Human Oversight in AI Use

Professor Dr Michael Pils, University of Konstanz[*]

From 2 August 2026, the European AI Act[1] will apply to high-risk AI systems[2]. A key provision for ensuring that AI protects fundamental rights is Art. 14 AI Act. It states that high-risk AI systems[3] must be designed and developed in such a way that they can be effectively supervised by humans, at the latest from the time they are placed on the market and in particular during their use (including with appropriate tools). Although human oversight is indisputably a fundamental principle of the AI Act, there is relatively little literature dealing with the implementation of human oversight.[5]

This Essay addresses the question of how human oversight is structured in the AI Act. Implementation guidelines will also be provided. First, the dichotomy of human oversight is briefly presented (see I. below). Since the wording raises numerous questions (see II. below), the author recommends an interpretation of the principle of human oversight based on its protective purpose (see III. below). The elements of human oversight mentioned in the text of the AI Act are then analysed (see IV. below), with a particular focus on risk assessment and the interaction between providers and deployers. The Essay concludes with the question of the delegation of supervisory duties, which is also not trivial in terms of labour law (see V. below).

I. Human oversight - safety net or obstacle to innovation?

Human oversight is intended as a last[5] safety net to avert or substantially reduce risks from the usage of high-risk AI systems. The regulator has deliberately placed the emphasis on humans. The choice of a human-centred approach is primarily in line with the programmatic objective of the AI Act to ensure that artificial intelligence is particularly trustworthy. Human oversight ensures control and takes into account the deeply rooted human need for control and security.[6]

[*] Michael Pils is a partner at the law firm TaylorWessing in Düsseldorf and a professor of employment and labour law at the University of Excellence in Konstanz (collective employment law and European employment and social security law, as well as, international law).

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144, and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Regulation on Artificial Intelligence) (OJ L, 2024/1689).

[2] Art. 113(2) AI Act in conjunction with Art. 113(3)(a) AI Act.

[3] For the definition, see Art. 6 et seq. AI Act in conjunction with Annexes I and III.

[4] Apart from the commentary literature, the following are particularly noteworthy: Hetmank/Meinel, KIR 2024, 127 et seq.; Bomhard, DSRITB 2024, 421 et seq.; Schemel, CB 2005, 353, 354.

[5] For an introduction to the regulatory framework, see: Chibanguza/Steege, NJW 2024, 1769; Ashkar/Schröder BB 2024, 771; Krönke, NVwZ 2024, 529.

[6] See Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st ed. 2025, Art. 14 KI-VO, para. 4 with further references.

However, the human-centred approach of the regulator is not without controversy. Human oversight can - depending on the complexity of the system - be harmful, or at least mostly pointless, and even convey a false sense of security.[7] Furthermore, the fear of machines seems anachronistic, as concerns about machine domination are not new. In art and literature, the 'fear of machines' has been a theme and cultivated since at least 1872[8]. It is not uncommon for the ultimate superiority of humans to be called into question. The human-centred approach was already criticised as anachronistic in Art. 22 GDPR.[9] Human oversight is - as the comic character Homer Simpson perhaps symbolises - anything but easy to achieve; this is merely discussed under the key word 'illusion of control'[10]. This illusion of control has two effects: firstly, it can convey a false sense of security that overlooks or underestimates the potential errors/risks of AI. Secondly, the illusion of control can reduce the willingness to invest in actual improvements to supervision or to take reasonable risks.[12] The legal obligation to implement human oversight 'at all costs' or - on the contrary - the unclear concept of human oversight in the AI Act can prevent investment in high-risk AI systems (also with respect to the potential breach of the AI Act and the high fines imposed by the AI Act - see IV.7 below), prevent potential efficiency gains from being exploited, or be a reason not to use the technology because of the associated liability and financial penalty risks.[12]

To put it bluntly, the weakness of high-risk AI systems could be the very 'human' factor that is supposed to supervise them. Scientists are increasingly recognising that humans are generally unable to concentrate on monitoring automated processes for longer than 30 minutes.[13] On the other hand, human supervision quickly becomes ineffective: Studies

[7] Colonna Faculty of Law, Stockholm University Research Paper 2024, 443, 444.

[8] Cs. Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st ed. 2025, Art. 14 AI Act, para. 3; the idea of a 'rule of machines' dates back to the early days of industrialisation. As early as the early 19th century, workers in England and Central Europe reacted against the mechanisation of production – for example, the so-called Luddites in England or the protests against Johann Heinrich Schüle's textile factory in Augsburg, which even led to the manufacturer being convicted by the city of Augsburg (see Grünsteudel/Hägele/Frankenberger, Stadtlexikon Augsburg, s.v. Schüle m.w.N.). Nineteenth-century literature continued to address the domination of machines over humans, for example in Samuel Butler's satirical essay 'Erewhon' (1872) . In the play 'R.U.R.' (1920), Karel Capek invented the term "robot" and told the story of artificial workers who rise up against humanity. Fritz Lang's film 'Metropolis' (1927) is widely known for its depiction of a high-tech city of the future. The bleak visions of the future in 2001: A Space Odyssey (1968), The Terminator (1984) and The Matrix (1999) are also considered influential, sharing a vision of complete technological domination by intelligent machines. Literature also deals with the fundamental dilemma of machines superior to humans: in I, Robot (1950), Isaac Asimov's robot laws offer an alternative to the machine apocalypse and attempt to conceive of a rationally regulated coexistence between humans and machines. In contrast, William Gibson's Neuromancer (1984) depicts a world in which artificial intelligences develop their own power structures in cyberspace; in the 1950s, mathematician Alan Turing warned of the danger 'machine'. Researchers seem to agree that human intelligence is possible even without biology, purely on a silicon basis; Walter, ifo Schnelldienst 8/2023, 19; for an introduction to the history of ideas about AI in art and culture, see Catani/Pfeiffer, Handbuch der Künstlichen Intelligenz, 2023.

[9] Among many others, Gola/Heckmann/Schulz, Art. 22 GDPR, para. 2.

[10] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st ed. 2025, Art. 14 AI Act, para. 5 ss., esp. 8.

[11] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st ed. 2025, Art. 14 AI Act, para. 8; Glaser, Risk in Management, p. 339.

[12] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st edd. 2025, Art. 14 AI Act, Rn. 5 ss., esp. 9.

[13] Baxter/Rooksby/Yang/Khajeh-Hosseini, Proceedings of the 30th European Conference on Cognitive Ergonomics, August 2012, 65, 66, available at: doi.org/10.1145/2448136.2448149.

[14] Brauner/Philipsen/Calero Valdez/Ziefle Behaviour and Information Technology 38 (2019), available at: doi.org/10.1080/0144929X.2019.1581258.

show that AI-supported suggestions are highly likely to be accepted, whether out of convenience or because the complexity and lack of transparency of AI makes concrete verification too difficult,[14] or in extreme cases even rejected entirely[15].

The implementation of human supervision raises further questions, in particular regarding the requirements to be placed on the supervisor. The qualification concept proposed by the regulator, which goes beyond AI competence (Art. 4 AI Act) and is intended to enable the person exercising human oversight, will be difficult to be obtained in many areas of application - just think of highly complex diagnostic or surgical instruments in medical technology; depending on the complexity of high-risk AI systems, it is difficult to ensure that the supervisor can actually adequately assess the risks. Sometimes, paradoxical results are even conceivable, whereby human supervision actually leads to safety risks, for example when intervening in a high-precision AI-supported robot used in eye surgery[16]. It therefore makes sense for a high-risk AI system to exclude human intervention for good reasons or, at least, to optimise results based on data obtained without human supervision.[17] After all, the complexity of high-risk AI systems poses a particular challenge for human supervision. How, for instance, should the supervisor decide whether the high-risk AI system is causing harmful or desirable anomalies and when is a desirable anomaly a malfunction?[19] Even with the best qualifications, there are limits to knowledge.

