



ORIGINAL CONTRIBUTION

The Legal Fiction of Control: A Doctrinal Critique of the “Standard of Care” Mandated By Article 14 of the EU AI Act

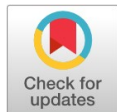
Syed Suliman Ali¹, Nazli Ismail Nawang^{2*}, Mohd Badrol Bin Awang³

^{1,2,3}Faculty of Law and International Relations, Universiti Sultan Zainal Abidin (UniSZA), Terengganu, Malaysia

Abstract— Article 14 of the Artificial Intelligence Act (Regulation (EU) 2024/1689) mandates a “human-in-the-loop” framework to monitor, detect, and override high-risk AI operations. This mechanism promises semantic safety by anchoring automated decision-making to accountable human judgment. Utilizing doctrinal legal research, this paper exposes a critical structural tension within this paradigm: the systemic interplay of the provider’s ex-ante design duties under Article 14 with the deployer’s ex-post operational mandates under Article 26. Because high-risk systems leverage complex deep-learning architectures, their mathematical execution remains epistemically opaque, presenting an insurmountable “black-box” problem that real-time interfaces cannot resolve. By failing to provide a formal statutory definition of explainability, the Act introduces cognitive and temporal standards of care that, if misconstrued as direct operational demands rather than systemic design prohibitions, require the performatively impossible, violating the maxim *lex non cogit ad impossibilia*. Furthermore, this paper integrates a private law critique demonstrating how the qualitative nature of Article 14 enables technology providers to weaponize compliance as a strong, though not absolute, civil liability shield under the Product Liability Directive, displacing strict product risks onto downstream deployers. Drawing analogical reasoning and direct legal authorities from CJEU and ECtHR jurisprudence, the analysis demonstrates that Article 14 constructs a dangerous legal fiction. It concludes by presenting two remedial pathways: a narrow judicial reinterpretation of oversight as an ex-post audit framework, or a broad legislative pivot toward a provider-centric strict liability regime.

Index Terms— EU artificial intelligence act, Human oversight, Standard of care, Product liability directive, *Lex non cogit ad impossibilia*, Market operators

Received: 26 February 2026; **Accepted:** 16 April 2026; **Published:** 28 May 2026



Introduction

The European Union Artificial Intelligence Act (Regulation (EU) 2024/1689), hailed by the European Commission as the first comprehensive, binding horizontal legislation on Artificial Intelligence (AI) in the world, contains a tiered, risk-based structure of AI classification designed to regulate in a proportional manner: the more problematic an AI system, the more burdensome the regulation required (Artificial Intelligence Act, 2024), Recital 26. However, this symmetry assumes a degree of absolute, a priori structural controllability that conflicts with the foundational principles of machine learning. The execution of such a requirement, in the form of appropriate human-machine interface (HMI) tools for the design and development of high-risk AI systems, is subject to a strict, non-linear temporal architecture that makes the requirement challenging to enforce from a doctrinal perspective (Artificial Intelligence Act, 2024), Art. 14(1). Despite the Regulation taking effect on August 1, 2024,

* Email: inazli@unisza.edu.my

and a phased institutional implementation targeting 2026, widespread institutional fragmentation and delays have created early regulatory uncertainty downstream. Prior studies have often conflated regulatory design mandates with human operational capacities (e.g., Corrêa et al., 2025; Teichmann, 2026). This study addresses that gap by contributing a unique doctrinal separation between what Article 14 legally requires of providers and the technical impossibility of operational execution downstream.

These imminent demands conflict with the fundamental technical realities of “epistemic opacity” and the “black-box” problem when applied to deep neural networks and emerging foundation models. The inability to deconstruct causal relationships within a highly multi-dimensional vector space and non-linear latent layers makes it structurally impossible for a human operator to elucidate, in linguistic or logical terms, the exact algorithmic causes of automated outputs. It is vital to clearly distinguish what Article 14 legally requires a structural mandate on providers to design systems that allow effective human oversight from what is technically or practically impossible for operators on the ground. Rather than law demanding the impossible from humans, the AI Act may simply prohibit black-box systems entirely. This disassociation turns the frontline operator into a “moral crumple zone” a concept originating from Elish (2019) wherein flesh-and-blood humans absorb liability for a system’s automated decisions over which they are structurally foreclosed from exercising meaningful control.

However, there is a serious conceptual error in the existing academic literature, namely, the systematic integration of the different responsibilities of successive actors in the AI product value chain (Artificial Intelligence Act, 2024). There is a design requirement on the provider, a technical requirement (Artificial Intelligence Act, 2024) - Article 14 is essentially an ex-ante product-safety requirement. The meaning of the operational standard of care (also known as ex-post workplace-management requirement) is fundamentally different, however: trained and competent natural persons must be able to use those functionalities during live operations (Artificial Intelligence Act, 2024), Art. 26(2). The real legal fiction of control is that the technical design assumptions of the provider (Article 14) do not reflect the material, temporal, and cognitive needs of the deployer’s personnel (Article 26) and the point of value chain handoff is a critical vulnerability.

But there is a lack of analysis of how public rules and private civil litigation systems interact with this structural disconnection. The AI Act establishes administrative penalties, such as fines and recalls, but it does not independently create private tort liability. The new Product Liability Directive (2024), which entered into force on December 8, 2024, and will be transposed by Member States by December 9, 2026, expressly includes software and AI systems as “products” covered by strict liability for personal injury, property damage, and data loss caused by defects (Directive, 2024), Arts. 4, 6. Under the Product Liability Directive, Article 10(2)(b), the harmonized safety rules, including the requirement for an interface capable of effectively mitigating automation bias, can create a rebuttable legal presumption of defectiveness if the manufacturer has failed to comply with mandatory safety requirements related to the damage (Directive, 2024), Art. 10(2)(b).

