International Law Review of AI, ML and Robotics
Volume 1 Issue 1, 2026

From the concept of "Human Control" in international law to the concept of "Human Oversight" in the current European Union Act on Artificial Intelligence

R. Shrikant Kaushik
Assistant Professor
Department of Artificial intelligence

Article Info
Received: 20-12-2025 Revised: 22-01-2026 Accepted: 05-02-2026 Published: 15-02-2026

Abstract: The term "human control" of autonomous systems has been replaced with "human oversight" of artificial intelligence, which is the topic of this article, which investigates the recent change in focus that has been taking place. While the first concept has been widely discussed in international legal talks for close to ten years, the second concept is the word that is being used in the European Union's most recent proposal for an artificial intelligence act. After that, the paper delves into the concepts of "control" and "human oversight" in relation to the European Union Artificial Intelligence Act. It places these concepts within the larger context of the Act, and it highlights both the shortcomings of the Act as well as potential solutions to enhance it.

1 Introduction

The word "oversight" has two distinct connotations in everyday speech. It first explains what it means to supervise something. Second, it describes an inadvertent failure during a task.1 As a result, when someone is in charge of anything, she either supervises it or makes a mistake, usually by forgetting important details. The draft EU AI Act2's article 14 promotes "oversight" as a normative standard for the interaction between artificial intelligence (AI) and human operators. The EU AI Act may not, however, adequately direct the path toward appropriate human oversight of AI. The second definition of supervision may perhaps be more accurately reflected in article 14. This is due to the fact that humans often fail to effectively supervise AI in intricate human-machine interactions. As a result, significant human control of AI may ultimately become difficult, making genuine understanding of its operations unattainable.

In light of this, the current chapter delves more into the problem of human monitoring of AI. In particular, it clarifies the connection between "human oversight," a new word suggested by the EU AI Act to

characterize the interaction between human operator and algorithm, and "human control," a phrase often used in international law when referring to intelligent systems. This is how the chapter goes. Article 14 of the proposed EU AI Act on human supervision, according to Section 1, is based on two narratives: The first is the issue over human control of autonomous weapon systems (AWS) under international law; the second is the extensive topic of ethical AI. The EU AI Act's choice of the phrase "oversight" instead of the more often used term "control" to distance itself from the debate over autonomous armed systems is then examined in Section 2. This shift in language might be explained by the fact that debates on AWS mostly focused on embodied systems and seldom produced any concrete outcomes. This interpretation, however, is incorrect since AI and system autonomy are intimately related. Rather, the Section proposes that the EU AI Act views control as the general framework and monitoring as one specific facet of human-machine interactions. To put it another way, monitoring is just one aspect of the pursuit of human control over AI; control is what the EU AI Act seeks generally. Section 3 then demonstrates how human monitoring, in accordance with Article 14 of the EU AI Act, strives to guaranty human agency and responsibility. It examines two important obstacles to accomplishing these goals: the peculiarities of human psychology and the potential for accountability gaps to arise. Even though these obstacles may appear manageable, Section 4 shows that the EU AI Act's Article 14 is based on erroneous assumptions. And the criticism doesn't stop there. The next two parts demonstrate that the apparent flaws relate to the EU AI Act's overall framework, which is both too aspirational in its claims and too narrow in its scope. In order to address the scope-related issue, Section 5 examines several shortcomings in the Act's risk-based methodology. The EU AI Act's failure to sufficiently safeguard basic rights is then explained in Section 6. In light of this, Section 7 makes the case for amending the EU AI Act—more especially, Article 14—to better reflect human capacities and uphold basic rights. Section 8 continues with some thoughts on the expected broad scope of the EU AI Act, both in terms of content and geography.

2 Two Narratives

The clause on "human oversight," included in Article 14 of the proposed EU AI Act, must be interpreted in light of two different narratives: (i) The conversations about autonomous weapon systems, which focus on using kinetic force via a specific kind of embodied system—weapons. (ii) The more general ethical discussion of AI, which addresses algorithms more thoroughly regardless of whether they are a component of embodied or disembodied systems.

