 18 to 24 months, whereby Art. 85 (2) of the draft AI Regulation provides for a transitional period of 24 months after the entry into force of the regulation. It is therefore likely that the regulation will not apply until 2025.[130] It is not too optimistic to assume that significant findings and experience in isolated AI certification can be achieved or gathered during this period. Corresponding projects, such as the "Zertifizierte KI"[131] (Certified AI) project in North Rhine-Westphalia, which we co-lead, or the "ExamAI – KI-Testing & Auditing"[132] (ExamAI – AI Testing & Auditing) project, have already begun their work. Dealing with the regulatory proposal and the requirements it sets out plays a central role in such projects. Against this backdrop, we suggest to directly entrust assessment bodies with the (optional) auditing of all stand-alone AI systems. This applies not least in view of the fact that the assessment bodies, which are responsible for the assessment of high-risk AI systems in the sense of Art. 6 (1) of the draft AI Regulation, are already in charge of this task, cf. Art. 43 (3) of the draft AI Regulation. These bodies may have extensive experience in examining the relevant harmonisation legislation. Nevertheless, the control of the requirements of the regulatory proposal is new territory for them. The fact that they are expected to have more experience with regard to AI systems at an earlier point in time than the notified bodies in the sense of the draft AI Regulation is therefore not convincing. Rather, equal treatment seems to be called for. At the very least, however, a clarification should be made regarding the provisional character of the different test procedures in the area of independent high-risk systems. The regulation and its recitals should make it even clearer that all high-risk systems within the meaning of Art. 6 (2) of the draft AI Regulation should be subject to an optional audit by notified conformity assessment bodies. Irrespective of any need for further specification, it can be stated that certification is already integrated as a central component in the concept of the draft regulation.

4. Extension of conformity assessments to low-risk AI systems (cf. Art. 52 of the draft AI Regulation)?

Conformity assessment procedures are only mandatory for certain high-risk AI systems in the current regulatory proposal. As explained, this is at least partly due to the fact that the corresponding test procedures are still in the development stage. Nevertheless, it can be expected that progress in this area will keep pace with further legislation, so that regulatory

[130] *Bomhard/Merkle,* RDi 2021, 276 (283).
[131] See the website at https://www.zertifizierte-ki.de/, last accessed on: 07.08.2022.
[132] See the website at https://www.rechtsinformatik.saarland/de/forschung/projekte/examai, last accessed on: 07.08.2022.

restraint is not necessary, at least for this reason.[133] Against this background, it must be asked whether conformity assessment procedures should not also be required to a certain extent with regard to other AI systems. The argument against this is that the obligation to carry out such assessments weighs heavily on the responsible party. It is therefore justified in principle to make such an increased obligation dependent on the criticality of the respective application. However, it must be taken into account that the classification of an application as a high-risk or no high-risk AI system does not always have to correspond to the real risk of the respective product. There are various conceivable reasons for this. On the one hand, it is possible that an AI system has been identified as particularly risky, but has not yet been included in Annex III of the draft AI Regulation. However, delays in this regulatory step can then have a significant negative impact on consumer rights. Especially since the current mandatory requirement in Art. 7 (1) a) of the draft AI Regulation of a reference to the areas already listed in Annex III could even completely block an expansion. Another reason why a system is wrongly not classified as a high-risk system may be that the associated risks have not yet been recognised or at least not adequately assessed. In our view, an example is provided by "deep fakes", which are currently only subject to the transparency obligation under Article 52 (3) of the draft AI Regulation. The associated risks for a free democracy do not seem to us to be adequately covered by the current draft regulation. This is not unusual, especially with new technologies. Which risks a society is prepared to accept under which conditions is always the result of (ongoing) negotiation processes. In this respect, AI applications pose a particular difficulty due to their dynamic mode of operation, which generally entails a lack of transparency for humans, which in turn can stand in the way of a realistic risk assessment.