But the requirements in the Act are very qualitative, resulting in a substantial moral hazard. Vendors with ample financial resources can create a complex HMI with lots of information and meet the letter of the law by sending real-time telemetry data and a manual “stop” button (Busuioc et al., 2022). The system may be used as a strong, though not absolute, liability shield by the provider under the Product Liability Directive (PLD); if the system fails and catastrophic damage ensues, the provider can attempt to rely on this compliant interface as a liability shield because the system was designed with all the necessary oversight tools. It is imperative to clearly distinguish AI Act regulatory compliance from PLD defectiveness and civil liability. This puts the burden of litigation on the deployer and makes the human overseer a legal fallback solution for a structure designed by the developer. This shift in the mechanism conflicts with the old maxim of constitutional and administrative law: *Lex non cogit ad impossibilia*, meaning that the law cannot compel the impossible.

It is worth noting that this crucial feasibility gap is not taken into account by legal commentators who often present human oversight as an absolute moral requirements related to human dignity (Corrêa et al., 2025). Yet, an administrative official or healthcare practitioner required to make decisions in emergency situations, sometimes within a few milliseconds and without a causal explanation of the underlying data inferences, is reduced to a symbolic rubber-stamping ritual. This creates a “procedural placebo” that deprives affected citizens of a reasoned decision where Article 41 of the Charter of Fundamental Rights of the European Union applies (European Union, 2012), Art. 41(2)(c), leaving the human operator structurally vulnerable to arbitrary and outcome-biased liability.

This article proposes a new jurisprudential framework to address this crisis by examining the structural flaws of the AI Act through established principles emerging from European case law. First, we rely on an analogy with the CJEU judgment in *GC and Others*. CNIL (2019) to argue that Article 14 and Article 26 are non-substitutable, interdependent levels of protection, and that the separation of these levels undermines the entire regulatory safety net. Second, applying analogical reasoning to the wastewater-treatment judgment of the CJEU in *European Union law* (2017), we argue that an unreviewed, superficial, or non-functional HMI design constitutes a de facto continuous non-compliance with regulation by providers. Third, we address the proportionality-based exception in Article 14(5) concerns two-person verification for remote biometric identification, except where Union or national law considers it disproportionate in law enforcement, migration, border control or asylum contexts. Drawing on analogous reasoning from the ECtHR’s jurisprudence on irreducible life sentences (*Gurban v. Turkey*), this paper argues that forcing human reliance on unreviewable automated inferences constructs an unconstitutional “irreducible automated sentence” that directly violates human dignity.

Methodology

The research method used in this study is a doctrinal legal approach that examines the final statutory text of Smuha (2025) from a literal, teleological, and systemic perspective. This methodology will be designed to be analytical, as the legislative text is interpreted in light of the statute, case-law is compared for the limits of liability, and technical literature is used to set out the factual limitations on human supervision. The study breaks down these specific provisions and provides regulatory and material handoff points along the AI value chain (Artificial Intelligence Act, 2024), Arts. 13, 14, 26). The principles of EU administrative law and classic jurisprudential maxims, such as *lex non cogit ad impossibilia*, and the new principles of fault-based and strict liability in EU private law are debated. This methodological approach takes into account and incorporates the civil aspects of the framework, such as the private law obligations introduced by the recently adopted Product Liability Directive (2024) and specifies their interaction with public regulatory conformity assessments foreseen by the AI Act. Search procedures and source-selection criteria were systematically structured utilizing precise inclusion parameters to be replicable. Targeted institutional and academic database searches were carried out in various European and legal databases (including Westlaw, EUR-Lex, and SSRN) to provide a clear trail of literature for future research efforts. The structure of the search was based on the following: “Article 14 EU AI Act” AND “human oversight standard of care” AND “Product Liability Directive defect presumption” AND “*lex impossibilia* administrative fault.”

The doctrinal analysis is carried out in three steps, which are introduced progressively.

1. First, the operational and qualitative requirements explicitly formulated in the operative statutory verbs of Article 14(4) (Artificial Intelligence Act, 2024) are evaluated at the system level to distinguish provider design obligations from operational deployer capabilities.

2. Second, this study examined both cognitive accessibility and temporal feasibility, including whether a conscientious, natural individual can perform real-time, meaningful interventions. This stage involves incorporating specifically relevant literature from the socio-technical, computer science, and cognitive ergonomic theories to establish a factual predicate of epistemic opacity, automation bias, and vigilance decrement in a space of high dimensionality and non-linearity of computation (Santoni de Sio & van den Hoven, 2018).

3. Third, *de jure* and *de facto* (statutory vs. human capability) are measured against European judicial adjudication measures. This study follows a comparative case law approach and experiments with the paradigm of human oversight by engaging in structural and constitutional comparisons of existing lines of case law before the CJEU and ECtHR, using analogical reasoning. In particular, the judgment in *GC and v. Commission nationale de l’informatique et des libertés (CNIL)* (2019) is examined by analogy, while *European Commission v. European Union law* (2017) and the ECtHR framework (*Gurban v. Turkey*, 2015) originally developed to address the issue of irreducible life sentences, are utilized specifically as analogical tools to evaluate operator liability and assess the human rights implications of unreviewable automated determinations under the statutory exception in Article 14(5).