(i) The Group of Governmental Experts (GGE) deliberations under the Certain Conventional Weapons Convention (CCW3) serve as the primary point of reference for the first narrative.4 Although non-governmental groups have been actively involved in influencing these conversations, they have been held in a traditional international law venue.5 Although the idea of controlling AWS6 has garnered a lot of interest in the conversations of the CCW-GGE, neither a specific definition of "meaningful human control"8 nor an internationally accepted definition of AWS7 have surfaced. The 11 Guiding Principles on Lethal Autonomous Weapons Systems are the sole concrete result.9 However, these principles remain ambiguous on the important topics and are not legally obligatory. However, there are three primary issues10 that have dominated the global discussion on autonomous weapon systems. First, there is the issue of legal compliance, or, in the context of war, how to make sure that the employment of autonomous weapon systems does not go against the fundamental tenets of international humanitarian law.Twelve The second issue is the potential for a responsibility gap.

when people are no longer able to decide whether to employ deadly force or carry out other activities that are sensitive to human rights.13. Third, a moral actor must constantly make judgments that may violate human rights in order to uphold human dignity14.15. Surprisingly, these three AWS-related issues have found resonance in other legal domains and have influenced the larger conversation about autonomous systems.16. As a result, the controversy surrounding AWS has had a significant impact on the global (legal) discourse around human control over intelligent systems.17.

(ii) The ethics of AI has been the subject of the second, more varied

story. These conversations have taken place in a number of forums, including academic institutions, civil society, private businesses, and regional organizations like the EU, the Council of Europe, and the OECD. These forums sometimes overlap. Discussions were first motivated by ethical issues, such as the issue of human control over AI.24. In many cases, the conversations have led to codes of behavior and ethical principles.25. These instruments lack legally enforceable standards and efficient enforcement mechanisms, despite the fact that they are by no means unimportant under international law.27.

In order to better confront AI, more recent national and regional initiatives seek to create hard legal requirements.28. These bids are often influenced by and evolve concurrently with soft law. The proposed EU AI Act establishes a general, legally enforceable norm for all EU member states.29. There will be minimal to no space for different national legislation, depending on the degree of harmonization.Thirty

3 Dissociation from Control of Autonomous Weapon Systems — or Control *and* Oversight of AI?

The EU AI Act seems to steer clear of the idea of control on a merely textual interpretation. Rather, it chooses the ostensibly weaker concept of supervision. Given that "human control" has been a major topic in AWS conversations and has also taken center stage in more general AI concerns, this shift in nomenclature is startling.31.

The EU AI Act may rely on the concept of "oversight" in order to avoid the deeply held beliefs about "human control" that have characterised AWS discussions. According to the CCW-GGE, embodied systems with a certain level of autonomy are linked to the pursuit of control. Therefore, it seems that the EU AI Act has excellent grounds to depend on "oversight" rather than "control" of AI. In this regard, the Act might be seen as a dissociative act.

On closer inspection, however, the dissociation theory is refuted. Since AI fuels a large portion of system autonomy, separating AI from autonomous systems would be unnatural and pointless.32. AI is responsible for many of the issues that system autonomy has raised.33 Interestingly, the legal issues surrounding the use of AI in a variety of industries are similar to

those brought up in relation to AWS.34. Therefore, separating human supervision under article 14 of the EU AI Act from earlier efforts to capture the essence of control would not be practical.

In fact, the EU AI Act's Title III, Chapter 1 seems like a checklist that may have resulted from the Geneva negotiations. However, compared to the majority of earlier recommendations for a human control standard in the CCW-GGE, the standards outlined in Article 14 of the EU AI Act are more comprehensive and specific.35 Article 14 draws on a number of strategies created by academics, practitioners, and civil society.38. Therefore, "human oversight" under article 14 of the EU AI Act should not be interpreted as a compromised or weaker version of "human control."

What function does "control" serve in the proposed EU AI Act, however, if "oversight" does not take the place of "control," particularly as the Act makes no mention of it? The Act seems to provide a high-level control plan with its general method of categorising AI systems into several risk-categories. In an effort to safeguard human health, safety, and basic rights, the EU AI Act40 forbids certain AI practices41 and imposes a wide range of obligations on suppliers and users of so-called high-risk AI systems.42. The maintenance of a risk management system, data governance, technical documentation and record-keeping, transparency obligations, duties pertaining to system accuracy and robustness, and human oversight are just a few of the desiderata that these duties flesh out and are reminiscent of the larger AI-control debate.44 When all of these responsibilities are carried out correctly, it should result in human control over AI. In this way, the European Commission's entire strategy for regulating AI reflects control. Article 14 on human supervision is just one element that contributes to human control of AI within this larger framework. However, Article 14 of the EU AI Act is quite significant: It is intended to provide meaningful human-machine interaction45, which is essential to human control of AI. In essence, article 14 on human supervision offers a fair level of detail for designing human-machine interaction. In that sense, human supervision is tangible and concerned with the real deployment of AI in particular circumstances. Conversely, control is more

expansive and abstract. It serves as the main goal of the EU AI Act by connecting the various elements. Expanding on a metaphor often used in the Geneva AWS debates, control may be seen of as the iceberg with all of its buried parts. In contrast, oversight just compensates for the tip that is visible.46.

