Reconstructing Liability in Distributed Architectures: The Revised Product Liability Directive vs. Federated Learning

Wiktor Wilkołaski[0009-0003-7810-9073]

Independent Researcher
wiktor042004@gmail.com

Abstract. The Revised Product Liability Directive engenders structural conflicts when applied to distributed data architectures. While software is now explicitly classified as an actionable product under the Revised PLD, Article 10(4) evidentiary presumptions face significant challenges when applied to privacy-preserving computational frameworks. This research isolates Federated Learning (FL) within high-risk credit evaluation systems to demonstrate how decentralized training environments vitiate the Directive's causal logic. In FL architectures, central providers (manufacturers under the Revised PLD) distribute a global model while local nodes (operated by deployers under the AI Act) compute localized weights, thereby depriving the provider of access to post-deployment training data. Consequently, the stringent evidentiary disclosure mandates of Article 9(1) place providers in a highly disadvantageous legal position. To resolve this regulatory impasse between ex-post liability frameworks and ex-ante privacy regulations, this paper proposes a Differentiated Liability Taxonomy *de lege ferenda*. This tripartite framework directly calibrates Article 10 presumptions of defectiveness and causality against quantifiable vectors of architectural control. Implementing this taxonomy as a legislative update or regulatory guideline secures exacting remedial mechanisms for algorithmic harm while preventing the disproportionate penalization of decentralized computational topologies.

Keywords: Directive (EU) 2024/2853 · AI Act · Federated Learning · Trusted Execution Environments (TEE) · Privacy-Enhancing Technologies (PETs) · Evidentiary Presumptions.

1 Introduction: The Systemic Impasse in Algorithmic Accountability

Artificial Intelligence (AI) governance frameworks suffer from an inherent regulatory dilemma. Transitioning from the European Union's ethical rules to tort liability reveals an inherent tension between ex-post liability and privacy-preserving computational design [17]. Specifically, the Revised Product Liability Directive

(Revised PLD) [14], i.e., Directive (EU) 2024/2853, represents this doctrinal transition through the reclassification of AI systems as products under a strict liability regime. Read alongside the Artificial Intelligence Act (AI Act), Regulation (EU) 2024/1689 [15], this creates two challenges for practitioners—algorithmic accountability and strict limitation of data usage, especially in high-risk systems such as automated credit scoring [16,10].

While centralized architectures comply with both regulations, distributed ones challenge traditional doctrinal approaches. By way of illustration, Federated Learning (FL) [21,23] was developed to implement the General Data Protection Regulation (GDPR) requirement for data protection by design and by default. The approach achieves this by uncoupling algorithm training from the underlying data processing. Edge nodes perform localized calculations of the gradient and send the resulting aggregated weights to the orchestrator without sending proprietary, raw financial data [37].

The evidentiary mechanism of the Revised PLD—specifically Articles 9 and 10—is predicated on the assumption that the manufacturer has continuous telemetric visibility of their post-market product [30]. Because FL deliberately shields the central orchestrator from local node input data, a structural evidentiary asymmetry emerges. According to Article 9, a failure by a manufacturer to produce relevant data creates rebuttable presumptions of defectiveness and causation under Article 10. This creates a doctrinal paradox where central orchestrators are strictly liable for localized issues such as edge-node data poisoning and bias which the orchestrator lacks architectural and cryptographic capacity to detect and manage.

In February 2025, the European Commission withdrew the Artificial Intelligence Liability Directive (AILD) [4], thereby exacerbating this regulatory gap. Consequently, the Revised PLD constitutes the sole harmonized instrument addressing algorithmic tort liability within the Union [13,28].

To address this schism between the realities of the distributed computational architecture and strict liability, this article adopts a methodology of legal doctrine combined with cryptographic threat modeling. Therefore, a Differentiated Liability Taxonomy is proposed *de lege ferenda* [42,7]. Contemporary European tort law relies extensively on binary demarcations of architectural control. The proposed taxonomy transcends this rigidity by integrating hardware-anchored Trusted Execution Environments (TEEs) and remote cryptographic attestation protocols [41,43] directly into the evidentiary calculus governing the burden of proof. By calibrating legal presumptions of defectiveness with cryptographically verifiable vectors of control, this framework maintains privacy-preserving computation while ensuring effective redress for algorithmic harms.