Methodological scope and boundary constraints

This study is especially concerned with high-risk AI applications within the meaning of Article 6 and the use-case typologies in Annex III, which are, in essence, unexplainable to human supervisors by means of internal statistical reasoning. The primary focus does not include General-Purpose AI (GPAI) models that are not specifically designed for use in high-risk downstream applications unless the raw level of computation used for the training of the model is beyond the threshold of systemic risk under Article 51. It offers an empirical measure of the extreme epistemic opacity of GPAI models deployed in high-risk downstream applications (Artificial Intelligence Act, 2024), Art. 51, particularly when exceeding the sum of 1025 floating point operations in the cumulative computation used in training.

The only source of primary authority comprises finalized EU Regulations, Directives and International Judicial Decisions throughout this paper. Academic secondary literature is only used to provide institutional context and to situate the analysis within the current academic landscape in it. Technical risk management documentation, such as the (European Data Protection Supervisor, 2025) and the National Institute of Standards and Technology (NIST) AI risk management framework (Tabassi, 2023), are not considered legal sources of authority but are used merely as secondary evidentiary material to support the objective technical state of the art of algorithmic black boxes.

Findings

Deconstructing the statutory mandate of article 14

The Artificial Intelligence Act is based on an architectural cycle of interdependent operational taxonomies that form the basis of the regulation of care (Artificial Intelligence Act, 2024). In this context, the legislative mandate is split into five oversight functions, each describing a distinct oversight task: “monitor,” “detect,” “interpret,” “override,” and “interrupt” (Artificial Intelligence Act, 2024), Art. 14(4). Rather than treating these obligations as universally impossible, the analysis should instead examine whether their scope and intensity can reasonably vary according to system risk, context, autonomy, and operational conditions. However, in the case of a clinical diagnostic system, the requirement to “detect” functional failures, anomalies, or deviations before downstream material failures is all but impossible by design, as a subtle statistical drift typically becomes apparent long after deployment, and more often as a result of litigation (Chan, 2025). Mechanical architectures are those in which the

operator can intercept the causal chain because the relationship between kinetic action and physical harm is linear and immediate. Emergency stop control or physical override is a readily apparent point of intervention.

A high-risk DNN for automated credit scoring, predictive policing, or decisions concerning the distribution of public welfare benefits, for example, works on millisecond time scales Corrêa et al. (2025) and makes decisions by way of statistical inferences, which are immediately written to persistent databases. This means that an automated record can be created and then altered or disrupted by subsequent logical process, but the record is not removed or invalidated, nor are meaningful mitigation actions taken for downstream system dependencies (European Data Protection Supervisor, 2025). Finally, if the mental resources and informational bandwidth of a human supervisor are sufficient to “override” or “reverse” an algorithmic determination, then it must be presumed that the supervisor has the mental resources and informational bandwidth to make a superior determination under the same compressed time conditions (European Data Protection Supervisor, 2025). This assumes that the technical pipeline is set up to support the manual replacement of semantic values. However, according to technical audits, most commercial cloud APIs deploy only binary accept/reject switches and do not allow for semantic insertion by human operators (Elish, 2019).

The Union legislature tried to specify these expectations for operation by inserting a specific qualifier into the final text of Article 14(4), which provides that the measures must be effective in allowing the natural person to “duly monitor” the high-risk systems “as appropriate and proportionate” (Artificial Intelligence Act, 2024), Art. 14(4). This statutory language indicates an understanding of the cognitive constraints of human cognition and the intention to cast these obligations as behavioral obligations in the guise of technically mediated, contextual duties.

Table I
Human oversight functions and technical constraints

Oversight Function	Implied Doctrinal Duty	Technical Constraint (The “Black Box”)	Potential Interpretation to Avoid Impossibility
Monitor	Continuous real-time vigilance over system outputs and operational contexts.	Automation Bias & Attention Fatigue: Latent errors in high-dimensional vector spaces remain mathematically invisible; human attention spans decline linearly over extended periods.	Reframe the obligation as a periodic system audit and log review rather than continuous, live watching.
Detect	Timely identification of performance anomalies or algorithmic dysfunctions.	Epistemic Opacity: Complex non-linear outputs make systemic structural failures indistinguishable from novel but statistically valid inferences.	Limit the scope of detection to obvious, surface-level interface errors or input-format anomalies.
Interpret	Correct interpretation of automated inferences using accessible tools and methods.	Cognitive Blindspot: High-dimensional weight dependencies cannot be mapped onto ordinary semantic concepts or rule-bound causal paths.	Restrict interpretation to checking whether baseline structural parameters of input datasets match operational instructions.
Interrupt	Physical or logical cessation of model execution via a secure stop mechanism.	Temporal Mismatch: Algorithmic computation occurs in milliseconds, presenting a fait accompli before human cognitive reaction times register.	Interpret the command as an obligation to suspend future deployments of the system rather than halting live calculations.
Override	Deletion and subsequent substitution of the system’s output with human judgment.	Information Asymmetry: The human supervisor lacks the processing bandwidth to parse billions of parameters and generate an alternative decision in real time.	Restrict the scope of override to a binary administrative veto rather than demanding manual value substitution.