4 Human Agency and Accountability through Oversight

Building on the idea that supervision contributes to the broader concept of control, human oversight may be said to support two significant goals: (i) ensuring and strengthening human agency47 and (ii) ensuring accountability.48.

(i) It is crucial to further define the concept of supervision in order to address the first goal. From a conceptual standpoint, supervision of AI refers to a certain kind of human intervention during the operation of an AI system.49. More precisely, supervision aims to guarantee that human judgement is present when an algorithm makes a choice that has an impact on a person.Fifty Human agency is involved in this situation since people can only provide a fair assessment when they are given a significant chance to understand the issue and make their own well-informed decisions.51. In this way, maintaining human autonomy is a prerequisite for human agency.52. Therefore, it is necessary to counteract the overpowering impact that AI tends to have on people, as shown by the excessive dependence of humans on the output of AI systems.53. Automation bias, another name for over-reliance,54 may lead to both commission and omission mistakes.55 The latter includes circumstances in which people mistakenly heed the AI system's guidance, while the former refers to inability to react to system abnormalities.56. Human psychology and mental capacities are the root causes of both mistakes: Reduced cognitive focus or a lack of system comprehension are common causes of omission errors.57 In contrast, errors of commission might be linked to a belief in the superiority of system judgements as well as a failure to take into account and appropriately interpret the information offered by the system.58. Another issue is that when people engage with an AI system, they can feel powerless, which would reduce their feeling of moral agency.59.

As a result, the EU AI Act and other (legislative) documents60 make an effort to appropriately account for

human psychology and cognitive capacities. This is especially crucial for recommender and decision support systems that are now used in a variety of different industries, such as banking, law enforcement, the court, transportation, medical, the military, and, of course, everyday life.67.

Human supervision may be broken down into three parts, notwithstanding the absence of a common definition.68. First, by getting sufficient information about the system's processing and results, the human operator69 will be able to comprehend how the system operates.70. This involves giving the operator the ability to analyse and challenge the algorithmic output and, as a result, add a little amount of judgement to the final choice.71. Second, the human operator will be able to override the system's output if they disagree with it.72. Third, the person working with the AI system must continue to be conscious of the automation bias tendency.73. The human operator is thought to actively engage with the system and maintain a feeling of empowerment if all three components are present in adequate amounts.74.

The three aforementioned components are included in article 14(4) of the EU AI Act. It asserts that "individuals to whom oversight is assigned"75 must have the ability to comprehend, analyse, evaluate, and veto an AI system's output. The person will also have the ability to stop the system's processing or even interfere with the operation.Thus, it seems that Article 14 of the EU AI Act77 does justice to the fundamental conditions required to guarantee and reinforce human agency. Additionally, it appropriately recognises the current technological trend that unites people with AI and allows them to work together naturally in accordance with their own skills.78. Therefore, human supervision under article 14 of the EU AI Act may be considered "meaningful," expressing a normative expectation utilising a major qualification for human control recognised from the AWS discussion.

(ii) The second goal of human monitoring is to guarantee responsibility.79. According to a transdisciplinary perspective, accountability is a kind of "backward-looking responsibility."81. It gives a retrospective account of what happened and assigns blame for behaviours, particularly those that cause damage to other people.82. However, in order to properly assign

responsibility for a specific result, an individual must have control over the action (or inaction) that resulted in that result.83. further precisely, three requirements must be met in order to hold someone responsible for a result (although national laws may add further nuance):84 (1) The human agent has to be capable of responsible behaviour. (2) Either an act or an omission must establish a causal connection between the actor and the result. (3) There must be a plausible suspicion that the agent did something improper on purpose or had the opportunity to stop anything improper from occurring.85.

Additionally, assigning responsibility fulfils inspection duties, such as requiring reasons and imposing consequences in the absence of such justifications.86. In this manner, victims of damaging actions or omissions will get a feeling of revenge.87.