On this basis, it can be considered to establish independent conformity assessment procedures as mandatory for other than high-risk AI systems. These can be procedures whose scope of testing is – if necessary, considerably – lower than the conformity assessment provided for high-risk systems. However, a corresponding regulation without further restrictions would have the consequence that ultimately every AI application would have to be subjected to a form of conformity assessment – and thus also such systems that pose a particularly low risk to consumers, if any at all. However, this would prove to be disproportionate. It must be taken into account that the providers of other AI systems are also subject to legal obligations with regard to the safety of the applications, the violation of which entails a considerable liability risk.

Nevertheless, it should be considered that the implementation of conformity assessment procedures has an advantage for the respective providers: In addition to the assurance of conformity with applicable law and possibly other ethical standards, this concerns in particular the competitive advantage associated with such procedures; this lies in the fact that consumers have a higher level of trust in tested products. Irrespective of this, the criticality

[133] For more detailed reasoning, see the comments under C. I. 3. c. cc.

of the respective application must be considered first and foremost in the debate on possible conformity assessment obligations. In this respect, however, it is noticeable that Art. 52 of the draft AI Regulation does contain AI systems that pose a particularly significant risk to the rights of consumers. Against this backdrop, an extension of the conformity assessments to AI systems within the meaning of Art. 52 of the draft AI Regulation seems worthy of consideration and ultimately welcome, unless the corresponding applications are upgraded to high-risk AI systems anyway. It would still have to be decided whether this additional obligation should apply to all of the systems mentioned in Art. 52 of the draft AI Regulation. This seems necessary at least for "deep fake" systems. Emotion recognition systems also have considerable risks, as explained earlier. In this respect, mere transparency cannot be a sufficient means to ensure the adequate protection of consumer rights. In this respect, a social debate is needed, which should be considered in the further legislative process.

By extending the audit obligations to systems in the sense of Art. 52 of the draft AI Regulation, the fact could be taken into account that such AI systems, as shown,[134] are also associated with considerable risks to European fundamental rights and values, which require additional safeguards in the form of conformity audits. In view of the fact that the requirements to be placed on such AI systems – in contrast to high-risk AI systems – are limited to mere transparency and information obligations, such an extension of the conformity assessments would not involve a disproportionate effort for the providers. Especially since they should also have the possibility to refrain from a review by third parties if they fully comply with the relevant harmonised technical standards.

5. New conformity assessment procedure in the case of a "substantial modification" in the sense of Art. 43 (4) of the draft AI Regulation

Because AI applications are often dynamic systems[135], this special feature must also be considered in the context of conformity assessments. The proposed regulation takes this into account in Art. 43 (4) of the draft AI Regulation. According to this, a particularity applies if an AI system subsequently modifies itself "substantially".

a. Presentation of the concept

This brings the term "substantial modification" into focus. Art. 3 No. 23 of the draft AI Regulation defines this as a change to the AI system following its placing on the market or putting into service which affects the compliance of the AI system with the requirements set out in Title III, Chapter 2 of this Regulation or results in a modification to the intended purpose for which the AI system has been assessed. Where these conditions are met, the conformity assessment procedure shall be repeated. This applies irrespective of whether the modified system is still to be placed on the market or whether it is to continue to be

[134] See already the comments on this subject under B. V. 1.
[135] See already the comments on this subject under A. II. 3.

used by the current user. However, a practice-relevant exception is to apply to self-learning systems that continue to learn after being placed on the market or put into service. Provided that the associated modifications were pre-determined by the provider at the moment of the initial conformity assessment and are contained in the information of the technical documentation in accordance with Annex IV No. 2 f) of the draft AI Regulation, they are not to be classified as a substantial modification.

b. Critical analysis and proposal for a concretisation of the requirements for carrying out a new conformity assessment

It has already been pointed out earlier that the dynamics of AI systems pose a challenge from a regulatory point of view. It is important not to undermine the potential inherent in the use of the technology from the outset by setting requirements that far exceed what can realistically be achieved by the providers or users. The very fact that later modifications often cannot be foreseen in advance corresponds to the special nature of self-learning systems.[136] The problem does not seem to be adequately covered in the current draft. Providers could be faced with insurmountable hurdles if they are required to repeatedly carry out test procedures, if necessary, at very short intervals. However, a solution does not lie in so-called "ad hoc conformity assessments", which are "carried out without human intervention"[137]. This is already not in line with an essential motive of the regulatory proposal, which is to protect personal autonomy from the risks posed by AI systems. Last but not least, Art. 14 of the draft AI Regulation requires effective human supervision with regard to high-risk AI systems. This basic value is not compatible with the idea of fully automated assessment procedures. In contrast, exceptions appear necessary, which should be included in the regulation in as concrete a form as possible. A regulation or explanation corresponding to the technological circumstances would be needed as to when changes in self-learning systems are still within the framework defined in advance by the provider.