2 The Legislative Landscape: Categorical Expansions and Strategic Withdrawals

The European Union's approach to digital liability has historically been characterized by a dual-track strategy: public law regulation through the AI Act,

and private law remediation through civil liability directives. Understanding the current structural incompatibility requires a forensic examination of recent legislative shifts, notably the expansion of the Revised PLD and the unexpected withdrawal of the AILD.

2.1 The Revised Product Liability Directive: Software as an Actionable Product

Adopted on October 23, 2024, and published in the Official Journal on November 18, 2024 (entering into force on December 8, 2024) as Directive (EU) 2024/2853, the Revised PLD fundamentally overhauls the legacy 1985 product liability regime (Directive 85/374/EEC) to encompass the realities of the digital age [14]. Member States are mandated to transpose the Directive into national law by December 9, 2026. The most critical doctrinal shift resides in Article 4(1) and Recital 13, which explicitly expand the definition of a "product" to include software, thereby capturing operating systems, firmware, applications, and complex AI systems [31], regardless of the mode of supply, whether integrated into physical goods, stored on a device, or provided as a cloud-based Software-as-a-Service (SaaS) [24,32].

This expansion resolves a decades-long accountability void regarding static, centralized software. By formally classifying AI as a product, the Revised PLD ensures that any natural person suffering damage—which now explicitly includes medically recognized psychological harm and the destruction or corruption of non-professional data under Article 6—can claim compensation by proving the defect and the causal link, irrespective of manufacturer fault. The liability period for latent damages has also been significantly extended to 25 years, accommodating the delayed manifestation of algorithmic harms. Furthermore, Article 11 explicitly bans any contractual attempt to limit or exclude liability for defective products, preventing suppliers from utilizing standard risk-management tools such as liability caps against consumers.

Economically, this framework incentivizes manufacturers to internalize algorithmic externalities, encouraging efficient precautions [27]. Legal scholars argue that the economic efficiency of strict liability depends entirely on the manufacturer's capacity to foresee risks, assess probabilities, and implement technical safeguards [26]. However, as the literature indicates, the new regulation may reduce the foreseeability of risks by presenting additional legal uncertainties, adverse selection, and moral hazard problems, particularly when software involves continuous learning or multi-agent modifications [7]. As the subsequent analysis will demonstrate, distributed computational architectures disrupt this fundamental risk-assessment capacity, challenging the economic rationale of the Revised PLD.

The European Commission originally envisioned the Revised PLD operating in tandem with a highly specialized instrument: the AI Liability Directive (AILD) [11,25]. The AILD was designed to adapt fault-based civil liability rules to the unique characteristics of AI [13]. However, prioritizing regulatory simplification, the Commission formally withdrew the AILD in early 2025 [4,12]. This withdrawal

leaves a structural gap, ensuring that AI-related liability falls predominantly under the strict liability regime of the Revised PLD [28]. Without the AILD's specialized mechanisms, the Revised PLD—a relatively broad instrument—now bears the primary burden of ensuring algorithmic accountability, making its evidentiary presumptions the primary battlefield for AI litigation.

The AI Act's Ex-Ante Framework and High-Risk Designations. Operating upstream of the Revised PLD is the AI Act (Regulation (EU) 2024/1689), which entered into force on August 1, 2024, with a phased implementation timeline stretching to 2027 and 2028 for various high-risk systems [15]. Unlike the Revised PLD's ex-post compensatory focus, the AI Act functions as a comprehensive product safety regulation mandating rigorous ex-ante compliance.