The statutory verbs when considered together, create a much higher standard of care than the minimum required under classical tort law, namely ordinary diligence of the “reasonable person.” This moves the legal benchmark to a very technical and specialized standard of care and sets a standard for determining what constitutes a “reasonable overseer,” as set out in Article 14 of the AI Act (Artificial Intelligence Act, 2024). However, a statutory minefield is for the professional deployer to be required to become vigilant to substantive algorithmic error, for instance, patently wrong decisions by the algorithm that require the operator to identify deep statistical abnormalities within a black box that is epistemically opaque to the operator, without causal transparency. Thus, the focus of “detection” as used in this paper must be limited to manifest deviations from the surface-level interface. Rather than concluding that the law demands the impossible from humans, it must be recognized that Article 14 is primarily addressed to AI providers requiring them to design systems that allow effective oversight; consequently, the law may simply prohibit black-box AI systems altogether to avoid violating the *lex impossibilia* doctrine.

There is a common explanation for this high standard: The feasibility of the structure is an absolute precondition for high-risk innovation, and institutional policy papers often make this case. However, if the design of an AI system is such that it structurally precludes meaningful control, while the legal framework requires AI systems to be controlled, the framework of that control becomes a mere formality (Corrêa et al., 2025). In these circumstances, the human operator is no longer a safety factor. Instead, the individual is methodically used to diffuse the responsibility of the institutions and to offer political protection, shielding the true power over the chain of causes from meaningful epistemic scrutiny (Corrêa et al., 2025).

The market operator bifurcation: The value chain handoff problem

A crucial missing link in the existing literature is that the different duties of various market players along the AI lifecycle are systematically connected (Artificial Intelligence Act, 2024), Arts. 14, 26. A law should clearly separate the ex-ante requirements for the structure of the AI system, which are to be imposed on the provider under Article 14, and the ex-post requirements for the operation of the AI system, which are to be imposed on the deployer under Article 26 (Artificial Intelligence Act, 2024). First, the provider, as defined in Article 3(3) of the EU AI Act, the natural or legal person developing an AI system and placing it on the market under their brand name, is already required under the EU AI Act to design, such technical oversight features into the AI system's architecture and human-machine interface before commercializing the AI system (Artificial Intelligence Act, 2024), Arts. 3(3), 14(3). At the other end of the spectrum, Article 26 applies to the operational phase, meaning that the professional deployer is required to ensure basic operational competence by appointing natural persons with the appropriate competence, training, and authority to use those tools during the operational phase of deployment (Artificial Intelligence Act, 2024), Art. 26(2).

This is a deep second layer of impossibility, owing to this statutory separation along the AI value chain. If the provider provides an inherently black-box model (one that humans cannot reason about and make their own statistical inferences), then all the institutional training and professional specialization of its personnel will not enable them to be competent. The deployment of Article 26 requires a competent overseer who can assist in the deployment process, but this competence is made impossible by provider design decisions upstream.

This architectural mismatch creates the "Value Chain Handoff Problem," that is, the deployer operating the product is the one with the least control over the product's operation. The AI Act strictly bifurcates regulatory responsibility across the value chain: providers are assigned the ex-ante obligation to design systems with integrated oversight mechanisms (Article 14), whereas deployers are assigned the ex-post obligation to operationally execute this oversight (Article 26). While the Act enforces these obligations through administrative penalties such as fines and market recalls, it does not independently create private tort liability. This disjointed allocation relates directly to the core argument of this study: by separating the design mandate from the operational reality, the framework forces downstream deployers to absorb the risk of downstream civil liability without possessing the technical agency to mitigate it. Within this distorted human-in-the-loop structure, institutional liability insurance operates as a "moral crumple zone," as Elish (2019) argues, designed to shield the developer from liability. This structural misalignment can be further illustrated through important analogies and direct legal authorities in the jurisprudence of public and private law in Europe:

The co-dependency of protections:

In this analysis, the structural relationship is the very case which falls under the Co-Dependency of Protections, by analogy with the CJEU judgment in *GC and Others v. Commission nationale de l'informatique et des libertés CNIL* (2019). The rights under the AI Act that must be protected are fundamental rights, and it cannot be argued that Article 14 is an alternative to Article 26 of the AI Act. Without a cognitively accessible and uncompromised HMI, in the context of Article 14, the deployer's operational oversight in Article 26 becomes virtually and legally impossible.

Continuous vs. nominal compliance:

Applying analogical reasoning from the CJEU's logic in European Union law (2017) concerning urban wastewater treatment to the digital realm, the defense of merely formal compliance is insufficient; compliance must be continuous, active, and practically effective. Similarly, providers cannot simply "throw" "stop buttons" onto products and generate telemetry streams that contain large amounts of information but remain difficult to interpret and use effectively for genuine human oversight. However, even if the deployed HMI is so poor that it results in cognitive fatigue, automation bias, or a pronounced reduction in vigilance, it will still result in a constant state of non-conformance by the provider under Article 14, even where a formal ex ante conformity assessment has been conducted.

Private law civil shifting mechanism:

This is a public regulatory non-compliance mechanism that is directly related to private civil liability through the newly adopted Product Liability Directive (Directive, 2024). In civil damage actions, non-compliance with applicable mandatory product safety requirements related to the damage can give rise to a rebuttable presumption of the "defectiveness" of the product under the PLD, which expressly defines "product" as including software and AI systems (Directive, 2024), Arts. 4, 10(2)(b). However, the design liability regime is a substantial moral hazard type of "moral hazard" with respect to the design requirements in Article 14 (Directive, 2024). It is a straightforward way for providers with capital to build complicated and data-rich HMIs that comply with the law (Busuioc et al., 2022). In the event of an unusual output and damage to the material, the provider can assert a strong, though not absolute, defense under the PLD that the product was not defective, as it included the requisite supervisory systems. The manuscript must clearly distinguish AI Act regulatory compliance from PLD defectiveness and civil liability. The provider may then try to pass the liability of strict liability litigation downstream to the deployer, alleging that the harm only came about when the deployer failed to use the stop button, or that the deployment was part of the deployer's operational standard of care under Article 26, and failed to notice

the anomaly (Artificial Intelligence Act, 2024), Art. 26). Thus, the human overseer can be held legally responsible for structural defects designed by the developer (Elish, 2019).