The term "substantial modification" is only helpful at first glance. The purpose of an AI application must be clearly stated in its conformity assessment – objectives that deviate from this in the meantime are decisive for a new conformity assessment. But even in this respect, the devil is in the detail: depending on how openly the purpose of the system is formulated, a change of purpose cannot be easily identified afterwards. For example, the purpose of a hoover based on AI technology could be formulated as "interior cleaning". If, after some time, the technology not only sucks up dust and other tiny particles from the floor, but also moves up walls to do the same, this is undoubtedly still a process of interior cleaning. Nevertheless, it is difficult not to assume a change of purpose, since the cleaning process of hoovers is usually limited to the floor area of an interior. In this respect, a further interpretation of the previous purpose is required, which necessitates a ratio-oriented con-

[136] *Bomhard/Merkle*, RDi 2021, 276 (281).
[137] *Bomhard/Merkle*, RDi 2021, 276 (281).

sideration. It must be taken into account whether the previous conformity test was exclusively related to such risks, which can arise when very small particles are sucked in at the floor of an interior. If this is the case and if it was not checked whether further risks are created when the hoover moves up the walls (for example, falling onto people or the like), then everything speaks in favour of a relevant change of purpose.

Against this background, it may prove to be a feasible way to make the necessity of a renewed conformity assessment dependent on whether the changed or extended functionality of the AI system gives rise to risks that are so relevant that they themselves require an independent review – and this was not yet the subject of the previous conformity assessment. The change of purpose can be an indication of this, but it is not necessarily so. In this respect, too, it is necessary to consider whether the change of purpose creates relevant new risks that differ significantly from the previous scope of testing. It is therefore proposed that high-risk AI systems should always be subject to a new conformity assessment if significant risks to the goods and interests of humans arise from their use that were not covered by the previous scope of testing. The fact that potential risks of self-learning systems often cannot be fully determined in advance does not per se stand in the way of this proposal. Even if it should not be possible to identify all conceivable risks in peripheral areas, this does not apply equally to the respective core area of application of the AI system. In any case, the risks associated with an intended use can be identified – at least predominantly – on a context-specific basis. The renewed testing obligation can apply, as the Commission proposal also provides, regardless of whether the modified system is still to be placed on the market or whether it is to continue to be used by the current user. This also makes Article 43 (4) sentence 2 of the draft AI Regulation superfluous, since a relevant risk that arises from further learning of the system, but was already covered by the previous scope of testing, does not entail a new conformity assessment test according to sentence 1. The exemption provision of Art. 43 (4) sentence 2 of the draft AI Regulation in its current version proves to be problematic anyway. According to this, there is no substantial change and no renewed obligation to test is triggered if the provider had specified any changes to the AI system and its performance in advance at the time of the original conformity assessment and noted them in the information of the required technical documentation. This concept harbours a considerable risk of abuse: providers could decide to specify and note any change in advance, no matter how absurd, in order to avoid a renewed obligation to test. For the purposes of effective protection of fundamental rights, it therefore seems more appropriate to focus on whether the risks in question have actually been reviewed – in the context of monitoring compliance with the specific requirements of Art. 8 et seqq. of the draft AI Regulation.