II. Presentation of the normative system of Art. 14 AI Act and open questions

Despite the dilemma presented, the AI Act requires effective human oversight of high-risk AI systems[20]. Art. 14(1) AI Act contains the basic principle of the necessity of human oversight and defines the basic requirements in the form of a general clause. As can be seen from a comparison with Art. 14(2) AI Act, the supervisor should avert risks to the triad of protected interests - health, safety or fundamental human rights - arising from the use of the high-risk AI system. Human oversight is thus designed as an external control body that the provider must implement and review as part of the risk management system (Art. 9 AI Act) and the quality management system (Art. 17 AI Act).

As already indicated, by establishing the supervisory person in the AI Act, the legislator wanted to strengthen the confidence of EU citizens in the use of high-risk AI systems.[20] The standard focuses on the technical and organisational measures that are intended to

[15] Parasuraman/Manzey Human Factors 52 (2010), 381.

[16] Aptly put by Gassner, MPR 2023, 5, 10; Cobbaert RF Quarterly 2021/2, 4 (13), available at www.raps.org/RAPS/media/news-images/Feature%20PDF%20Files/21-6_RFQ-2_Cobbaert_4-26.pds.

[17] Gassner, MPR 2023, 5, 20, refers to this as a ban on innovation.

[18] Linardatos, GPR 2022, 58. 66.

[19] Recital 73 AI Act, para. 1, first sentence.

[20] Recital 64 sent. 1 AI Act; critical: Laux, Institutionalised distrust and human oversight of artificial intelligence: towards a democratic design of AI governance under the European Union AI Act, AI & Society (2023), 1 s.

monitor the high-risk AI system in such a way that human supervision is also possible in a meaningful and risk-minimising manner[21]. Art. 14(1) AI Act imposes the obligation of providers to develop high-risk AI systems in such a way that they can be effectively supervised.[22] Art. 14(2) AI Act focuses on the programmatic goal of human oversight as a basic principle, according to which human oversight ultimately prevents risks to health, safety or fundamental rights posed by high-risk AI systems. Human oversight is not a static system, but a flexible one that also depends on the particular high-risk AI system.[23] The individual requirements for human oversight are mentioned in Art. 14 (3) to (5) AI Act, with Art. 14 (3) AI Act containing the technical organisational requirements for establishing human oversight, while Art. 14(4) AI Act, which is closely related to Art. 13 AI Act,[24] describes the functionality requirements for ensuring human oversight and is intended in particular to prevent so-called automation bias[25]. Human oversight must be technically implemented in particular in the 'stop-bottom' function.[26] For biometric third-party identification, Art. 14(5) AI Act contains a special provision (not considered in detail here) which enshrines the dual control principle.[27]

The complexity of the provision in Art. 14 AI Act, which is apparent at first glance, raises numerous questions of interpretation. In the literature, the provision is therefore often criticised as unclear. Art. 14 AI Act does not establish a clear and unambiguous regulation in every respect.[28] In particular, the following questions could be asked:

- Must human oversight always be established? Or can or must human oversight even be omitted if the technical design of the high-risk AI system makes the system appear safe or - taken to the extreme - if the establishment of human oversight makes the system unstable or prone to errors, for example in the case of medical devices?
- What criteria should be applied to determine appropriateness in the meaning of the AI Act? Should appropriateness be based solely on the supervisory measures (as stated in the wording of Art. 14 AI Act) or must the precautions referred to in Art. 14(3) AI Act also be appropriate?[29]

[21] Regarding the measures, see Sterz/Baum/Biewer/Hermanns/Lauber-Rönsberg/Meinel/Langer, FAcct '24: Proceedings of the 2024 ACM Conference on Fairness, p. 2500 ss.

[22] BeckOK KI-Recht/Buchner, 4th ed. 2025, Art. 14 Rn. 2.

[23] Aptly BeckOK KI-Recht/Buchner, 4th ed. 2025, Art. 14 KI-VO Rn. 4.

[24] Hilgendorf/Rothläsig, KI-VO/Kumkar, § 6 No. 25.

[25] Examples include: Wickens et al., The Journal of the Human Factors and Ergonomics Society, 2015, 728.

[26] Spindler, CR 2021, 361, 367: panic bottom.

[27] Cs. Art. 3 No. 41, Art. 14(5) subparagraph 1 AI Act; Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, margin note 3, 134 and passim; Martini, NVwZ-Extra 1-2/2022, 1, 5; BeckOK, AI Law/Buchner, Art. 14 AI Act, para. 6. Hüger, Artificial Intelligence and Discrimination, p. 103 ss.; Zußner, in: Chan/Ennusschat, Artificial Intelligence and Public Commercial Law, p. 95 ss.; Schindler, Biometric Video Surveillance, p. 51 ss.

[28] Hetmank/Meinel, KIR 2024, 127, 133.

[29] Hetmank/Meinel, KIR 2024, 127, 128 s.

- Does Art. 14(3) and (4) AI Act provide for a tiered system of responsibilities for the provider and - separately - the deployer?
- To what extent can the provider (partially) delegate its obligations to the deployer, or do both ultimately remain responsible?[30] What obligations does the provider have with regard to ensuring that measures are also implemented by the deployer ?[31] Must the provider intervene if the high-risk AI system continues to develop after being placed on the market without reaching the threshold for change under Art. 43(4) AI Act?
- Must every technically conceivable option for establishing appropriate human supervision be implemented, or can financial or time-related organisational circumstances, for example, lead to one of the safeguard measures in Art. 14 AI Act being omitted? Is the additional financial burden an element of proportionality, as some of the literature[32] assumes?
- Should suitability be determined subjectively, taking into account the deployer's context, or abstractly and objectively?
- How can one assess whether human oversight is effective within the meaning of Art. 14(4) AI Act?
- When is the supervisor enabled in an appropriate and proportionate manner to understand the performance/limitations of the high-risk AI system, be aware of and counteract automation bias, correctly interpret the results of the high-risk AI system, or consciously decide against using the high-risk AI system, and intervene in the high-risk AI system (at any time and ultimately also justifiably (i.e. for an objective reason)?
- Can the high-risk AI system be handed over to a private individual who is not considered a deployer according to the definition?
- With regard to software ergonomics, how should be the warning system for human oversight be designed that it is not too confusing or causes an overload of acoustic, visual and tactile signals?[33]

III. Protective purpose-related interpretation principle and systematic interpretation

The open questions of interpretation must be answered against the background of the protective purpose of Art. 14 AI Act.

The protective purpose is the primacy of human action and human supervision, as the historical development and recitals make clear.[34] Human supervision is seen as a

[30] Controversial, cs. Hetmank/Meinel, KIR 2024, 127, 129.

[31] Hetmank/Meinel, KIR 2024, 127, 130 s.

[32] Hetmank/Meinel, KIR 2024, 127, 130.