Under Annex III, AI systems utilized to evaluate the creditworthiness of natural persons or establish their credit score are explicitly classified as "high-risk" systems [16]. Providers and deployers of high-risk systems are subject to a complex set of obligations. Most notably for this analysis, Article 10 mandates rigorous data governance and quality standards, while Article 12 requires technical capabilities for the automatic recording of events (logs) over the system's lifetime. These logging capabilities must record the period of each use, the reference databases against which input data has been checked, and instances of substantial modification.

Notably, Article 26 delineates the specific obligations of deployers. It dictates that the responsibility to maintain these logs defaults entirely to the deployer if the logs remain under their operational control (Article 26(6)). In scenarios involving multi-institutional collaboration—such as federated credit scoring networks—the AI Act structurally permits and even encourages the decentralization of data control to respect fundamental rights and GDPR minimization principles [8,10]. This creates a distinct and formal legal separation between the entity developing the overarching model (the provider/manufacturer) and the local financial entity executing it and holding the data (the deployer).

3 Federated Learning in Financial Infrastructures: Technical and Legal Asymmetries

3.1 The Architectural Decoupling of Models and Data in Credit Scoring

Federated Learning (FL) structurally decouples model training from data aggregation, pushing computation to the network's edge rather than centralizing raw data. In a cross-institutional credit scoring context, local nodes (e.g., retail banks) train a globally distributed model (W_{global}) strictly on their proprietary, localized data, generating local weight updates (W_{local}). The central orchestrator aggregates only these cryptographically secured updates, never accessing the underlying raw financial data [39,22,9].

Advanced implementations integrate privacy-enhancing technologies like Differential Privacy or homomorphic encryption [5,36,35]. This provides formal

privacy guarantees, making it computationally infeasible for the central server to reverse-engineer gradients to infer sensitive data, achieving rigorous compliance with GDPR data minimization (Article 5(1)(c)). Studies confirm these frameworks maintain high accuracy (e.g., $> 96\%$) while ensuring regulatory compliance without centralized data aggregation [38,33].

3.2 Identifying the "Manufacturer" and "Deployer" in FL Topologies

Applying European Union regulation to this technical topology reveals an asymmetric and highly problematic distribution of legal responsibilities. As Woisetschläger et al. systematically analyze [37], FL represents cooperative learning in which clients and servers naturally share legal responsibility, yet the regulatory texts enforce rigid, singular categorizations (see Table 1).

Under the definitions of the AI Act and the Revised PLD, the entity designing and orchestrating the central FL server acts as the "provider"[1] and the "manufacturer"[2]. This central entity dictates the secure system design, manages the aggregation algorithms, defines the hyper-parameters, and ultimately places the final global model on the market.

Conversely, the local financial institutions operating the edge nodes act as "deployers"[3]. Because the AI Act emphasizes secure overall system design rather than granular control over individual data processing phases, the core provider obligations—and subsequently, the strict manufacturer liability under the Revised PLD—default teleologically to the central server operator. The deployers control the actual training vectors, possess the raw data, and maintain the local execution logs pursuant to Article 26 of the AI Act.

This architectural reality creates a significant legal vulnerability: the central manufacturer is held strictly liable under the law for the outputs of an algorithm shaped by localized data vectors to which they have no physical or technical access.

Trust Assumptions and Threat Model. To formalize this legal-technical intersection for rigorous analysis, we must define the cryptographic and architectural threat model typical for high-risk federated credit scoring networks. The network operates under specific trust assumptions:

- **Central Orchestrator (Manufacturer):** Considered *Honest-but-Curious*. The orchestrator faithfully executes the aggregation protocol and maintains the global model. It seeks to optimize the model without intentionally breaching deployer privacy, but it is incentivized to evade strict liability under the Revised PLD by shifting blame to the edge nodes.
- **Local Edge Nodes (Deployers):** Considered *Malicious/Rational*. They possess the proprietary raw financial data. A malicious or highly negligent

[1] Regulation (EU) 2024/1689 (AI Act), Article 3(3).
[2] Directive (EU) 2024/2853 (Revised PLD), Article 4(1).
[3] AI Act, Article 3(4).

deployer may attempt to manipulate the local gradient (e.g., via data poisoning) to bias the model for competitive advantage, subsequently attempting to exploit the evidentiary asymmetries of Article 10(4) of the Revised PLD to shift liability onto the central Manufacturer.