The focus thus shifts to the concept of the relevant risk. In this respect, an orientation can be made on the basis of the other value judgement of the draft regulation. Thus, a relevant risk can always be assumed if a risk is added or an existing risk is increased in a way that in itself justifies the classification as a high-risk AI system and would thus entail the obligation

to conduct a conformity assessment. Insofar as the occurrence of a new risk is in question, this prerequisite is often – although not always – fulfilled in the case of a change of purpose. One should think of constellations in which the AI system is now used in another of the areas specified in Annex III. If the threshold advocated here is not exceeded, it can be left to the provider's discretion whether to carry out a new conformity assessment. If he does not do so, a special liability risk should be linked to this.

The following example can be used for the variant of an increase in risk which makes a renewed conformity assessment necessary: According to Annex III No. 3 a) of the draft AI Regulation, AI systems that are intended to be used for decisions on the access or allocation of natural persons to education and training institutions are high-risk AI systems. It is conceivable that such a system was used for access to higher education and had previously undergone a conformity assessment. As a result of an update, the system now takes into account not only the categories of grades, waiting periods after leaving school and social commitment, but also the parents' income. The latter could be used as a category to favour graduates from financially weak families in the allocation of university places. However, this criterion raises considerable ethical and legal questions. It entails increased risks with regard to the fair distribution of university places. The risk generated by the system so far is modified by the new category in a way that in itself makes a new conformity assessment – at least with regard to this additional aspect – appear appropriate. The control consideration here is the isolated view of a system that would carry out study allocations solely on the basis of this criterion. This would be classified as a high-risk AI system in the sense of Annex III No. 3 a) of the draft AI Regulation. The increase in risk associated with the extended selection category therefore justifies the necessity of a further conformity assessment in principle.

This risk-based consideration – borrowed from the basic concept of the draft regulation – with regard to the obligation to carry out a new conformity assessment can have far-reaching consequences for the person responsible. Conformity assessment procedures impose considerable requirements – they mean a great deal of effort for the responsible party, not least in economic terms. In this respect, the reasons that can usually lead to significant changes or a significantly increased risk in an AI system must be taken into account. It is possible that the architecture of an AI system – for example its neural network – is subsequently modified. This can have far-reaching consequences that also affect the risk assessment. However, this is probably the less significant case in practice. In contrast, the focus shifts to a "continued learning" of the system based on machine learning with a previously unknown set of data after the conformity assessment has been carried out. This data may have been obtained in different ways – for example, through an independent data collection of the AI system in operation or through external acquisition. In any case, this process involves the risk of substantial modification to the AI system – new rules may be devised and implemented that modify the way the programme works in ways not previously imag-

ined by humans. This also carries risks that may require new compliance assessment. Depending on how often an AI system relies on this process of further learning in order to maintain its high technical standard or to technically adapt to new, significant developments, this can result in a short sequence of conformity assessment obligations for the responsible party. The current draft does not adequately take this into account. The reason for this is the fact that the draft AI Regulation follows an "all or nothing" principle with regard to conformity assessment procedures. If a conformity assessment has to be carried out again, this includes all processes of the assessment procedure – no specific shortening of the procedures is considered. It is quite conceivable that systems that continue to learn again and again, but use very similar data sets of high quality in each case, do not need a complete conformity assessment in order to provide sufficient assurance of their ongoing security. It therefore seems advisable to also adapt the obligation for a repeated conformity assessment to the requirements of proportionality. In specific circumstances, a shortened conformity assessment may be necessary to safeguard consumer rights. Last but not least, this has considerable practical advantages both for the interests of consumers worthy of protection and for the companies. As long as repeated conformity checks are carried out according to the "all or nothing" principle, it seems conceivable on the one hand that a "substantial modification" as a trigger for a renewed assessment will only be assumed under very narrow conditions, so as not to cause an unjustified burden on the person responsible for the AI system. This bears the risk that the necessary conformity assessments are not carried out because the threshold for this is set too high. However, this proves to be disadvantageous for the rights of consumers. On the other hand, the current design could also have the opposite effect, i.e. that those responsible, in view of the liability risks described, would assume a "substantial modification" too often, i.e. partly wrongly, and carry out renewed assessments as a precautionary measure. This scenario would be associated with considerable losses, especially of a financial nature, for the companies, which in turn could have a detrimental effect on their innovation potential. In view of the dual objective of the draft, on the one hand to preserve European fundamental rights and values, and on the other hand to make Europe the centre for trustworthy AI, none of the scenarios presented is convincing. It is therefore preferable to carry out more frequent conformity assessment procedures, which could then be repeated under limited conditions. It would be welcome if a corresponding regulatory system were to be drafted in the further legislative process.