[33] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st edition 2025, Art. 14 KI-VO, margin note 14.

guarantee eliminate or reduce the weaknesses of high-risk AI systems, namely the lack of 'common sense', empathy and self-reflection, and phronesis as moral wisdom and the ability to take all relevant factors into account in an socially appropriate manner in everyday life.[35] At the same time, according to the regulator's idea, human oversight ensures that the consequences of AI-based decisions are assessed and evaluated on the basis of the polyphony of human perception, which is considered superior.[36] Ultimately, human oversight is intended to prevent deception in the results or decision-making of the high-risk AI system, if necessary.[37] The interaction between humans and machines should - ideally, of course[38] - exploit the performance potential of the high-risk AI system and prevent risks arising from humans[39], for example due to arbitrariness, emotion, prejudice, etc.[40]

The scope of the triad of protected interests[41] (health, safety, fundamental rights) poses a challenge for the practical implementation of the risk analysis approach underlying Art. 14 AI Act. The triad of protected interests is primarily linked to the fundamental rights enshrined in the European Charter of Fundamental Rights.[42] Health[43], for example, refers to physical and mental integrity (Art. 3(1) of the European Charter of Fundamental Rights). The concept of security as the integrity of the legal order also has a strong fundamental rights reference in its definition. Strong fundamental rights protection can be found in various places in the AI Act, most prominently in Recital 1 AI Act, which explicitly refers to the protection of democracy, the rule of law and the environment, or more strongly in Recital 48 AI Act. The reference to fundamental rights mentioned last in the triad should be regarded as a general clause. This has direct consequences for interpretation: Although the wording of Art. 14(3) AI Act states that the balancing criteria are exhaustive (which has been criticised[44]), it should be clear from the interpretation under Art. 14(2) AI Act that this cannot be meant to be an exhaustive list due to the difficulty of classifying high-risk AI systems.

[34] With reference to the historical derivation, BeckOK, AI Law/Buchner, Art. 14 AI Act, para. 8 ss. and passim, also on recitals 27 and 73.

[35] Martini Blackbox Algorithmus, p. 59.

[36] Martini Blackbox Algorithm, p. 59.

[37] Martini Blackbox Algorithm, p. 59.

[38] Critical of human-machine supplementation and the so-called MABA-MABA trap, EDPS TechDispatch on human oversight of automated decision-making, 2025, p. 11.

[39] Martini Blackbox Algorithm, p. 47 s.

[40] BeckOK AI Law/Buchner AI ActArt. 14 Rn. 31 f s.

[41] The protection triad and the reference to the EU's normative value system can already be found in Art. 1 (1) AI Act, but also in numerous other places, cs. Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 64.

[42] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 72 'Overlap'.

[43] On the conceivable health risks arising from and in connection with high-risk AI systems already mentioned in the recitals, see Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 65 ss.

[44] Hetmank/Meinel, KIR 4, 127, 129.

The strong focus on fundamental rights is dogmatically remarkable, as the AI Act makes fundamental rights applicable inter privatos, i.e. it swifts the public-law related ‘defense’ rights into a ‘general principle’ to be respected for the addressees of the AI Act. Principally, the European Fundamental Rights only apply inter privatos via general clauses or, in European law, only need to be taken into account with regard to the interpretation of fundamental freedoms (in this context, the European Court is said to be a motor of the promotion of the idea of fundamental rights). Although the AI Act only provides for a limited human rights impact assessment, it must be ensured, with regard to human oversight, that the aforementioned protected interests are taken into account in any case.

Systematically, European secondary law should be used and utilised for the interpretation of Art. 14 AI Act. The concept of human oversight can be found in various places in European secondary law, in particular[45] in the prohibition of fully automated decisions in Art. 22 GDPR[46]. This strict and penalized prohibition of Art. 22 GDPR is applicable in addition to Art. 14 AI Act, as its legal nature and scope of application are different[47], although the basic protection objective is comparable[48]. The three exceptions to the basic prohibition[49] mentioned in Art. 22(2) GDPR have in common is that the controller must take appropriate safeguards to preserve the rights and freedoms and legitimate interests of the data subject.[50] The obligation for the controller to intervene personally is considered a key safeguard.[51] Similar to Art. 14 AI Act, Art. 22 GDPR requires that decision-making processes in automated systems must be transparent and comprehensible and that, to this end, the controller and the data subject must be provided with information about the logic used in the decision-making process.[52]

Similarly, European occupational health and safety law also stipulates human supervision as a basic principle. The basic standard in Section 3 of the German Occupational Safety and Health Act (ArbSchG which bases on various EU directives) obliges employers to take

[45] Similar concepts are found in secondary legislation, e.g. in Art. 11 JI-RL, in Art. 22 Digital Service Act (Regulation (EU) 2022/2065), Art. 10 Platform Work Directive (EU 2024/2831) and Annex II No. 2.1 lit. b Aviation Safety Agency Regulation; Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 21 passim.

[46] On protection against being merely the object of an automated decision, see generally Paal/Pauly/Martini DS-GVO/BDSG Art. 22 para. 1.

[47] Aptly: Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 17 et seq., with analysis of the scope of application, according to which Art. 14 AI Actalready applies at the decision-preparation stage. Human supervision can exist independently of the automated decision; on the other hand, an automated decision does not exclude human supervision; for more details, see also Schemel, CB 2005, 353, 354 ss., Köhler, EuDIR 2025, 16, 21, whereby Schemel and Köhler specifically address the SCHUFA ruling of the ECJ (judgment of 7 December 2023, C-634/21, CB 2024, 71), which is not considered here.

[48] BeckOK, AI Law/Buchner, 4th ed. 2025, Art. 14 AI Act, para. 27.

[49] Conclusion/performance of a contract (Art. 22(2)(a) GDPR), legal basis (Art. 22(2)(b) GDPR) or explicit consent (Art. 22(2)(c) GDPR).

[50] Cf. Art. 22(2)(b) aE and (3) GDPR.

[51] Among many others, Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 15 s. ; Paal/Pauly/Martini GDPR/BDSG Art. 22 para. 35; Martini/Nink NVwZ-Extra 10/2017, 1, 3 ss.

[52] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, margin note 19 et seq.; Martini/Nink NVwZ-Extra 10/2017, 1, 5; Paal/Schulz ZfDR 2025, 89, 103 and 107.

[53] See Sections 3, 6 BetrSichVO for operating resources/materials.

the necessary occupational safety measures based on the risk assessment (Section 5 ArbSchG which also bases on various EU directives), which may include human supervision as a measure.[53] However, human supervision is not a mandatory requirement of H&S law, as according to the TOP principle, technical measures must be taken before organisational and the later before personnel measures. However, the occupational health and safety concept is based on the idea that the actors must work together to achieve the maximum level of protection and continuous improvement and adaptation. Since high-risk AI systems must ultimately also be understood against the background of occupational health and safety law, providers and deployers in particular must cooperate, as will be shown shortly. Since the high-risk AI system is, by definition, changeable, the risk assessment would have to be carried out on a rolling basis, which also requires the cooperation of all actors.

The fact that, under occupational health and safety law, work processes must be designed in such a way that humans control machines (and not vice versa) is also reflected in the basic idea of human supervision. Despite the undisputed advantages of high-risk AI systems, they should be regarded as nothing more than a tool within the framework of human-controlled work processes. If an employer were to use a high-risk AI system in such a way that it dominates human work (rather than merely supporting and facilitating it), this system would be difficult to justify under occupational health and safety law.

As an interim conclusion, it can be said that the systematics of Art. 14 AI Act and the interaction with accompanying European regulations, in particular data protection and occupational health and safety, make it clear that human oversight must be more than just a ‘rubber stamping of the machine's decision’ or a ‘formality’.[54] Human oversight cannot be designed as a one-off, static element, but must - especially in the interaction between provider and deployer - be dynamically geared towards mutual cooperation. Otherwise the protective purpose of Art. 14(2) AI Act would be rendered meaningless. If the AI Act is also understood as a regulation that takes occupational safety law into account (which is necessary in any case due to Art.14(2) AI Act, given its reference to fundamental rights), then the principles of occupational safety under European law must be taken into account in its interpretation and implementation. However, this may also mean that human oversight is not always necessary if, due to the technical-organisational measures, a risk to the protected interests referred to in Art. 14(2) AI Act can be ruled out against the background of a methodologically correct risk assessment. Similar to occupational health and safety law, human supervision of high-risk AI systems also has a third-party protection effect, as it is also a matter of protecting the legal positions and legitimate interests of third parties.