In light of this, two issues with joint human-AI job fulfilment stand out. First, "responsibility, resp. accountability gaps"88 emerge when it is impossible to assign blame. In highly developed human-machine interactions, this is probably going to be the case. It is infamously difficult to determine the true source of injury due to the complexity and unparalleled speed of data processing.Ninety In a similar vein, when crucial aspects of the joint job fulfilment are assigned to the machine, it is debatable whether the human operator still has the ability to behave responsibly and may legitimately be accused of misconduct.91 Machine learning exacerbates the issue since it might be difficult, if not impossible, to comprehend the system's processing and output.92. Therefore, "accountability gaps" are far from fiction as it is only possible to hold someone legally liable if it can be precisely determined who intentionally or deliberately caused a negative event.

Second, there is a problem with inspection. Sanctions may only be applied to people who are deemed accountable for a certain result, even when an intelligent system may provide explanations or excuses.95. Retributive justice cannot be achieved in any other way.96. However, as it is difficult to pinpoint accountability in joint human-AI job completion, it would not be appropriate to penalise individuals who participate in human-machine interactions in a single, perhaps subordinate role.

As a result, it has become common practice to include a human operator

in the "decision and control loop."97. This loop suggests that the human operator must always be able to keep an eye on and comprehend the system's functioning, interpret its output, and have a significant impact on the conclusion.98. To put it another way, no decision or action can be made without significant and verifiable human participation in collaborative human-AI job completion.99 It could be conceivable to assign accountability and place blame on the shoulders of the human operator under these principles.

As a result, monitoring under article 14 of the EU AI Act seems to fairly effectively handle issues of accountability and inspection. The following concept forms the basis of the article: The human operator will gain the decisive position in combined human-machine decision-making if they satisfy the agency-promoting standards of article 14(4). As a result, they may be held responsible for a negative outcome.A hundred

However, private individuals will not be able to file claims against AI system providers or users because the proposed EU AI Act is intended as an administrative measure governing the relationship between state authorities and those offering AI systems commercially on the market or using them for professional purposes.102 Therefore, supervision does not create responsibility, even while it seeks to guarantee accountability.

5 Does Oversight Overlook Something?

Article 14 of the EU AI Act is likely to clear the path out of the complexities of managing AI systems with human agency and responsibility clearly in sight. Article 14(4) seems to establish the groundwork for meaningful human-machine interaction by outlining five elements that are seen to be essential for guaranteeing human agency and, subject to that, for legitimately assigning liability to the human operator. This is in line with the Act's overarching goal of creating a thorough framework for human oversight over AI. Recent studies, however, indicate that some prudence is necessary. Human abilities might fall short of the normative obligations.103 Both human agency and responsibility are affected by the ensuing disparity.104 Fulfilling the three composite criteria mentioned in Section 4(i) is expected to lead to human agency. In reality, however, the conditions are either difficult to fulfilll or unlikely to enhance human decision-making in collaborative job completion.105 The ability of humans to comprehend, analyse, and challenge the output of AI systems is still restricted, despite all attempts to

improve system transparency and provide explanations. The comprehension of AI systems by a human operator is seldom sufficient.107 This relates to the fact that AI systems' data mining skills are greater than human cognitive capacity, which is why AI is used in the first place.108 In addition, explanations are a two-edged sword. Even when the system is flawed, they could have the negative consequence of encouraging more faith in it.109 Another issue is the possibility that people might interfere with AI system processing or alter its results. Research indicates that it is difficult for human operators to accurately assess the quality of algorithmic output.110. Therefore, there is no assurance of a better result even if an operator chooses to overrule an algorithmic output.111 Lastly, even with instructions to confirm an AI system's output, automation bias is difficult, if not impossible, to eliminate, particularly when time is of the essence.112.

> The human operator may ultimately wind up just rubber-stamping the judgements recommended by the AI system, despite the standards outlined in article 14(4) of the EU AI Act being promising.113. Human supervision will ultimately make a false, or at least too optimistic, promise if normative expectations and actuality diverge. In some situations and circumstances, it won't maintain human agency.114.