This threat model highlights the necessity for a verifiable mechanism to resolve disputes when algorithmic harm occurs.

4 The Evidentiary Deficit and the Article 10 Paradox

The structural tension between privacy-preserving FL architectures and European tort law is most pronounced within the evidentiary mechanics of the Revised PLD. The Directive explicitly seeks to alleviate the historically heavy burden of proof placed on consumers facing complex technological products by instituting stringent disclosure mandates and broad rebuttable presumptions. However, in distributed architectures, these very mechanisms inadvertently turn privacy compliance against the manufacturer.

4.1 Disclosure Mandates Under Article 9(1)

Article 9(1) of the Revised PLD empowers national courts to order a defendant manufacturer to disclose "necessary and proportionate" evidence that is "at its disposal" to assist the claimant in proving defectiveness [31,30]. In centralized AI architectures, retaining exhaustive system telemetry, comprehensive training datasets, and detailed inference logs constitutes standard industry practice. If a centralized provider refuses to disclose this data during a liability claim—for instance, regarding a discriminatory or defective credit decision that caused financial exclusion—they face immediate and significant legal repercussions.

However, in Federated Learning architectures, the central entity is cryptographically and architecturally blind to the local data inputs that caused the algorithmic drift or defect. The manufacturer intentionally designed the system to comply with GDPR data minimization and the AI Act's decentralized governance frameworks. Consequently, the local training data evades the central infrastructure and genuinely falls outside the scope of evidence "at its disposal."

While a national court may acknowledge the factual reality that a manufacturer cannot physically disclose what it does not possess, this lack of disclosure does not absolve the manufacturer; rather, it triggers a substantial cascade of strict liability presumptions.

4.2 The Rebuttable Presumptions of Defectiveness (Article 10)

Article 10 of the Revised PLD introduces a series of rebuttable presumptions designed to substantially shift the burden of proof in favor of the claimant.

- **Article 10(2)(a)** explicitly presumes a product is defective if the defendant fails to comply with the disclosure obligations under Article 9.

- **Article 10(3)** presumes causation if defectiveness is established and the damage is of a kind typically consistent with the defect in question.
- **Article 10(4)** triggers a presumption of both defectiveness and causality whenever a claimant faces "excessive difficulties" in proving their case, "in particular due to technical or scientific complexity," provided the claimant demonstrates a mere *prima facie* likelihood of a defect.

AI systems, particularly the high-dimensional deep neural networks utilized in advanced FL credit scoring, are the primary example of "technical or scientific complexity" due to their inherent opacity, non-linear decision-making, and "black box" characteristics [24,32]. Furthermore, the non-IID (independent and identically distributed) nature of data across different financial institutions in an FL network exponentially increases the complexity of auditing the model [8]. Therefore, plaintiffs alleging algorithmic harm will almost universally, and successfully, invoke the Article 10(4) presumption.

4.3 The Article 10 Paradox: Telemetric Blindness and Strict Liability

The interaction of these legislative articles creates a systemic paradox for FL manufacturers. Once the claimant invokes the Article 10(4) presumption due to the complexity of the distributed credit model, the burden of proof shifts to the manufacturer to rebut the presumption of defectiveness.

To mount a successful legal defense—specifically relying on the exemption under Article 11(1)(c), which requires proving that the probable defect did not exist at the time the model was placed on the market or that it emerged post-deployment due to factors entirely outside the manufacturer's control—the manufacturer requires forensic access to the local training logs and data vectors.

Because the FL architecture cryptographically precludes the central server from exfiltrating this localized telemetry from the deployer's nodes, the manufacturer is technically unable to gather the exculpatory evidence required by law. Thus, engineering a system for maximum privacy and strict AI Act compliance legally vitiates the provider's evidentiary capacity to defend themselves under the Revised PLD. The rebuttable presumption inevitably hardens into unmitigated strict liability, disproportionately penalizing the manufacturer for anomalies, biases, or data poisoning introduced entirely by the local deployer.