[54] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 47.

IV. Elements of human oversight

As shown, the regulator has regulated human oversight in a fragmentary manner. Numerous undefined legal terms make it necessary to interpret the provision. Those basic elements of human oversight are presented now.

1. Limited monitoring concept

The AI Act is based on a limited monitoring concept for high-risk AI systems with the establishment of human oversight. In Art. 14(1) AI Act, the regulator makes it clear that human oversight must be ensured as early as possible and throughout the duration of the application of the high-risk AI system (full 'life cycle').[55] Human oversight is therefore not only required from the time the system is placed on the market or becomes market-ready. Rather, human oversight must also be ensured during the experimental stage of development, especially since this is when potential risks are regularly discovered. The specific time at which the control is to take place is not regulated, nor is the minimum level of control[56], as this is to be defined in the context of risk assessment.

Furthermore, not every conceivable risk associated with the use of the high-risk AI system needs to be prevented or reduced, but only the obvious, foreseeable risks associated with its use. Furthermore, human supervision does not, in principle, require the natural person to be able to understand all the processes of the high-risk AI system.[57]. Foreseeability however must be defined also in the light of Art. 14 (2) AI Act. As soon a risk emerges - even by an other deployer or its customer - this risk is foreseeable. Thus, as mentioned and as to be shown later, a cooperation between deployer and provider is utmost important. On the other hand: Human oversight is therefore not an element that is necessary at all times, but merely a limited 'safety net' that compensates for the - assumed - weaknesses[58] of a high-risk AI system.

2. Human oversight as a design requirement for high-risk AI systems

Human oversight is designed as a fundamental design requirement for high-risk AI systems. Art. 14(1) AI Act defines human oversight as a design requirement for high-risk AI systems in order to classify AI as human-centric and trustworthy.[59]. Developers of high-

[55] Recital 73, sent. 1 AI Act; Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, paras. 51 and 55.

[56] Laux, Institutionalised distrust and human oversight of artificial intelligence: towards a democratic design of AI governance under the European Union AI Act, AI & Society 2023, 3.

[57] Ex Art. 14(4)(a) AI Act; Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 48.

[58] Regarding Legal philosophy, Martini Blackbox Algorithm, p. 59.

[59] See Recital 1, sent. 1 AI Act; on the establishment of the principle and the changes in the context of its development, as well as on the concepts of human in the loop, human on the loop and the ultimately adopted concept of human in command,

risk AI systems (at least the providers) must integrate human oversight into the programme code from the outset.[60] Human intervention and human interaction must therefore be possible immediately, or at least from the point of 'work readiness'.

Taking into account the findings of occupational health and safety software ergonomics, the perception of human oversight must be designed to be intuitively understandable and consistent ('user-friendly') for the user and effectively anchored via a human-machine interface.[61] The high-risk AI system must therefore be capable of supervision.[62] If Art. 14(1) AI Act is interpreted in conjunction with Art. 14(4) AI Act, there is much to suggest that the high-risk AI system must be equipped in such a way that the supervising person can supervise the high-risk AI system, including all its effects, e.g. on the economy, society and basic ethical principles, and must be able to decide both on its use and on the interruption of its use.[63] Although this primarily requires the supervisor to have the appropriate skills, the system must provide the information in such a way that supervision is ensured.[64]

As will be shown shortly, the specific design of human supervision depends on the AI risk assessment, so that other supervisory concepts with more or less intensive involvement of the 'human' factor may also be necessary for the specific high-risk AI system in the specific context of use.[65] Depending on the design and specific use of the high-risk AI system, a conceivable result of the risk assessment may also be that human oversight is dispensable.[66]

3. Risk assessment in the sense of an "AI risk assessment”

In order to understand the specific requirements for human oversight, providers, but also indirectly deployers, are required to analyse the risks and manage the risks in a manner appropriate to the legal interests that are foreseeably endangered by the high-risk AI system. Human oversight must be set up in such a way that the objective in Art. 14(2) AI Act can be achieved at any time. Conceptually, this objective means that the effectiveness of human oversight must be ensured on a dynamic, rolling basis. Risks to the aforementioned protected interests must be methodically recorded, gaps in knowledge or

Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 30 passim; BeckOK, AI Law/Buchner, 4th ed. 2025, Art. 14 AI Actpara. 15 ss.

[60] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st ed. 2025, Art. 14 AI Act, para. 5 ss., esp. 10

[61] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 53 s.; Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st ed. 2025, Art. 14 AI Act, para. 14.

[62] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st edition 2025, Art. 14 AI Act, para. 5 ss., esp. 11 ss.

[63] BeckOK AI Law/Buchner, Art. 14 AI Act, para. 36 et seq.

[64] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st ed. 2025, Art. 14 AI Act, para. 5 ss., esp. 11 ss.

[65] Similar view BeckOK KI-Recht/Buchner, Art. 14 KI-VO para. 34 and 35 as well as 37.

[66] Bomhard, DSRITB 2024, 421, 427.

any factors of uncertainty/unpredictability in the high-risk AI system must be identified, and appropriate countermeasures must be developed. The methodologically complete and systematically coherent recording of risks is an essential point in answering the question of whether human oversight is specifically effective, suitable and appropriate. The audit direction of the provider and deployer may, but does not have to, be identical, as the deployer may have different fundamental rights considerations in the specific context of use than the questions arising from the provider's perspective.

In addition to risk analysis, however, a risk management system is also required.[67] If, for example, the context of use changes or the high-risk AI system alters the working environment, the effectiveness of human supervision must be reassessed, just as it would be if new risks arose from use or the data set required for the AI changed.

The system of continuous, or at least event-driven, risk assessment is well known in European occupational safety law. The process of this effectiveness control can also be seen expressed in Art. 14(2) AI Act in order to define the foreseeability. Thus, this AI risk assessment should be implemented in practice via a standardised work process, similar to the risk assessment in occupational health and safety law. In order to avoid and reduce the threat of sanctions under Art. 99 AI Act, comprehensive cooperation between provider and deployer is also recommended here, but also with regard to any third parties (e.g. customers) that may be involved or affected by the high-risk AI system.

The risk assessment underlying Art. 14 AI Act does not mean that providers and deployers must prevent risks to the central protected assets *ex ante*. Rather, the hierarchical relationship between ‘prevent’ and ‘minimise’ recognises that although the primary duty is to prevent hazards to key protected assets, if the hazard cannot be prevented, it only needs to be minimised.[68] The correctness of this approach is supported by the parallel to occupational health and safety law, where, depending on the activity, H&S risks can often only be minimised. On the other hand, a prohibition of usage of the particular high-risk AI system may even be conceivable if there are changes in the state of the technical art, occupational medicine or occupational science. It may therefore be the case that a previously permissible use of the high-risk AI system becomes impermissible over time due to new findings, or that modifications become necessary. This must be taken into account in both the licence agreement and the description of the usage process in accordance with occupational health and safety law. In this context, the German Bundesanstalt für Arbeitsschutz und Arbeitssicherheit (Federal Institute of Work Protection and Safety) has established a scientist group for interdisciplinary analysis of the AI-related H&S risks. It is necessary to continuously monitor the results of the working

[67] With reference to Recital 65 AI Act, BeckOK AI Law/Buchner, Art. 14 AI Actpara. 39.

[68] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 63.

group since the results are the relevant standard of technical, occupational medicine or occupational science to be observed by H&S law.