II. Requirement of complementation through (continuous) monitoring procedures?

The draft Regulation imposes a wide range of obligations on those responsible for high-risk AI systems, which relate to monitoring their use even outside conformity assessments. These include, in particular, various documentary obligations, record keeping obligations, requirements for human supervision, etc. In addition, the focus is on the obligation to set up a risk management system (Art. 9 of the draft AI Regulation) and a quality management system (Art. 17 of the draft AI Regulation). In this context, the question arises as to whether, in addition to the rules on the conformity assessment procedure, there is a need

for additional regulation imposing even more stringent obligations in relation to the monitoring procedure already set out in the draft. The term "monitoring" generally describes the monitoring of specific processes by a variety of methods, such as documentation, measurement, observation, etc. Various technical aids and other observation systems may be used.[138] In the context of AI systems, monitoring procedures are in principle an important tool for safeguarding consumer rights. As explained above, a new conformity assessment procedure under the current draft will only be considered if there has been a substantial modification to the AI system or, according to our understanding, if there is a relevant new risk associated with the application. In the meantime, however, there is a need for protective mechanisms which are less intrusive for those responsible for the system, but which at the same time allow for continuous monitoring of the potential risks involved. This is the only way to take account of the process and dynamics of this technology, given the current state of science, which does not (yet) allow reliable real-time monitoring of AI systems.

Monitoring procedures are carried out largely without the involvement of third parties. In terms of guaranteeing consumer rights, this proves to be less protection than external control. However, it should also be taken into account that the breach of obligations to monitor the high-risk AI system entails a substantial liability risk. This is particularly important in view of the fact that the establishment and ongoing implementation of a monitoring system involves a considerable amount of expenditure, including financial expenditure. Essentially, therefore, the combination of continuous monitoring and conformity control obligations provided for in the draft Regulation appears to be an appropriate means of safeguarding consumer rights. As a result, consumer-friendly continuous monitoring procedures by external parties are likely to prove disproportionate. Such long-term and therefore particularly intensive interventions in the legal sphere of the providers, for example with regard to their legitimate interests of confidentiality, can hardly be legitimised as long as the providers themselves are obliged to implement effective internal monitoring procedures.

III. Specifics of the liability of conformity assessment bodies with regard to AI systems

The draft AI Regulation itself does not explicitly provide for the liability of conformity assessment bodies with regard to audited AI systems. This applies both with regard to possible violations of the law by consumers as well as by providers and other persons responsible for the product. In that respect, reference should be made to the foregoing explanations on general liability rules under national law. With regard to AI systems and the conformity assessment bodies operating in this respect, it would be important to assess whether they act in a public capacity. Only on this condition a claim to official liability can be considered. There is also the question of how the relationship between the inspection body and the person responsible for the AI system is regulated internally – for example, whether a general exclusion of liability can be agreed upon. This requires in-depth studies, which will,

[138] Onlinemarketing.de-Lexikon, term "monitoring" (See the website at https://onlinemarketing.de/lexikon/definition-monitoring, last accessed on: 30.08.2022).

however, be linked to the specific design of future conformity assessment bodies and are therefore not the subject of this report.

D. Synthesis: Guidelines for a consumer-friendly design of the conformity assessment procedures in the draft AI Regulation

I. Preliminary considerations

The European Commission's draft AI Regulation has now been extensively commented on by the Committee on the Internal Market and Consumer Protection and the Committee on Civil Liberties, Justice and Home Affairs in the European Parliament. Their suggestions will be taken into account in our own subsequent proposals if they are taken up by us. This is indicated by the enumeration referring to the document "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 – COM/2021/206 final". In addition, some aspects relevant from a consumer point of view, which are touched upon in the context of other amendments, which are not included here, will be examined in greater detail.