The second goal of human supervision, accountability, is not improved by Article 14 of the EU AI Act. Accountability is dependent on the exercise of human agency since it necessitates the exact allocation of duties. However, it would not be appropriate to (just) hold the human operator liable since human agency can barely be guarantyd to a sufficient degree. Therefore, a "moral crumple zone" 116 could develop if politicians adhere to the appealing but ultimately unachievable concept of supervision.117. The phrase refers to a situation in which a human actor with little influence over the autonomous system is mistakenly held accountable for a certain result.118. Therefore, in very complicated human-machine interactions, the human operator may be reduced to a simple anchor for attribution of culpability in the event that the system malfunctions.119. Human operators eventually become "liability sponges," or scapegoats.

They often fail to avoid algorithmic mistakes.122 In high-stakes decision-making, human monitoring therefore runs the potential of becoming a fig leaf or cover-up device for risk-related issues.124

Article 14 of the EU AI Act, however, highlights more issues than only the practical viability and efficacy of human control. The EU AI Act's risk-based strategy and the protection provided for basic rights are additional issues. These issues relate to the larger framework of the EU AI Act, particularly its Title III on high-risk AI systems, even if they may be found in Article 14 of the Act. Overall, these worries highlight possible shortcomings in the high-level control approach of the EU AI Act, which involves regulatory125 and social126 aspects in addition to addressing the human operator of AI systems.

6 The Risks of a Risk-based Approach

The EU AI Act employs a sophisticated risk-based approach, which may be the cornerstone of the high-level control strategy it pursues. This approach is motivated by the aim to limit the risks arising from diverse AI applications. Depending on the threats AI systems pose to human health, safety, and basic rights, the Act imposes an increasing number of requirements on providers of AI systems.127 Starting at the "top," a few AI systems that provide intolerable dangers to basic rights are prohibited under Title II, Article 5. The EU AI Act's most comprehensive provision, Title III, which also includes Article 14 on human supervision, identifies AI systems that pose a significant danger to protected goods128 and subjecting them to the regulations outlined in Chapter 2. Bots, emotion detection and biometric classification systems, and apps creating synthetic content (also known as "deepfakes") are subject to specific disclosure requirements under Title IV, Article 52.129 Lastly, standards of behaviour for any other AI systems that are not thought to present a significant danger may be freely enacted under Title IX, article 69.

One may argue that defining distinct risk classifications stratifies problems and worries.130 The EU AI Act, at least on its surface, provides sufficient space for innovation while addressing the most pressing risks AI presents to human autonomy, well-being, and basic rights131.132 It seems obvious sense to include categories of undesirable and high-risk AI systems133 in this balancing process.134 Notably, disembodied AI systems that pose a danger to health, safety, or basic rights are included in the high-risk category in addition to embodied ones. Decision-support and recommender systems in fields like education, career training, and employment are among the variety of disembodied high-risk AI systems. Additionally,

this category includes advising systems for law enforcement, such as those that assess additional risk factors or forecast recidivism.135

However, the risk-based approach's present architecture has several dangers. Remarkably, the EU AI Act does not define "high-risk AI."136 Rather, it provides AI providers with legal clarity by providing lists of high-risk AI applications137.138 Nevertheless, the more nuanced truth may be missed by these rankings.139In essence, whether the lists of high-risk AI systems in Annexs II and III are adequately thorough and current determines how much protection the EU AI Act provides for health, safety, and fundamental rights.140

Regulating a dynamic technology and staying abreast of the threats it brings are undoubtedly challenging. The fact that the EU AI Act already seems out of date in certain ways, although still being drafted, serves as evidence of this. Certain AI applications that seriously jeopardise basic rights141 do not fit under any of the eight headings142 in Annex III, nor do they feature in the lists in Annexs II and III.143 This discrepancy is not shocking since dangers are context-dependent and may only become evident gradually, making it challenging to fully recognise them beforehand.144. Furthermore, there will undoubtedly be hazards that we are now unaware of as the science of artificial intelligence develops quickly and produces ever-new, creative uses.145

The EU AI Act gives the European Commission the authority to modify the scope of Title III on high-risk AI systems in recognition of this dynamic.146 However, the Commission's authority is restricted: Article 7(1) of the EU AI Act states that only Annex III (listing high-risk AI by domain areas) may be updated; Annex II (listing EU law that, if relevant, defines AI as high-risk) cannot.148 However, even with regard to Annex III, the Commission's authority is limited since it may only modify the list of high-risk AI systems within the boundaries of the current eight headings (or domain areas).149 The Commission cannot create a new domain area to cover a danger that a new AI application poses.150 The application may only try to fit inside the current categories.151. For instance, an AI system intended to keep an eye on kids at home may prove to be quite dangerous. If the Commission wants to address the danger in this situation, it can

Try to fit the application under a title like "education and vocational training."152 However, the Commission would have to initiate a comprehensive modification of the Act involving all of the EU's legislative bodies if it thought this was too risky.