Human supervision is a basic principle of high-risk AI systems. However, it is questionable whether it can always be guaranteed, for example if the AI risk assessment concludes that all requirements for the safety of a high-risk AI system have been met[69]. It appears to be arguable that the human oversight might be replaced if the results of the AI risk assessment allow the deviation from the law. The regulator merely wanted to use the principle of human oversight to create a safety net that applies in addition to the other requirements for high-risk AI systems.[70] It must therefore be ensured that, depending on the area of application of the high-risk AI system (think, for example, of high-precision medical devices based on a high-risk AI system), human oversight cannot lead to damage, i.e. the opposite of the standard under Art. 14(2) AI Act.[71] In this respect, human oversight is at odds with AI applications that are specifically designed to exclude continuous monitoring by humans.[72]

4. Cooperation between providers and deployer s

Although the wording of Art.14 AI Act differentiates between providers and deployer s, the better arguments favour an obligation of cooperation between provider and deployer. With regard to the obligation of human oversight, the AI Act primarily[73] imposes obligations on two actors: the provider developing the high-risk AI system[74] and the deployer using the high-risk AI system[75]. At first glance, the provider, as the developer of the high-risk AI system, is obliged to ensure human supervision in Art. 14(1) and (3) AI Act, which corresponds to the goal of avoiding risks as early as possible.[76] However, in the context of cooperation to ensure the purpose of human supervision in the meaning of Art. 14(2) AI Act, there is no implicit restriction of the protective purpose arising from the distribution of roles underlying the AI Act.[77]

In accordance with the basic concept of product safety law, the provider must ensure that human supervision can be carried out in accordance with Art. 14 AI Act (Art. 16(1) AI Act),

[69] On the interactions between human oversight and other requirements for high-risk AI systems, Enqvist, Law, Innovation and Technology 15:2 (2023) , 517 s.

[70] Hetmank/Meinel, KIR 2024, 127, 128.

[71] For example, a robot used for eye surgery, Gassner, MPR 2023, 5, 11.

[72] Level 3 and Level 4 vehicle systems could serve as examples, see BeckOK AI Law/Buchner, Art. 14 AI Act, para. 61

[73] See Dienes, MMR 2024, 456, 457s.; Hetmank/Meinel, KIR 2024, 127.

[74] Art. 3 No. 3 AI Act: "Any entity that 'develops or has developed an AI system [..] and puts it into service under its own name or trademark"; Recital 64, p. 2 AI Act.

[75] Art. 3 No. 4 AI Act.

[76] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 56 ss.; Hetmank/Meinel, KIR 2024, 127; Enqvist, Law, Innovation and Technology 15:2 (2023), 518.

[77] Rightly so, Hetmank/Meinel, KIR 2024, 127, 130.

while the deployer must, in particular, take appropriate technical and organisational measures to ensure that the high-risk AI system is used in accordance with the provider's instructions for use (Art. 26(1) AI Act). In addition, the deployer is obliged to assign supervision to humans who have the necessary competence, training and authority, as well as the support to supervise the high-risk AI system (Art. 26(2) AI Act).

It would be simplistic to separate the responsibilities of the provider and the deployer with regard to human oversight. It is true that both must ensure human oversight in different ways for their respective areas of responsibility. However, only through cooperation can it be ensured that the basic principle of human oversight is upheld. If unforeseen risks arise, the high-risk AI system must be adapted immediately, which affects not only the technical programme code but also the specific exercise of human supervision.

Even after the high-risk AI system has been placed on the market, the provider remains obliged to ensure system conformity or - in the event of a significant change (Art. 43(4) AI Act) - to re-establish conformity. Art. 20 AI Act obliges providers of high-risk AI systems to take immediate action if the system may no longer meet the requirements.[79]

The place that must define the cooperation between the provider and the deployer is usually the licence agreement. In addition to the specific cooperation obligations already existing in the AI Act[80], the obligation to cooperate in order to ensure the objectives of human oversight at all times should be formulated as a general clause. For example, it should be stipulated that (with respect to the trias of protected interests relevant) information from the use of the high-risk AI system regarding the effectiveness of human oversight must be shared immediately.

If there is no deployer of the high-risk AI system, but the high-risk AI system is made available to a private individual[81], the deployer obligations under the AI Act are meaningless *per definitionem*: Deployer does not include a private individual In that very case, only the other safety measures of the AI Act apply[82]. However, since high-risk AI systems can pose risks that require independent human supervision, it is likely to be highly risky in terms of liability law if providers leave high-risk AI systems to private individuals. It seems questionable whether it is sufficient in such cases to rely solely on warnings or detailed instructions for use[83]. If a provider agreed to the usage of the high-risk AI system by an individual person, the legal risks, especially regarding damages and penalties, should be taken into account. Given the legal uncertainty, it could be the case that

[78] For details, see Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 59.

[79] Hetmank/Meinel, KIR 2024, 127, 132.

[80] The duty to cooperate must be described as clearly as possible, taking into account the clear distinction that is necessary from a German perspective due to the otherwise imminent threat of covert temporary employment.

[81] Obligations of private deployer s are excluded, Art. 2(10) AI Act; on this issue, see Hetmank/Meinel, KIR 2024, 127, 132 et seq.

[82] Hetmank/Meinel, KIR 2024, 127, 132.

[83] However, see Hetmank/Meinel. KIR 2024, 127, 133.

providers will not make high-risk AI systems accessible to private persons. This would be a limitation of business chances and a limitation for international AI providers regarding the European market. Legal certainty in this respect is utmost desirable.

5. Effectiveness and appropriateness of human oversight

Oversight of the high-risk AI system must be effective and appropriate in accordance with Art. 14(1) and (3) AI Act, in each case with regard to measures and precautions[84]. The effectiveness of human supervision refers to its efficacy[85], which ensures the protection of the legal interests referred to in Art. 14(2) AI Act. Effectiveness and its control must be conceptually and organisationally ensured in the high-risk AI system. The more autonomously the system operates, the higher the risks to the legal interests protected under Art. 14(2) AI Act, and the stricter the supervisory measures that are necessary. The countermeasures for risk reduction or prevention must be taken by the provider as efficiently as possible[86] and within the limits of what is technically feasible. As described above, the provider (as well as the deployer) does not have to prevent every conceivable risk identified in the risk analysis; only those risks that result from the intended use of the high-risk AI system or its reasonably foreseeable misuse[87] or - despite compliance with all requirements (Art.s 8 et seq. AI Act)[88] - persist.[89] It is precisely the latter case that allows for experimental, innovative high-risk AI within limited boundaries - a zero-risk concept is not necessary[90].

In practical terms, it may be important to determine whether the provider is free to choose from the options available under Art. 14(3) AI Act and/or to what extent the criterion of technical feasibility must be determined objectively or subjectively from the provider's point of view.[91] Correctly - due to the protective purpose of Art. 14(2) AI Act - there is only a

[84] On the linguistically unsuccessful version, see Hetmank/Meinel, KIR 2024, 127, 129; Dienes, MMR 256, 458.

[85] Hetmank/Meinel, KIR 2024, 127, 128; Sterz/Baum/Biewer/Hermanns/Lauber-Rönsberg/Meinel/Langer, FAcct '24: Proceedings of the 2024 ACM Conference on Fairness, p. 2496 ss.; Green criticises the criterion of efficiency, Computer Law & Security Review, Volume 45 (2022), 105681, 9 ss.

[86] Hetmank/Meinel, KIR 2024, 127,130.

[87] Legally defined in Art. 3 No. 13 AI Act, i.e. expected incorrect conclusions by users during operation or unprogrammed functions (emergence) produced by high-risk AI systems themselves, cs. Recital 65, pp. 4 and 7 AI Act.

[88] In other words, risks that occur despite the risk management system (Art. 9 AI Act) or despite qualitatively sufficient training data (Art. 10 AI Act) and despite the technical documentation (Art. 11 AI Act), recording, and despite the transparency and provision obligations (Art. 12 and 13 AI Act) and despite compliance with the requirements for accuracy, robustness and cybersecurity (Art. 15 AI Act).