1. Limiting harmonised standards to technical specifications and procedures (AM 2126)

Amendment 2126 proposes to limit harmonised standards to technical specifications and procedures. Work organisation and ethical considerations should not be applicable. Indeed, it is true that the definition of ethical or even legal guidelines is not a matter for standardisation.[139] However, we have already pointed out in the course of the elaboration so far that, not least for reasons of consumer protection and the preservation of democratic legislative processes, it seems necessary for the draft Regulation to be more specific than it is currently the case with regard to the requirements for various AI systems. Ethical and legal values must be considered in such a concretization. If this desirable step will not be taken, however, it would be problematic to remove ethical considerations entirely from the standardisation process. This would run the risk of completely neglecting relevant social concerns.

2. Proposals to limit the Commission's powers with regard to common specifications

A number of amendments propose limiting the Commission's powers to adopt common specifications, varying in detail. For example, Amendment 2129 proposes the complete deletion of Article 41 of the draft AI Regulation. A modification of the draft regulation to take account of such proposals should be warned against. Regarding safeguarding consumer rights, it is beneficial if the competence to determine the content of the requirements to be imposed on AI systems is not largely or completely entrusted to standardisation institutions. Indeed, their expertise is of enormous importance for the development of such requirements. But as has been shown, the involvement of the European legislator in the sense of

[139] See already the comments on this subject under B. IV. 3.

further specification in the text of the regulation itself takes much better account of consumer rights in particular. If this is not done, it seems necessary not to allocate competences unilaterally, but to give the Commission a relevant role in terms of specifications. Especially when the Commission (which should be self-evident) draws on the expertise of experts – particularly from standardisation institutions – this proves to be a greater guarantee of consumer rights than a procedure that is unilaterally dominated by a few non-legislative actors.[140]

Against this backdrop, AM 2130 must also be viewed critically. The aim is to depend the specification competence of the EU Commission, among other things, on the inaction of European standardisation organisations. AM 2131 then calls for the automatic replacement of the common specifications adopted by the EU Commission in the event of the intervention of a European standardisation organisation. Such a one-sided power relationship between standardisation organisations and the European Commission cannot be justified even in view of the high level of technical expertise of the standardisation institutes.

AM 2134 also proves to be problematic for similar reasons and is therefore not acceptable in our view. It is intended to depend the specification competence of the EU Commission on the existence of international standards. The average high level of European standards for the protection of consumer rights is rarely matched by international standards. In this context, it is a retrograde step to reduce the Commission's competence in favour of standards which may not correspond to the level of European ideas on consumer protection.

In addition, AM 2137 cannot be approved. The implementation of this proposal would mean that the Commission would not be able to intervene by means of technical specifications if the standards adopted as such are sufficient for the subject matter in question, even if the Commission wishes to act in respect of entirely new and different risks. This is not very effective, not least for reasons of consumer protection.

3. Protection of children by common specifications (AM 2132)

AM 2132 requires the Commission to adopt common specifications defining how risk management systems can address the specific concerns that may arise when AI systems interact with children. This request is worthy of approval. However, we consider it advisable to include this general consideration in the draft Regulation already in the context of the requirements for risk management systems under Art. 9 of the draft AI Regulation. We are strongly in favour of this.

[140] For this reason, we also call for the involvement of standardisation institutions in the process of developing common technical specifications, see our proposal for the introduction of a § 41 (1a) of the draft AI Regulation.

II. Amendments

Text proposed by the Commission	*Amendment*
Article 40 High-risk AI systems which are in conformity with harmonised standards or parts thereof the references of which have been published in the Official Journal of the European Union shall be presumed to be in conformity with the requirements set out in Chapter 2 of this Title, to the extent those standards cover those requirements.	**Maintain**
	AM 2125: Article 40 – paragraph 1 a (new) **The Commission shall issue standardisation requests covering all essential requirements of this Regulation in accordance with Article 10 of Regulation 1025/2012 no later than 6 months after the date of entry into force of this Regulation.**
Article 41 paragraph 1 Where harmonised standards referred to in Article 40 do not exist or where the Commission considers that the relevant harmonised standards are insufficient or that there is a need to address specific safety or fundamental right concerns, the Commission may, by means of implementing acts, adopt common specifications in respect of the requirements set out in Chapter 2 of this Title. Those implementing acts shall be adopted in accordance with the examination procedure referred to in Article 74(2).	AM 2136: Where harmonised standards referred to in Article 40 do not exist or where the Commission considers that the relevant harmonised standards are insufficient or that there is a need to address specific safety, **accessibility** or fundamental right concerns, the Commission may, by means of implementing acts, adopt common specifications in respect of the requirements set out in Chapter 2 of this Title. Those im-