The EU AI Act's Article 84(1) expands on the clauses restricting the Commission's flexibility. It says that the list in Annex III may only be reviewed once a year, which makes it difficult to respond quickly to emerging hazards. Lastly, only high-risk AI systems are subject to the Commission's authority to amend the list of dangerous AI systems.153 It cannot add AI systems that need more transparency to article 52.155 or broaden the list of illegal AI activities in article 5 of the EU AI Act 154.

With its categorical differences and limited opportunities for significant modifications, the list-based method consequently inherits some degree of inertia.156 The list-based method as it exists now is probably going to overlook the procedural complexity of developing, deploying, and using AI systems.157 It may not properly recognise how AI affects people's everyday life, including the fulfilment of their basic rights.158

Based on this, it is safe to challenge the list-based approach to hazards associated with AI. It is noteworthy, in fact, how lightly the EU AI Act treats low-risk AI systems. This is not to say that every AI system has to be strictly monitored since there can always be dangers. However, rather than a straightforward (and incomplete) list of the most hazardous AI uses, a discretionary evaluation in light of the dangers each particular application poses for basic rights would better serve the goals of the AI Act.159 For example, certain recommender systems are exempt from article 14 of the EU AI Act on human monitoring since they do not fit into the high-risk category of Title III. But if an AI system has a big influence on people's everyday life and basic rights, shouldn't the people who are impacted by it have some control over it? The EU AI Act's possibly deceptive claims on the safeguarding of basic rights follow from this finding.

7 Protection of Fundamental Rights

The issues and deficiencies pertaining to both article 14 of the EU AI Act (i) and the intended control strategy of the EU AI Act as a whole (ii) should be addressed by a number of initiatives.

(i) In order to prevent the human operator from just rubber-stamping the choices made by the AI system and becoming a mere anchor for responsibility attribution220 in

the event that anything goes wrong, article 14 of the EU AI Act has to be improved. In essence, human insight should be the basis for supervision under article 14 of the EU AI Act, which mandates that the system be successfully monitored in the specific use case. This might result in the exclusion of certain use cases and situations where it is not possible to ensure that an AI system would be used in a way that complies with the law. To put it another way, an AI system shouldn't be employed if insight is unattainable in certain situations.222 For instance, the outcome will be obvious and unavoidable if the system is inadequately dependable, the work is crucial, the response time is really quick, and there is a serious risk to basic rights: The usage of the corresponding AI system in the current situation is prohibited since it cannot be implemented legally.223 This should be made explicit in Article 14 of the EU AI Act. It does not, however, imply that the system or its usage must be completely banned. Therefore, there is now no need to extend article 5 on forbidden AI methods since article 14 of the EU AI Act is more context-sensitive. Notably, an amended version of Article 5 of the EU AI Act that permits

Since the Union's jurisdiction currently rarely covers the substance of today's article 5, it would hardly be possible to add to the list of illegal AI techniques.

However, how can effective supervision be conducted by human insight? First and foremost, article 14 should provide significant practical training requirements,224 clearly outlining the responsibilities of both the supplier and the user of high-risk AI systems.225 Training need to be tailored to the specific application situations.226 Additionally, before exerting supervision, human operators must to be required to demonstrate their degree of training and expertise.227. Additionally, they must to be able to ask for assistance.228 Additionally, as human operators must comprehend and interpret an AI system's output, article 14 should take into account the status of explainable AI (XAI) to clarify provider and user responsibilities in this area.229 Last but not least, article 14 should mandate ongoing observation and simple procedures for reviewing compliance.230

(ii) At a higher level, several aspects of the EU AI Act should be re-examined, notwithstanding the fact that its risk-based approach seems appropriate.232 A key component of the risk-based approach should include an evaluation of basic rights.234 Therefore, several changes will be necessary if the EU maintains the list-based approach to AI-induced risks for legal certainty: 235 Annexs II and III should likewise allow for flexible changes under the EU AI Act.236 Therefore, the current titles of Annex III should not necessarily limit the Commission's authority under article 7(1) of the EU AI Act. In order to respond to recent developments, the Commission should be given the authority to quickly but thoughtfully add additional titles to the list in Annex III.237 For this

For this reason, the Commission need to evaluate Annex III on a regular basis without being constrained by a yearly schedule.238 Additionally, Annex III's list of AI systems might be constructed more transparently, with the mentioned AI systems serving merely as examples of legal standards. The reasoning is obvious: There would be greater leeway for interpretation if the list of AI systems were not a concrete, closed set. Lastly, it ought to be feasible to amend Title IV, Article 52 to increase its flexibility.239 Overall, this would make it feasible to keep the planned high-level control approach from falling behind actual advancements.