[89] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 75.

[90] On the balance between innovation and risk avoidance, see point 1.1 of the explanatory memorandum to the AI ActE-Kom; Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, paras. 2 and 81.

[91] Hetmank/Meinel, KIR 2024, 127, 130.

[92] Hetmank/Meinel, KIR 2024, 127, 130; Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 86; no genuine right to choose: BeckOK, AI Law-Buchner, Art. 14 AI Act para. 40.

[93] Hetmank/Meinel, KIR 2024, 127, 129.

choice between options for reducing/preventing risks if these options are objectively equally effective.[92] In this respect, contrary to the wording of Art. 14(3) AI Act, the provider cannot completely shift its obligations to the deployer.[93] This has been already made clear with respect to the joint cooperation principle and the AI risk assessment. However, the AI risk assessment may conclude that certain measures can only be effectively implemented by the deployer.[94] According to the TOP principle derived from European occupational health and safety law, technical measures must always take precedence over organisational or personal measures. Art. 14 (paras. 3 and 4) AI Act must be understood to mean that technical possibilities must be exhausted in any case and that the provider cannot hide behind the argument that 'it is too complicated or too costly'.[95] Providers must examine the extent (maximising the area of protection) to which measures for effective human supervision are already built in 'ex works' and must be available to the deployer from the outset. Human oversight must therefore generally be programmed into the AI code as supervision by design in such a way that it is more than just a placebo effect.[96]

Art. 14(3)(b) AI Act provides that the provider may also specify precautions that are to be implemented and enforced by the deployer in a reasonable manner within the specified framework.[97] The wording is not to be understood as a genuine alternative, but rather, due to the underlying principle of technical measures taking precedence over organisational and personal measures, should be regarded as a supplement at most[98] if the technical measures have been exhausted but, in view of the protective purpose, there are gaps in protection that the deployer can reasonably and efficiently close. It is disputed whether, in this case, the provider must also ensure that the deployer implements the measures and, if necessary, is liable for violations, which is largely denied.[99] At first glance, there is much to be said for placing the responsibility for implementation on the deployer and - in accordance with general liability principles - not holding the provider responsible for 'third-party actions' in this respect. However, as explained above, Art. 14 AI Act is based on the principle of cooperation between the provider and the deployer. Correctly, the wording does not imply a fundamental prohibition that the provider has no responsibility for the implementation of the measures transferred to the deployer . If Art. 14(3) AI Act is interpreted in the light of Art. 14(2) AI Act, which maximises the scope of protection, there should in any case be an obligation to check that the deployer complies with and is able

[94] Hetmank/Meinel, KIR 2024, 127, 129.

[95] A different view, according to which no preference is given to technical feasibility, is expressed by Hetmank/Meinel, KIR 2024, 127, 130.

[96] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 88.

[97] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 92 ss.

[98] Hetmank/Meinel, KIR 2024, 127, 130.

[99] No obligations to monitor implementation by the deployer are seen by Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 94, and BeckOK, AI Law-Buchner,Art. 14 AI Act, para. 42; differentiating in this respect Hetmank/Meinel, KIR 2024, 127, 131.

[100] See also Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 94 Hetmank/Meinel, KIR 2004, 127, 131.

to comply with the specified standards. In any case, the provider must ensure that the implementation of the supervisory measure is ‘easily’ possible for the deployer, i.e. that the technical path is largely smooth for the deployer.[100] In other words: The provider must reassure itself that the measures it has delegated to the deployer are understandable and implementable for the deployer . Due to the special nature of the high-risk AI system, it cannot exculpate itself by claiming that it did not address the foreseeable deployer risks in this regard and only licensed the high-risk AI system without considering its effects in the specific application environment and in the specific manner of use.

The wording of Art. 14(3)(b) AI Act is unclear in other places as well: the term ‘appropriate’ could also be understood to mean that it depends solely on the provider's point of view and could therefore also refer to measures that the deployer can only implement with external or additional help[101]. Since the associated question of responsibility and liability is disputed in this respect, this point should be regulated in the licence agreement for the high-risk AI system. Providers and deployers of high-risk AI systems are well advised to clarify responsibilities internally and explicitly regulate cooperation and information obligations.

The provider must also communicate the identified risks to the deployer in an understandable form and specify in the instructions for use (Art. 13 AI Act) which measures the deployer must take to avoid the risks. The recitals require the provider to ensure comprehensibility, i.e. the high-risk AI system must be designed in such a way that supervision is possible, i.e. the deployer understands the functionality of the system and can evaluate it, particularly with regard to its limitations and strengths.[102] Consequently, the evaluation of human oversight measures, including the associated technical measures, is part of the technical documentation of the high-risk AI system (Annex IV No. 2 (e) AI Act). However, the provider's obligation is not limited to providing information. Rather, in the case of high-risk AI systems, the provider must ensure, by means of an operational restriction, that the system cannot be used for purposes other than those intended or that the deployer cannot override the system (whether out of ignorance or intentionally).[103]

From the first impression, one could argue that the measures in Art. 14(3) AI Act (and also the requirements set out in Art. 14(4) AI Act) are incomplete with respect to the principles of fairness and anti-discrimination[104]. However, Art. 14(1) and (2) AI Act ensures that, for instance, the principle of fairness or the constitutionally protected prohibition of discrimination must always be taken into account in high-risk AI systems.[105]

[101] Hetmank/Meinel, KIR 2024, 127, 130 with accurate reference to the imprecise wording.

[102] Recital 72, sent. 2 AI Act.

[103] Recital 73, sent 3 AI Act.

[104] See Ebers et al., RDi 2021, 528, 534.

[105] BeckOK AI Law/Buchner, Art. 14 AI Act, para. 45.

6. Measures to enable supervision by the supervisor

Art. 14(4) AI Act contains further specific measures that must be taken to enable the supervisor to exercise human oversight. The following core duties regarding the measures can be detected:

Pursuant to Art. 14(4)(a) AI Act, the provider must ensure that the human supervisor can regularly and thoroughly check the high-risk AI system for typical AI errors (anomalies, malfunctions, unexpected performance/results).[106] The error control requirement in Art. 14(4)(a) AI Act corresponds to the transparency requirement for high-risk AI systems, according to which the deployer must be enabled to interpret and use the results/outputs of the high-risk AI system appropriately.[107]

Automation bias (Art. 14(4)(b) AI Act)[108] must be prevented in particular by technical protection mechanisms, e.g. activation mechanisms that consciously force the user to make decisions and make them aware of the limitations of the high-risk AI system by means of questions or information (cf. Art. 14(4)(c) AI Act).[109] Reference banners or a burden of proof are also proposed, which clearly demonstrate that the AI requirement has been addressed, in order to achieve the awareness of the need to interpret the result generated by the high-risk AI system, as required by Art. 14(4)(c) AI Act.[110] In the medical device sector, depending on the area of application, it may be necessary for the recommendation of the high-risk AI system to be independently validated by two experts or even by a team (e.g. cancer diagnostic and medical treatment).[111] The more the AI

[106] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 101 ss.; Possible measures include runtime monitoring, statistical methods for detecting unusual inputs/outputs (outlier detection), checking for deviations from training and input data (out-of-distribution detection), rule-based validation, tracking factors that influence AI decisions (feature attribution (SHAP, LIME)), visualisation tools, testing tools (A/B testing), detailed in Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st edition 2025, Art. 14 AI Act, para. 22.

[107] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st edition 2025, Art. 14 AI Act, para. 23.

[108] Automation bias refers to trust in machine-based decisions, fundamentally: Guijaro Santos, ZfDR 2023, 23, 28; EDPS TechDispatch on human oversight of automated decision making, 2025, p. 10; Martinie, Blackbox Algorithm, p. 59; Hill, DÖV 2014, 213 220; above footnote 10.