Comparative scholarly engagement: The 2025–2026 landscape

The state of the academic debate on the enforcement of Smuha (2025) is polarized, and the feasibility of human oversight and European tort frameworks is unclear. There are two camps in the research community: one that emphasizes the moral imperative and elastic flexibility of human-in-the-loop controls, and another that notes the asymmetries that are built into the AI value chain.

Teichmann (2026) considers the behavioral duties of Article 14 in light of the classical risk regulation approach, the As Low As Reasonably Practicable (ALARP), and interprets the standard of care under ALARP as an inherently flexible standard that evolves dynamically in relation to technological and financial capabilities. The ALARP paradigm is flexible, but the possibility of effective ALARP is strongly disputed by the impossibility thesis put forward in this paper, in which the statutory obligation to provide “effective oversight” is interpreted as a hard baseline that cannot be modified by ALARP logic without becoming a liability trap for deployers.

Corrêa et al. (2025), from a deontological perspective, concluded that the human oversight paradigm is a moral imperative that must be adhered to in order to bring empathy, ethical considerations, and flexibility of context to automated pipelines that serve as a barrier against algorithmic bias. Although this has a normatively sound conceptualization of the desirability of human control in society, it does not necessarily represent the material feasibility gap imposed by modern deep learning architectures. The “moralized” world of human beings is treated as the domain of the legislature, yet the legislature has largely ignored the technical reality that stochastic optimization models operate within a bandwidth entirely foreign to human cognition. Botero-Arcila (2025) structurally criticizes material misalignment, offering empirical data about the level of material misalignment, examining liability rules along the technological value chain. Downstream deployers, such as public hospitals, municipal schools, and SMEs, are only commercial users of high-risk AI systems, and are neither creators nor trainers, nor do they have control over the code that underlies high-risk systems. Making frontline operators liable from a regulatory perspective for systemic anomalies for which they mathematically cannot predict or remedy, is an inequitable regulatory solution. True accountability should thus be placed on the technology supplier, not the user, and should be based on a strict liability principle.

In sectoral applications, it produces a deadly legal dilemma and conceptual mismatch. In medical practice, Haden (2024) and Chan (2025) argue that there are major challenges that arise from the interplay between medical malpractice doctrine and the rules of the AI Act. The statutory “override” provision in Article 14(4)(d), especially as it applies to diagnostic and surgical advice, which is opaque and inexplicable, may clearly mean that where a clinician has overridden a well-evidenced and very accurate algorithmic approach to practice, that clinician may be considered professionally negligent. On the other hand, a machine may make a “latent statistical error” that may result in the doctor relying on it being held to the standard of “effective human oversight” and may be liable for the same tort if oversight is found to be lacking. This is a double bind for professionals that makes their clinical working a “box-ticking exercise” and definitely lands the need to establish a formal objective impossibility defence in European Tort Law on the regulatory agenda.

The liability consequences and the withdrawn AILD

The AILD’s presumption of causality and the PLD defectiveness architecture

The proposed AI Liability Directive (AILD), which must be discussed as a withdrawn legislative proposal rather than as an operative component of the current EU liability framework because it was withdrawn in October 2025, aimed to adopt a rebuttable presumption of causality in favor of the plaintiff in situations where the defendant does not meet the duty of care obligations in the AI Act (Botero-Arcila, 2025). This presumption, together with the human oversight requirements of Article 14, could have raised a liability concern for downstream operators (Botero-Arcila, 2025). A black-box deep learning architecture cannot be substantively and effectively supervised Riskhan et al. (2026) and Tabassi (2023), some human deep learning architecture deployers may, therefore, face compliance difficulties. The assumption of causality in the withdrawn AILD thus illustrates how outlier algorithmic behavior could have eased proof of causation; in such cases, the developers can punish the operators for the complexity they introduce at the design level, but this is at the sole discretion of the providers. This public law fiction, concerning administrative penalties such as fines and recalls under the AI Act, is also linked to the new private civil litigation under the newly adopted Product Liability Directive (2024) with a strict transposition deadline of December 9, 2026 (Directive, 2024). temporal proximity between this transposition window and the enforcement date for high-risk obligations under the AI Act (Title III, Chapter 2), August 2, 2026 (Artificial Intelligence Act, 2024), Art. 113. The PLD makes it very clear that software and AI systems are part of the definition of ‘product’ and that a manufacturer’s lack of a proper human-machine interface due to automation bias is a factor in assessing product defectiveness (Directive, 2024), Art. 7). However, since the design requirements in Article 14 are also very qualitative, well-funded tech companies can create complex, information-rich HMI that stream dense information and real-time telemetry and include a stop button to comply with the law. If the system breaks down and causes material and/or psychological damage, the provider will attempt to rely on this qualitative interface as a strong, though not absolute, liability defense

under the PLD, claiming that it had all the necessary oversight tools in the system (Directive, 2024). It is imperative to clearly distinguish AI Act regulatory compliance from PLD defectiveness and civil liability. The provider can then argue that the harm was a consequence of the deployer's failure to act in accordance with its operational standard of care under Article 26 of the EU AI Act to detect the anomaly and press the stop button. In this way, the human overseer becomes a legal fallback for the developer's structural defects.