	plementing acts shall be adopted in accordance with the examination procedure referred to in Article 74(2).
Article 41 paragraph 1a (new)	**When deciding to draft and adopt common specifications, the Commission shall consult the Board, the European standardisation organisations as well as the relevant stakeholders including consumer protection agencies. These organisations and stakeholders shall be regularly consulted while the Commission is in the process of drafting the common specifications.**
Article 41 paragraph 2 The Commission, when preparing the common specifications referred to in paragraph 1, shall gather the views of relevant bodies or expert groups established under relevant sectorial Union law.	 The Commission, when preparing the common specifications referred to in paragraph 1, shall **consult** relevant bodies, expert groups **and other relevant stakeholders** established under relevant sectorial Union law **including consumer protection agencies**.
Article 41 paragraph 3 High-risk AI systems which are in conformity with the common specifications referred to in paragraph 1 shall be presumed to be in conformity with the requirements set out in Chapter 2 of this Title, to the extent those common specifications cover those requirements.	**Maintain**
Article 41 paragraph 4 Where providers do not comply with the common specifications referred to in paragraph 1, they shall duly justify that they have adopted technical solutions that are at least equivalent thereto.	AM 2148 Where providers do not comply with the common specifications referred to in paragraph 1, they shall duly justify that they have adopted technical solutions that **meet the requirements referred to in Title III, Chapter 2 to a level** at least equivalent thereto.
Article 42 paragraph 1	**Maintain**

Taking into account their intended purpose, high-risk AI systems that have been trained and tested on data concerning the specific geographical, behavioural and functional setting within which they are intended to be used shall be presumed to be in compliance with the requirement set out in Article 10(4).	
Article 42 paragraph 2 High-risk AI systems that have been certified or for which a statement of conformity has been issued under a cybersecurity scheme pursuant to Regulation (EU) 2019/881 of the European Parliament and of the Council63 and the references of which have been published in the Official Journal of the European Union shall be presumed to be in compliance with the cybersecurity requirements set out in Article 15 of this Regulation in so far as the cybersecurity certificate or statement of conformity or parts thereof cover those requirements.	**Maintain**
Article 43 paragraph 1 For high-risk AI systems listed in **point 1 of** Annex III, where, in demonstrating the compliance of a high-risk AI system with the requirements set out in Chapter 2 of this Title, the provider has applied harmonised standards referred to in Article 40, or, where applicable, common specifications referred to in Article 41, the provider shall follow one of the following procedures: (a) the conformity assessment procedure based on internal control referred to in Annex VI;	For high-risk AI systems listed in Annex III, where, in demonstrating the compliance of a high-risk AI system with the requirements set out in Chapter 2 of this Title, the provider has applied harmonised standards referred to in Article 40, or, where applicable, common specifications referred to in Article 41, the provider shall follow one of the following procedures: (a) the conformity assessment procedure based on internal control referred to in Annex VI;