[109] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 106.

[110] BeckOK, AI Law/Buchner, Art. 14 AI Act, para. 55; dissenting opinion Bomhard/Merkel, RDi 2021, 276, 281: Awareness of automation bias is an obvious minimum requirement.

[111] Bomhard/Pieper/Wende/Bomhard/Witzke, 1st ed. 2025, Art. 14 AI Act, para. 25.

[112] For the critical view of the Austrian data protection authority on the AI system in job placement, see Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 117.

[113] Paal/Hüger, MMR 2024, 540, 541; Schwartmann, AfP 2024, 1, 7 (anchoring bias).

[114] Paal/Hüger, MMR 2024, 540, 541.

[115] Instead of many Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 107 ss.; Pilnoik, JZ 2022, 1021, 1027; Paal/Hüger, MMR 2024, 540, 541; BeckOK AI Law/Buchner, Art. 14 KI-VO Rn. 53; Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st ed. 2025, Art. 14 KI-VO, Rn. 24 s.

[116] Interpretation does not mean explainability, Dienes, MMR 2024, 456, 459.

[117] BeckOK KI-Recht/Buchner, Art. 14 KI-VO Margin. No. 57.

[118] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1. Aufl. 2025, Art. 14 KI-VO, Margin No. 26.

system only makes recommendations, the greater the requirements for these recommendations to be comprehensible. From an organisational perspective, circumstances that promote ‘blind trust’ in the high-risk AI system (e.g. time pressure[112], anchoring effects[113], pressure to justify rejection of the AI result[114]) should be excluded.[115] The supervisory authority's ability to interpret[116] the outputs of the high-risk AI system is not further described in the AI Act.[117] The possibility of interpretation is also related to the transparency requirement in Art. 13 AI Act. Technically, for example, the result could be explained with explanations in such a way that the deployer can assess its accuracy.[118]

Employers who use a high-risk AI system must not indirectly prevent critical questioning of the system by a supervisor by, for example, interpreting critical questioning as a negative contribution via KPIs or setting individual targets for the variable remuneration of employees that run counter to human supervision. A well-known negative example of the complexity of the elements presented is the application tool used by Amazon, which led to discrimination in the results.[119]

Art. 14(4)(d) AI Act provides for further elements of human supervision. It stipulates that human oversight must override and/or reverse the result generated by the high-risk AI system.[120] The legislative requirements reach their limits when the execution of the AI decision has already created a fait accompli, e.g. in high-frequency trading, AI-based control of vehicles/aircraft or AI-based traffic management systems/ power grids. Recommendation systems in particular quickly create facts that can no longer be reversed: if, for example, radicalising content is shown on TikTok or YouTube based on the evaluation of user behaviour, human control usually comes too late. Social media in particular try to keep users on the platform for as long as possible through recommendations, which works all the better neurobiologically through the human reward system the closer the recommendations to user behaviour. The length of time spent on the platform thus subliminally triggers mechanisms in the user that exclude or at least significantly impede their ability to disregard the AI results or to recognise errors and counteract them.[121]

In support of this, Art. 14(4)(e) AI Act introduce a stop button that allows human intervention in the system at any time.[122] However, the stop button is not practical in many high-risk AI systems because it interferes with business-critical processes, for example. Particularly in the case of AI-based control of vehicles/aircraft or in medical technology, the stop button can have fatal consequences if it is not designed to be context-sensitive.[123]

[119] Genovesi/Kaesling/Robbins/Gössl, Recommender Systems: Legal and ethical issues, p. 15; Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 118.

[120] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 124 ss.

[121] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st ed. 2025, Art. 14 AI Act, para. 28.

[122] Martini/Wendehorst, 2nd ed. 2026, Art. 14 KI-VO Rn. 130 ss.

[123] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st ed. 2025, Art. 14 KI-VO, Rn. 29; Bomhard/Merkle. RDi 2021, 276, 281.

Some literature argues that the requirements in Art. 14(4) AI Act can be relativised with reference to the specific characteristics of the high-risk AI system or its use.[124] The wording of paragraph 4 is unhelpful in this respect[125], but in the case of Art. 14(4) AI Act (as in the case of Art. 14(3) and (5) AI Act), the general principle of Art. 14(2) AI Act must always be taken into account. A deviation would only be conceivable if a threat to the protected goods in paragraph 2 is logically impossible.[126] Whether this is actually the case can only be answered on the basis of the above-mentioned AI risk analysis. When implementing high-risk AI systems, the effects of behavioural research and psychology, as specified by the regulator, must also be taken into account.[127] It may be useful to use psychological explorations to determine when and under what conditions human oversight functions efficiently with the measures provided for in Art. 14(4) AI Act, while at the same time minimising the burden on the user.[128]

7. Liability issues and burden of proof

The ambiguities in the wording of Art. 14 AI Act must be assessed in light of the high penalties imposed by the AI Act. Violation of the obligations referred to in Art. 14 AI Act is - indirectly - subject to administrative penalties.[129] Although the central provision of Art. 99 AI Act does not mention the human supervision obligation under Art. 14 AI Act, the provider is indirectly subject to the risk of a fine of up to €15 million or 4% of its global annual turnover via Art. 16 AI Act, to which Art. 99(4) AI Act refers; furthermore, the provider is indirectly subject to fines of up to €15 million or 4% of global annual turnover. Deployer s are also indirectly subject to fines under Art. 26(2) in conjunction with Art. 99(4)(e) AI Act. The same applies to the deployer.

In practice, however, the risk of being fined is - according to practitioners and Experts - frequently considered low in view of the open points in the wording of Art. 14 AI Act, at least until a conclusive administrative practice has been established. In this respect, there is reason to hope for a 'grace period', at least until end of this year. This is because the enforcement of the standard must be prepared. Although the undefined legal terms are to

[124] See, for example, the approach taken by Sterz/Baum/Biewer/Hermanns/Lauber-Rönsberg/Meinel/Langer, FAcct '24: Proceedings of the 2024 ACM Conference on Fairness, p. 1, 9.

[125] Fink, SSRN 202t, 1, 10.

[126] This is the conclusion reached by Fink, SSRN 2025, 1, 10.

[127] Gier-Reinartz/Kenning/Zimmermann, KIR 2024, 170, 170 s.

[128] For a clear illustration of this, see the study on the assessment of hate speech, Gier-Reinartz/Kenning/Zimmermann, KIR 2024, 170, 173.

[129] On the system of sanctions, see Martini/Wendehorst, 2nd ed. 2026, Art. 14 KI-VO Rn. 7-10.

[130] Art. 96(1)(a) AI Act.

[131] Notifying authority within the meaning of Art. 3(19) AI Act; market surveillance authority, Art. 3(26) AI Act.

[132] Art. 70(3)(48) AI Act.

be clarified by guidelines for practical implementation,[130] these guidelines are at best an aid to interpretation, i.e. courts must review the protection of fundamental rights, which is strongly emphasised in Art. 14 AI Act, in each individual case. The relatively low staffing levels of the authorities, at least in Germany, are an additional obstacle to the enforcement of the standard. The AI Act does oblige the national authorities[131] to enforce the requirements of Art. 14 AI Act.[132] In Germany, according to the planned draft law, the monitoring obligation is to be transferred to the Federal Network Agency. The current draft of the law implementing the AI Act[133] only provides for the creation of 33.2 permanent positions at the Federal Network Agency, and other authorities will not be staffed with additional personnel[134]. It seems questionable whether the staffing levels will be sufficient given the complexity of AI-related tasks and whether the authority will review the largely undefined legal terms of Art. 14 AI Act. Furthermore, the draft law implementing the AI Act[135] does not provide for any further direct sanctions for violations of human oversight.