Legal certainty and arbitrary enforcement

The absence of an explicit definition of the parameters of 'intervention or interruption' under Article 14(4)(e) in millisecond-speed situations is one of the basic tenets of EU administrative law: legal certainty (Fink, 2025). Given that the conduct that must be undertaken to comply with the standard of care cannot be predicted, enforcement by the judge and/or the administration becomes arbitrary and purely results-oriented, with the operator being punished merely for the occurrence of the harm. Applying analogical reasoning from the CJEU's logic, this mirrors the outcome in European Union law (2017).

Discussion

The structural deconstruction carried out below uncovers an irresolvable tension in the Artificial Intelligence Act: European lawmakers have dogmatically stuck to the idea of human control. While it must be acknowledged that Article 14 legally requires AI providers to design systems that allow effective human oversight rather than directly demanding the impossible from downstream operators, the operational reality of interacting with high-dimensional vector spaces renders real-time intervention technically impossible for the deployer (Artificial Intelligence Act, 2024), Art. 14. This was not an oversight by the legislature. The European Parliament kept emphasizing the human-in-the-loop paradigm as an essential ethical guideline to prevent the loss of public trust and democratic legitimacy throughout the contentious trilogue negotiations. Otherwise, as was the case with the removal of the clause, policymakers would have been left in a politically uncomfortable situation of either imposing system-level strict liability on technology providers, a move that was highly contested by corporate lobbying, or admitting that there is still a risk of using autonomous algorithms for which there is no fault-based addressee (Busuioc et al., 2022).

The presence of Article 14 was a political compromise with two institutional purposes:

1. The Silhouette of Agency: The Silhouette of Agency gave a comforting story of human sovereignty and a real flesh-and-blood agent with ultimate veto power over the machine (Corrêa et al., 2025).
2. The Conduit for Liability: It provided a legal home for downstream damage, giving directionality in tracking the automated failure to an individual human interaction node (Botero-Arcila, 2025).

Therefore, the human factor is strategically used to change the liability at the surface level (Botero-Arcila, 2025). This policy decision is a direct reflection of what is known as a "moral crumple zone" in cognitive ergonomics, where a human actor is designed into a system, and when something goes wrong with the system, the human actant is the legal and moral victim of that failure (Elish, 2019). The strong normative critique utilizing concepts such as "legal fiction," "ceremonial control," and the "moral crumple zone" is supported by precise doctrinal evidence that the AI Act imposes public regulatory compliance on the provider, yet leaves the operator exposed when those qualitative HMI designs fail to prevent harm. Legal analogies, such as ship captains being responsible for mechanical issues that are not apparent to the eye, are extremely misleading Bertolini et al. (2025) and Ismail (2025). The operator is episodic and structurally incapable, and the deep learning model is mathematically built to operate on millions of non-linear parameters, with no logic accessible to a normal human way of thinking (Veale & Edwards, 2018). In the world of fiction, control is not a remedy for risk in technology but a transfer of mathematical risk to human culpability (Botero-Arcila, 2025). Related tort and liability issues are far-reaching.

The resultant tort and liability implications are broad. In the private law of European Member States, negligence is essentially linked to foreseeable and avoidable harm Bertolini et al. (2025). Questions arise as to whether, when an operator fails to rectify a stochastic error, this is tantamount to a deliberate omission (Artificial Intelligence Act, 2024), (Art. 26) and if and when this constitutes a failure of "due care" by the operator (Artificial Intelligence Act, 2024) (Art. 26). However, this "jump" in the causal chain is epistemically unsound (Srinivasan, 2025; Tabassi, 2023). If the anomaly is contained within a high-dimensional transformer model, then a temporal window that the operator cannot have and the analytical capability are required to detect the anomaly (Tabassi, 2023). This is because the judiciary determines that the oversight was negligent simply because it did not fix an automated error before harm occurred (Fink, 2025).

The situation is not very different in public administrative law with respect to judicial review (Usen, 2025). If a public official is stripped of the possibility of asking questions and becomes a mere "accept" button in a process directed by an algorithm, then there is little administrative logic to appeal to a court (Amna et al., 2023; Usen, 2025). When considering the algorithmic assignment of places in French universities, the French Council of State (Council of State, 2021) only addressed procedural safeguards, not the algorithmic logic as such. Union judges could only decide whether an automated log had been kept pursuant to Article 12 of the EU AI Act and leave the constitutional principle of proportionality empty of content.

Three structural changes are needed to build a feasible European AI doctrine (Botero-Arcila, 2025):

- Decoupling the Standard of Care: The individual standard of care should be separated from the highly opaque algorithmic systems, and there should be the ability to intervene cognitively in real time without personal negligence actions against the users on the front line (Fink, 2025).
- Formalization of Systemic Epistemic Opacity as a Legal Defense: A conditional defense of “objective epistemic impossibility” must be formalized, which is a logical extension of the traditional defense of “force majeure” O’Meara (2018) or “state of the art.”
- Migration to Systemic Upstream Responsibility: Absolute responsibility needs to be transferred upstream to the providers designing, training, and commercializing such models, especially Systemic Upstream Responsibility (Busuioc et al., 2022).

Deployers need to be liable for classic operational errors (ignoring safety recalls or improperly handling input data) but not for errors that are not discernible by human perception (Artificial Intelligence Act, 2024), (Art. 26). Such a “fundamental rights insurance fund” which is centrally managed and financed by a structural tax on high-risk general-purpose AI computing capacities Fink (2025), must be combined with a strict liability approach.