To resolve this paradox without compromising data minimization principles, legal frameworks must integrate technical mechanisms of verifiable trust, as discussed in the following section.

5 Reconciling Ex-Post Liability with Ex-Ante Compliance: The Role of Cryptographic Attestation

To resolve this legislative incompatibility without abandoning the privacy and security guarantees of Federated Learning, the legal framework must evolve to interface directly with advanced cryptographic mechanisms. Traditional legal contracts between the central server and the edge nodes are fundamentally

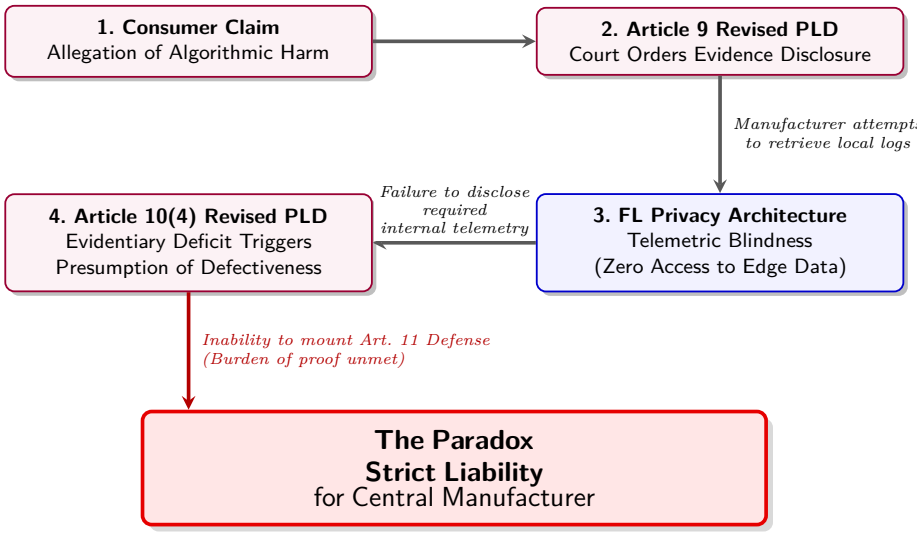

Fig. 1. The Article 10 Paradox: A flowchart illustrating how privacy-by-design inadvertently leads to strict liability.

insufficient, as post-hoc contractual indemnification does not solve the primary evidentiary deficit in tort litigation against a consumer. Instead, hardware-based cryptographic attestation offers a verifiable, mathematical proxy for data visibility.

5.1 Hardware-Enforced Cryptographic Attestation (TEEs)

Trusted Execution Environments (TEEs)—such as Intel SGX, AMD SEV, and ARM TrustZone—provide hardware-enforced secure enclaves that cryptographically isolate computations (e.g., local training epochs) from the host machine's operating system [43,40,19]. Rather than relying on post-hoc contractual indemnification, TEEs supply a tamper-evident, cryptographically signed audit trail of the local model's execution state. By deploying the FL client application within a TEE, the central manufacturer secures an indisputable proof of execution integrity. This mathematical verification serves as the precise legal mechanism required to trigger an Article 11 exemption under the Revised PLD, establishing fault localization without compromising GDPR minimization mandates or requiring invasive access to the underlying raw financial data [41,20].

The cornerstone of this technical resolution is Remote Cryptographic Attestation [41,20]. Attestation is a mechanism by which a TEE generates a cryptographically signed report—anchored securely by a hardware root-of-trust embedded in the silicon—proving to a remote verifier (the central manufacturer) that a specific, unmodified piece of software has been securely loaded into a genuine enclave.