Another factor of uncertainty is civil liability, i.e. liability for damage caused by a violation of human oversight. The AI Act does not contain any specific provisions in this regard. Civil liability remains primarily subject to the general rules[136]. The new Product Liability Directive[137] is intended to cover AI software and provides for strict liability in this respect. However, there are already doubts as to whether the directive, which is to be transposed into national law by the end of 2026, is suitable for the specificities of the AI sector is already being questioned[138], especially since the proposed AI liability directive with reversal of the burden of proof and a limited disclosure approach has currently failed. Nevertheless, a violation of the requirements of Art. 14 AI Act may lead to an easier enforceable claim for damages under Sections 823 and 831 of the German Civil Code (BGB). If the safety requirements for high-risk AI system as protective laws, then fault is in any case given if there is a lack of an effective and appropriate system of human supervision or if the system has been designed in a faulty or contradictory manner. If, for example, damage is caused by the high-risk AI system and it turns out that the supervisor did not have enough time to perform their duties or that there were structural ambiguities regarding their responsibilities, then the deployer can be assumed to be at fault in this respect. Insofar as the high-risk AI system licence agreement is classified as a contract

[133] See bmds.bund.de/fileadmin/BMDS/Dokumente/Gesetzesvorhaben/260209_RegE_KI-MIG_final_barr.pdf; on the legislative process, see bmds.bund.de/service/gesetzgebungsverfahren/gesetz-zur-durchfuehrung-der-ki-verordnung.

[134] E.g. no additional positions for the Federal Financial Supervisory Authority, 13 permanent positions for the Federal Office for Information Security, 8 permanent positions for the Central Office for Information Technology in the Security Sector, no additional positions for the Federal Commissioner for Data Protection and Freedom of Information; 5 permanent positions at the Federal Institute for Occupational Safety and Health.

[135] See § 15 bds.bund.de/fileadmin/BMDS/Dokumente/Gesetzesvorhaben/260209_RegE_KI-MIG_final_barr. pdf

[136] On the question of whether the security requirements for high-risk AI systems constitute duties of care under liability law or protective laws within the meaning of Section 823 (1) and (2) of the German Civil Code (BGB); Grützmacher CR 2021, 433; 437 ss.

[137] Directive (EU) 2024/2853.

[138] Bertolini, Artificial Intelligence and Civil Liability, 2025, p. 83 ss.; Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 12.

with protective effect in favour of third parties, since the protection of legal interests expressly applies inter omnes, a contractual claim for damages could also be considered. The interpretation of the extent to which Art. 14 AI Act protects third parties is therefore left to the courts. This can inhibit innovation.

The AI Act does not contain any information on the burden of proof. The burden of proof for the appropriateness of measures to ensure human supervision lies with the provider or, downstream, with the deployer : This follows from the principle that human oversight must be anchored as a guiding principle in every high-risk AI system. Against the background of Art. 14(2) AI Act, the provider cannot withdraw from its obligation under Art. 14(3) AI Act.[139] This is because economic risk is inherent in any development, and Art. 14 AI Act explicitly does not provide for any exceptions for economic considerations.

V. Delegation concept

Structurally, this concept is similar to delegation under occupational health and safety law. For effective delegation, Art. 14(4) AI Act. However, the supervisor must also be in a position to perform the supervisory task in terms of time, budget and instructions, which appears questionable in practice.[140] Conflicts of interest must be anticipated and must not interfere with the performance of supervisory duties. The delegation of human supervision is only effective if it is clearly defined and delimited in relation to other areas of responsibility. If human supervision is delegated, its correct exercise must be monitored by management. As the use of high-risk AI system, the scope and structure of human supervision must also be adapted to changed circumstances and/or regulatory conditions. The transfer of supervisory duties can - obviously - be transferred to employees of the deployer or to external personnel or to a central office. Unlike in the case of transfer under occupational health and safety law, no special form is prescribed, but for documentation purposes, it is generally advisable to put the transfer in writing. To ensure the effectiveness of supervision, care must be taken to ensure that a functioning substitution rule is in place. The AI Act assumes that human oversight must be delegated to natural persons (Art. 14(4) sentence 1 AI Act). In terms of content, the commentary attaches great importance to emphasising the expertise and specialist knowledge of the person exercising supervision, as the AI Act sets a high bar in this respect - despite the correction to the even stricter Commission draft - sets a high bar.[141] The supervisor must have in-depth knowledge of

[139] However, Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 91 takes a different view, arguing proportionality; however, this overlooks the fact that the focus on the three pillars of protection does not mention financial burdens, but only technical restrictions. If economic reasonableness were a factor, the legislator would have had to could have explicitly mentioned this in Art. 14 AI Act.

[140] Bomhard/Pieper/Wende/Bomhard/Wietzke, 1st ed. 2025, Art. 14 AI Act, para. 5 ss., esp. 12 s.

[141] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 97.

[142] Disagreeing Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 98.

the specific high-risk AI system, the algorithms and programming used in it, error identification and methods, which is likely to go beyond simply studying the operating instructions.[142] To a certain extent, these requirements conflict with the provider's need for confidentiality with regard to the algorithms underlying the high-risk AI system. The dilemma can be resolved by means of a test phase, in which the deployer - under the guidance of the provider - tests the high-risk AI system in the specific operating environment.[143] The information provided by the provider must be sufficiently comprehensible to enable the supervisor to take it into account on the basis of the objective recipient's perspective.

If there is a works council, is subject to co-determination in accordance with occupational health and safety law with regard to the selection criteria, the training criteria and the individual personnel measures usually associated with this (e.g. transfer, Section 99 Works Constitution Act (BetrVG)). However, the content of the transfer is not likely to be subject to co-determination, as the AI Act is considered a law within the meaning of the introductory sentence of Section 87 (1) BetrVG. In practice, it appears that the position of the supervising employee is regulated in the IT and AI framework works agreement.

VI. Summary

Although human oversight is designed as a ‘safety belt’ in Art. 14 AI Act, its implementation is not easy, so that its practical effectiveness can be doubted[144]. The wording raises a number of questions. The aim of the regulation to establish a safety net in this respect has only been partially implemented in terms of technical standards.[145]

The fundamental rights requirement calls for, requires an interpretation that allows for human intervention and supervision in all cases, unless the duty of supervision appears to be reasonably excluded after a comprehensive risk assessment based on the TOP principle. Depending on the technical complexity of the high-risk AI system, a more sensible departure from human supervision and a greater focus on technical perfection may also be a conceivable solution.[146] As in occupational health and safety law, the use of high-risk AI systems requires cooperation between providers and deployer s in the sense of activating the persons affected by the high-risk AI system for the purpose of protection.

Another decisive factor for the implementation of human supervision is a coherent and controllable delegation concept, which must be effectively implemented in accordance

[143] Martini/Wendehorst, 2nd ed. 2026, Art. 14 AI Act, para. 98

[144] Green, Computer Law & Security Review Volume 45 (2022), 1 ss.; Koulu, Maastricht Journal of European and Comparative Law 27 (2020), 7s9 s.; Elish, Engaging Science, Technology and Society (2019), 40; overview also in BeckOK, AI Law/Buchner, 4th ed. 2025, para. 31.

[145] Hetmank/Meinel, KIR 2024, 127, 133.

[146] Bomhard, DSRITB 2024, 421, 427.

with labour law. In addition to the qualification of the supervisor, this requires, in particular, the necessary (financial and time) budget to be able to carry out human supervision efficiently.

It appears to be necessary to discuss the concept of human oversight in international context. For instance, the experiences and guidelines of Japan could also be helpful to develop a beneficial and mutual trusted AI. In this context, conferences like this Conference in Kyoto can serve a s a platform to develop world-law-standards combining academic and practitioners' views.