	AM 2171:
(b) the conformity assessment procedure based on assessment of the quality management system and **assessment of the** technical documentation, with the involvement of a notified body, referred to in Annex VII.	(b) the conformity assessment procedure based on assessment of the quality management system and technical documentation, with the involvement of a notified body, referred to in Annex VII.
Where, in demonstrating the compliance of a high-risk AI system with the requirements set out in Chapter 2 of this Title, the provider has not applied or has applied only in part harmonised standards referred to in Article 40, or where such harmonised standards do not exist and common specifications referred to in Article 41 are not available, the provider shall follow the conformity assessment procedure set out in Annex VII.	
	AM 2177:
For the purpose of the conformity assessment procedure referred to in Annex VII, the provider may choose any of the notified bodies. However, when the system is intended to be put into service by law enforcement, immigration or asylum authorities as well as EU institutions, bodies or agencies, the market surveillance authority referred to in Article 63(5) or (6), as applicable, shall act as a notified body.	For the purpose of **carrying out** the conformity assessment procedure referred to in Annex VII, the provider may choose any of the notified bodies. However, when the system is intended to be put into service by law enforcement, immigration or asylum authorities as well as EU institutions, bodies or agencies, the market surveillance authority referred to in Article 63(5) or (6), as applicable, shall act as a notified body.
Article 43 paragraph 2	**Delete**
For high-risk AI systems referred to in points 2 to 8 of Annex III, providers shall follow the conformity assessment procedure based on internal control as referred to in Annex VI, which does not provide for the involvement of a notified body.	

For high-risk AI systems referred to in point 5(b) of Annex III, placed on the market or put into service by credit institutions regulated by Directive 2013/36/EU, the conformity assessment shall be carried out as part of the procedure referred to in Articles 97 to101 of that Directive.	
Article 43 paragraph 3 For high-risk AI systems, to which legal acts listed in Annex II, section A, apply, the provider shall follow the relevant conformity assessment as required under those legal acts. The requirements set out in Chapter 2 of this Title shall apply to those high-risk AI systems and shall be part of that assessment. Points 4.3., 4.4., 4.5. and the fifth paragraph of point 4.6 of Annex VII shall also apply.	**Article 43 paragraph 2 (new)**
Article 43 paragraph 4 High-risk AI systems shall undergo a new conformity assessment procedure whenever they are substantially modified, regardless of whether the modified system is intended to be further distributed or continues to be used by the current user. For high-risk AI systems that continue to learn after being placed on the market or put into service, changes to the high-risk AI system and its performance that have been pre-determined by the provider at the moment of the initial conformity assessment and are part of the information contained in the technical documentation referred to in point 2(f) of Annex IV, shall not constitute a substantial modification.	**Article 43 paragraph 3 (new)** **High-risk AI systems that have already been the subject of a conformity assessment shall be subject to a new conformity assessment procedure,** **(a) where a modification creates a new risk which in itself justifies categorization as a high-risk AI system. This is the case, in particular, if the modification enables use in a sector of Annex III that deviates from the previous categorization of the high-risk AI system;** **(b) where, as a result of a modification, the risk associated with the high-risk AI system, which may be covered by the categorization in an Annex III sector, increases to an extent that in itself justifies a conformity assessment.** **A new conformity assessment shall take due account of the results of the**

	previous assessment, allowing in particular for a limitation of the scope of the assessment to newly added risks justifying the classification as a high-risk AI system or to newly added risk-relevant factors.
Article 43 paragraph 4 a (new)	**Article 43 paragraph 3 a (new)** AM 2197: **The specific interests and needs of the small-scale providers shall be taken into account when setting the fees for third-party conformity assessment under this Article, reducing those fees proportionately to their size and market size.**
Article 43 paragraph 4 b (new)	**Article 43 paragraph 3 b (new)** AM 2198: **Any provider may voluntarily apply for a third-party conformity assssment regardless oft he risk level of their AI system.**
Article 43 paragraph 5 The Commission is empowered to adopt delegated acts in accordance with Article 73 for the purpose of updating Annexes VI and Annex VII in order to introduce elements of the conformity assessment procedures that become necessary in light of technical progress.	**Article 43 paragraph 4 (new)**
Article 43 paragraph 6: The Commission is empowered to adopt delegated acts to amend paragraphs 1 and 2 in order to subject high-risk AI systems referred to in points 2 to 8 of Annex III to	**Delete**

the conformity assessment procedure referred to in Annex VII or parts thereof. The Commission shall adopt such delegated acts taking into account the effectiveness of the conformity assessment procedure based on internal control referred to in Annex VI in preventing or minimizing the risks to health and safety and protection of fundamental rights posed by such systems as well as the availability of adequate capacities and resources among notified bodies.	