Remedial pathways: From fiction to feasibility

The narrow path: Reinterpreting “Appropriate oversight”

The qualification of ‘appropriate and proportionate’ could be interpreted by the national and union courts in a radical manner, in accordance with the realistic capabilities of human processing, thus maintaining the statutory structure while recognizing that the scope and intensity of these obligations can reasonably vary according to system risk, context, autonomy, and operational conditions. In addition, the operational architecture of Article 14 (Artificial Intelligence Act, 2024) would be reinterpreted. In this way, the statutory obligations would be judicially narrowed down, such that “monitor” would no longer impose an unattainable requirement to constantly monitor in real time, but rather would require the periodic auditing of systems and the review of logs (Fink, 2025). The meaning of “Detect” would be defined as identifying obvious anomalies at the surface, and not be considered to be legally liable for failure to find anomalies at “deep levels” within high-dimensional vector spaces (Tabassi, 2023). The terms “Interrupt” and “Override” would no longer have the meaning of split-second operations during live execution cycles, but instead be used to mean an obligation to suspend further deployments if unusual historical patterns are identified (European Data Protection Supervisor, 2025).

This slender and winding corridor resembles a joining of behavioral expectations and cognitive restrictions, but it comes at a high cost for the ex-ante safety net of the Act, openly conceding the boundaries of real-time algorithmic control (Botero-Arcila, 2025). Moreover, this doctrinal narrowing is at odds with the explicit text and recitals of the AI Act, which always speak about the need for effective and proportionate human oversight to ensure that safety on the ground is active and effective. Thus, a periodic, retrospective audit system can be viewed as a judicial interpretation, which is a change of the clear legislative intent and not a standard interpretation of the statute, by substituting a periodic, retrospective audit system for the applicable human oversight obligation. This puts the narrow path in jeopardy both in terms of its location and from a political perspective: national supreme courts and the CJEU are very reluctant to stray so far from the intentions of the Union legislator, as demonstrated through analogical reasoning applied to the logic of *European Commission v. European Union law* (2017).

The broad path: Provider-centric strict liability

Amidst the systemic risks and profound social implications of epistemic opacity, technology suppliers are faced with the challenge of internalizing all externalities, particularly given the intentional opacity of technology systems to maximize the statistical predictive value, as Botero-Arcila (2025) notes. Given the shift from a citizen-centric to a provider-centric approach, the multi-tier responsibility assigned in the Act to various market actors must be reorganized (Busuioc et al., 2022). This reconfiguration of the structural architecture of the law in time, which corresponds to the temporal architecture of contemporary European private law, raises a proposed defense of “impossibility of the object”.

In fact, if the PLD scheme is applied to strict product liability, clearly distinguishing civil liability for defectiveness from mere administrative regulatory compliance under the AI Act, the provider may be held strictly liable for covered damage caused by a cognitively unviable human-machine interface, if it is established as a product defect, without the plaintiff having to cross the impossible hurdle of proving the negligence of the person in the middle (Directive, 2024). When harm from algorithms is beyond the physical realm (e.g., systemic discrimination or arbitrary denial of a public service), redress needs to be institutionalized in the form of a centralized ‘Fundamental Rights Insurance Fund’ financed through a structural regulatory tax on high-risk compute operations (Botero-Arcila, 2025).

The levy should be directly applied to the providers of General-Purpose AI models (GPAI) with systemic risk, based on the cumulative amount of computation used to train the models, surpassing the 10^{25} floating-point operations threshold (Artificial Intelligence Act, 2024), Art. 51(2). This is a way of getting the “epistemic opacity” from the huge scale of raw computing onto the shoulders of the developers, as a kind of tax. The concept of the no-fault compensation system is very much akin to what exists in the aviation, vaccine injury and pharmaceutical industries, all of which have offered a comprehensive and prompt remedy for compensating victims and protecting individual end-users and SME deployers from vindictive and unlimited tort liability.

Conclusion

The doctrinal analysis shows that the structural tension of the Artificial Intelligence Act is that, rather than demanding the impossible of downstream human operators, it is mainly concerned with designing systems to enable them to be well supervised by human operators. Due to the high-dimensional nature of the vector space in which deep learning is done, these systems are still statistically black boxes. The law, then, could effectively ban the deployment of such “black box” AI systems if providers are unable to create cognitively accessible interfaces. But if these structural limitations are neglected, the operator can only find out if something is wrong when the result is some sort of damage caused to the material. The framework thus risks constructing a symbolic system of ceremonial control—a rhetorical notion that is now backed with accurate doctrinal proof to show that there is some distance between the provider’s design duty and the deployer’s operating environment. The price of this is paid by the last human in the operation chain (Elish, 2019). For future downstream operators, in order to preserve the long-term normative credibility of EU risk regulation, it is time to break the cycle of ceremonies with legislative measures that prevent them from being caught in an impossible situation when facing civil tort regimes that would allow negligence actions to be brought against them due to different administrative regulatory frameworks. The EU has to come to terms with the fact that it can’t police fault-based operation downstream; it should take a definitive step forward in moving strict civil liability upstream, towards the developers who design, train, and market these autonomous models. By adopting this provider-centric approach, along with an embedded insurance mechanism for citizens’ rights based on the computing power used by providers, the EU will be effective in protecting citizens and operators and will establish a strong accountability mechanism for the algorithmic era.