Even more daring initiatives to defend basic rights shouldn't be abandoned without careful thought. To make up for the power imbalance in respect to those who are subject to AI systems, one option is to make AI providers and users direct bearers of basic rights duties240.241

Individual rights to seek remedy might be included within the EU AI Act itself, together with strong oversight and enforcement procedures.242 In addition, a system for filing complaints about violations of basic rights might be set

up.243 Last but not least, two significant rights for people might support the proposed procedural rights: a right to information that enables impacted parties to effectively assert their rights and a right not to be exposed to AI systems that disproportionately harm basic rights244.245 These recommendations will undoubtedly improve the EU AI Act's standing for the individual human being and make it truly human-centric.246 In essence, the EU AI Act's social control component would become more prominent and effective by improving the status of impacted persons via directly enforceable rights.

The EU AI Act may be able to achieve the aim of genuine human control of AI if it is updated in accordance with the following recommendations on a refined definition of human supervision and a revised risk-based strategy that adequately recognises fundamental rights. However, it is important to remember that the EU AI Act will continue to be a kind of guidance. A readable act of appropriate dimensions can only be maintained by outsourcing some details. For instance, sector-specific regulations are more suited to address the issue of what specialised training is needed. In a metaphorical sense, all AI-related topics will be covered by the EU AI Act's broad control framework, which will also provide direction for defining sector-specific laws.

8 Oversight based on Insight

The issues and deficiencies pertaining to both article 14 of the EU AI Act (i) and the intended control strategy of the EU AI Act as a whole (ii) should be addressed by a number of initiatives.

(i) In order to prevent the human operator from just rubber-stamping the choices made by the AI system and becoming a mere anchor for responsibility attribution220 in the event that anything goes wrong, article 14 of the EU AI Act has to be improved. In essence, human insight should be the basis for supervision under article 14 of the EU AI Act, which mandates that the system be successfully monitored in the specific use case. This might result in the exclusion of certain use cases and situations where it is not possible to ensure that an AI system would be used in a way that complies with the law. To put it another way, an AI system shouldn't be employed if insight is unattainable in certain situations.222 For instance, the outcome will be obvious and unavoidable if the system is inadequately dependable, the work is crucial, the response time is really quick, and there is a serious risk to basic rights: The usage of the corresponding AI system in the current situation is prohibited since it cannot be implemented legally.223 This should be made explicit in Article 14 of the EU AI Act. It does not, however, imply that the system or its usage must be completely banned. Therefore, there is now no need to extend article 5 on forbidden AI methods since article 14 of the EU AI Act is more context-sensitive. Notably, an amended version of Article 5 of the EU AI Act that permits

Since the Union's jurisdiction currently rarely covers the substance of today's article 5, it would hardly be possible to add to the list of illegal AI techniques.

However, how can effective supervision be conducted by human insight? First and foremost, article 14 should provide significant practical training requirements,224 clearly outlining the responsibilities of both the supplier and the user of high-risk AI systems.225 Training need to be tailored to the specific application situations.226 Additionally, before exerting supervision, human operators must to be required to demonstrate their degree of training and expertise.227. Additionally, they must to be able to ask for assistance.228 Additionally, as human operators must comprehend and interpret an AI system's output, article 14 should take into account the status of explainable AI (XAI) to clarify provider and user responsibilities in this area.229 Last but not least, article 14 should mandate ongoing observation and simple procedures for reviewing compliance.230

(ii) At a higher level, several aspects of the EU AI Act should be re-examined, notwithstanding the fact that its risk-based

approach seems appropriate.232 A key component of the risk-based approach should include an evaluation of basic rights.234 Therefore, several changes will be necessary if the EU maintains the list-based approach to AI-induced risks for legal certainty: 235 Annexs II and III should likewise allow for flexible changes under the EU AI Act.236 Therefore, the current titles of Annex III should not necessarily limit the Commission's authority under article 7(1) of the EU AI Act. In order to respond to recent developments, the Commission should be given the authority to quickly but thoughtfully add additional titles to the list in Annex III.237 For this