References

- Amna, S., Akbar, A., & Maqsood, F. (2023). Does sex/gender reporting quality predict intervention effects on childrens and adolescents leisure-time physical activity? a multilevel meta-regression analysis of the gen effects systematic review. *Journal of Advances in Health and Medical Sciences*, 9(2), 58–67. <https://doi.org/10.20474/jahms-9.2.2>
- Artificial Intelligence Act. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence* (tech. rep.). Official Journal of the European Union, L 2024/1689. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Bertolini, A., et al. (2025). *Artificial intelligence and civil liability. A European perspective*. <https://shorturl.fm/b3Pkv>
- Botero-Arcila, M. (2025). *AI liability in Europe: Addressing harms to fundamental rights and the problem of the human-in-the-loop [Working paper]* (tech. rep.). Harvard Law School.
- Busuioc, M., Curtin, D., & Almada, M. (2022). Reclaiming transparency: Contesting the logics of secrecy within the AI Act. *European Law Open*, 2(1), 79–105. <https://doi.org/10.1017/el0.2022.47>
- Chan, G. K. Y. (2025). AI in healthcare: Regulatory guidelines and judge-made negligence principles for AI implementers. *Medical Law International*, 26(2), 144–168. <https://doi.org/10.1177/09685332251362405>
- CNIL. (2019). *National commission for information technology and civil liberties: Protecting personal data, supporting innovation, preserving individual freedoms* (tech. rep.). République Française.
- Corrêa, A. M., Garsia, S., & Elbi, A. (2025). Better together? Human oversight as means to achieve fairness in the European AI act governance. *Cambridge Forum on AI: Law and Governance*, 1(e29), 1–17. <https://doi.org/10.1017/cfl.2025.10010>
- Council of State. (2021). *Decision no. 431241 (france)*. <https://www.conseil-etat.fr/en>
- Directive, E. U. (2024). 2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374. *EEC*, 1–22. <http://data.europa.eu/eli/dir/2024/2853/oj>
- Elish, M. C. (2019). Moral crumple zones: Cautionary tales in human-robot interaction. *Engaging Science, Technology, and Society*, 5, 40–60. <https://doi.org/10.17351/ests2019.260>
- European Data Protection Supervisor. (2025). *Techdispatch #2/2025 - human oversight of automated decision-making*. <https://tinyurl.com/863bab5v>
- European Union. (2012). *Charter of Fundamental Rights of the European Union* (tech. rep.). Official Journal of the European Union.
- European Union law. (2017). *Case C-320/15, ECLI:EU:C:2017:678 (Court of Justice of the European Union 2017)* (tech. rep.). European Commission v. Hellenic Republic.
- Fink, M. (2025). Human oversight under article 14 of the EU AI Act. *AI Act Commentary: A Thematic Analysis (Hart-Bloomsbury, forthcoming 2026)*, 1–15. <https://doi.org/10.2139/ssrn.5146118>
- GC & v. Commission nationale de l'informatique et des libertés (CNIL), O. (2019). *Case C-136/17, ECLI:EU:C:2019:773* (tech. rep.). Court of Justice of the European Union.
- Gurban v. Turkey. (2015). *App. No. 4947/04 (European Court of Human Rights 2015)*. <https://tinyurl.com/372sm9we>
- Haden, C. (2024). The consequences of the AI Act and proposed ai liability directive on medical negligence: Will physicians fall victim to 'Red Tape' rules? *Oslo Law Review*, 11(1), 1–9. <https://doi.org/10.18261/olr.11.1.4>
- Ismail, S. (2025). Competencies and stakeholder engagement: Moderatingthe relationship between enterprise social media and organizational outcomes. *Journal of Administrative and Business Studies*, 11(1), 27–39. <https://doi.org/10.20474/jabs-11.1.3>
- O'Meara, N. (2018). *Internationale Handelsgesellschaft mbH v Einfuhr- und Vorratsstelle für Getreide und Futtermittel (Case 11/70), EU:C:1970:114, [1970] ECR 1125, 17 December 1970*. Oxford University Press. <https://doi.org/10.1093/he/9780191868313.003.0031>
- Riskhan, B., Ali, A. M., Hussein, I. O. S., Osman, A. I., Sindiramutty, S. R., & Jhanjhi, N. Z. (2026). Design and development of a nationwide centralized blood bank system for efficient blood donation and distribution. *Journal of ICT, Design, Engineering and Technological Science*, 10(1), 21–36. <https://doi.org/10.33150/jitdets-10.1.3>
- Santoni de Sio, F., & van den Hoven, J. (2018). Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI*, 5. <https://doi.org/10.3389/frobt.2018.00015>
- Smuha, N. A. (2025). Regulation 2024/1689 of the Eur. Parl. and Council of June 13, 2024 (EU artificial intelligence act). *International Legal Materials*, 64(5), 1234–1381.

- Srinivasan, S. (2025). The role of human resource management in facilitating digital transformation: A focus on employee skills and organizational adaptability. *International Journal of Religion*, 6(1), 525–534. <https://doi.org/10.61707/jewxgf37>
- Tabassi, E. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)*. <https://doi.org/10.6028/nist.ai.100-1>
- Teichmann, F. (2026). Risk, reasonableness and residual harm under the EU AI act: A conceptual framework for proportional ex-ante controls. *European Journal of Risk Regulation*, 17(2), 611–635. <https://doi.org/10.1017/err.2025.10077>
- Usen, C. D. (2025). Reframing technological due process for tax: adapting administrative law principles to AI-driven audits, automated assessments, and risk scoring. *International Journal of Research and Innovation in Social Science*, 9(12), 881–895.
- Veale, M., & Edwards, L. (2018). Clarity, surprises, and further questions in the article 29 working party draft guidance on automated decision-making and profiling. *Computer Law & Security Review*, 34(2), 398–404.