For this reason, the Commission need to evaluate Annex III on a regular basis without being constrained by a yearly schedule.238 Additionally, Annex III's list of AI systems might be constructed more transparently, with the mentioned AI systems serving merely as examples of legal standards. The reasoning is obvious: There would be greater leeway for interpretation if the list of AI systems were not a concrete, closed set. Lastly, it ought to be feasible to amend Title IV, Article 52 to increase its flexibility.239 Overall, this would make it feasible to keep the planned high-level control approach from falling behind actual advancements.

Even more daring initiatives to defend basic rights shouldn't be abandoned without careful thought. To make up for the power imbalance in respect to those who are subject to AI systems, one option is to make AI providers and users direct bearers of basic rights duties240.241

Individual rights to seek remedy might be included within the EU AI Act itself, together with strong oversight and enforcement procedures.242 In addition, a system for filing complaints about violations of basic rights might be set up.243 Last but not least, two significant rights for people might support the proposed procedural rights: a right to information that enables impacted parties to effectively assert their rights and a right not to be exposed to AI systems that disproportionately harm basic rights244.245 These recommendations will undoubtedly improve the EU AI Act's standing for the individual human being and make it truly human-centric.246 In essence, the EU AI Act's social control component would become more prominent and effective by improving the status of impacted persons via directly enforceable rights.

The EU AI Act may be able to achieve the aim of genuine human control of AI if it is updated in accordance with the following recommendations on a refined definition of human supervision and a revised risk-based strategy that adequately recognises fundamental rights. However, it is important to remember that the EU AI Act will continue to be a kind of guidance. A readable act of appropriate dimensions can only be maintained by outsourcing some details. For instance, sector-specific regulations are more suited to address the issue of what specialised training is needed. In a metaphorical sense, all AI-related topics will be covered by the EU AI Act's broad control framework, which will also provide direction for defining sector-specific laws.

9 Conclusion: Wide and Deep

The EU AI Act is not yet the comprehensive, future-proof, and basic rights-sensitive piece of law that it claims to be. This does not negate the Act's significant advantages. It does provide a significant framework for control that addresses the most important threats AI presents to human safety, health, and basic rights.

Because of this, the EU AI Act will probably be used as a model for global regulatory measures.247 Regardless of such measures, manufacturers and suppliers of AI systems outside the EU will still need to comply with the requirements set out by the EU AI Act and cope with its extraterritorial impacts.248

The Act's geographic reach is broad, but its content is profound. The obligations outlined in the EU AI Act will probably stretch beyond the type of high-risk AI systems that the Act specifically regulates,

notwithstanding the Act's concentration on especially dangerous AI uses. For example, in the event of a true compliance pull, military systems—which are not covered by the EU AI Act because to competence249—will likely need to meet the requirements set out by the EU AI Act. Additionally, AI systems that aren't presently classified as high-risk might, mutatis mutandis, be vulnerable to at least

some of the duties outlined in the Act, mostly by way of judicial interpretation. This may be especially true for applications that might have a significant social effect.250 Because highly changeable AI systems may fall inside the purview of the EU AI Act's responsibilities after a simple modification of the use intention, they may also need to be developed with the obligations for high-risk AI in mind.

All things considered, the duties outlined in the EU AI Act are rather extensive. Programmers and software developers must reconsider how they operate in light of the new responsibilities associated with risk management. As a liberation from earlier battles to establish rules for human-machine interaction, this includes the need for human monitoring.

The art now rests in improving both article 14 and the EU AI Act in conformity with basic rights theory and understanding from practical research addressing the viability of human supervision, based on the critique expressed in this chapter. By doing this, it will be feasible to prevent human supervision as defined by the 14 EU AI Act from turning into just another oversight in a long string of unsuccessful efforts to address the intricate relationship between man and machine.

The Swiss National Science Foundation, National Research Program 77 "Digital Transformation," provided funding and support for the project "Meaningful Human Control of Security Systems – Aligning International Humanitarian Law with Human Psychology" (grant no. 407740_187494, co-PIs François Fleuret and Markus Christen); Juliane Beck is a PhD researcher working on the same project; Serhiy Kandul, Markus Kneer, and Vincent Micheli are other partners and researchers.