At the conclusion of a local training epoch, the TEE generates a signed cryptographic proof defined as:

$$\sigma = \mathsf{Sign}_{\mathsf{SK}_{\mathrm{TEE}}}(\mathsf{Hash}(\mathrm{Binary}) \parallel \nabla \mathbf{W}_i \parallel H(D_i) \parallel \mathrm{Timestamp}) \tag{1}$$

This payload binds the exact software binary, the generated output gradient ($\nabla \mathbf{W}_i$), and a binding hash commitment of the local data provenance ($H(D_i)$). Through attestation, the manufacturer can mathematically verify that the local deployer's node adhered to agreed-upon data preprocessing standards, did not inject malicious or unformatted gradients, and utilized the correct neural network architecture [6,1].

Verifiable Claims against Data Poisoning and Byzantine Faults. In distributed credit scoring, the most severe risks leading to Revised PLD liability are Data Poisoning and Byzantine faults [3,18]. Data poisoning occurs when a malicious or highly negligent local deployer injects manipulated training data (e.g., highly biased demographic loan data intended to alter the model's fairness metrics) to subvert the global model. Because the central server in standard FL only observes aggregated gradients, backdoor and targeted poisoning attacks are notoriously difficult to detect, often masquerading as natural non-IID data variance.

By integrating TEEs and Remote Attestation, the FL framework evolves into a Zero-Trust Federated Learning environment [34]. Frameworks such as *Sentinel* employ code instrumentation to track control-flow and monitor critical variables within the local TEE, generating the attestation report (σ) that is securely transmitted to the server. The attestation report acts as a tamper-evident cryptographic audit trail. If the central server detects anomalous gradients or a severe degradation in model fairness metrics, it can query the attestation logs.

If the attestation fails, lacks a valid signature, or reveals a breach of protocol, the manufacturer isolates the anomaly and discards the gradient. Crucially, in the context of the Revised PLD, this cryptographic proof serves as the exact legal mechanism required to trigger an Article 11 exemption, providing the evidentiary link required to establish fault localization without compromising GDPR minimization mandates. Importantly, while TEEs introduce minor computational overheads, modern enclaves easily accommodate the memory constraints of tabular data models used in financial networks, making this a highly feasible technical tradeoff.

6 Proposing a Differentiated Liability Taxonomy (*Lex Ferenda*)

The application of the Revised PLD's Article 10 presumptions relies heavily on an outdated, binary understanding of software architecture: either the manufacturer possesses full control, or they do not. As established, distributed AI systems exist on a highly nuanced spectrum of control. Relying solely on the unpredictable judicial interpretation of "technical complexity" by national courts on a case-by-case basis will inevitably result in significant legal fragmentation across the EU and chill investment in critical privacy-enhancing technologies [17,42].

Therefore, this research report formalizes a legislative or regulatory integration of a Differentiated Liability Taxonomy (*de lege ferenda*). This framework supersedes the rigid binary application of Article 10, precisely calibrating the legal

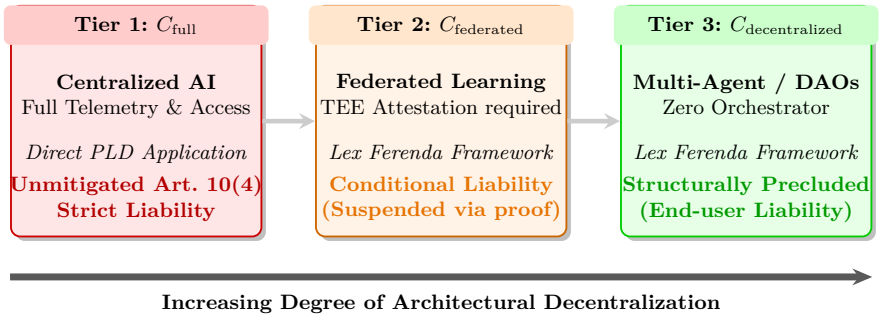

Fig. 2. The Proposed Differentiated Liability Taxonomy (*Lex Ferenda*): Calibrating the severity of Article 10(4) evidentiary presumptions directly against the manufacturer's verifiable architectural control.

burden of proof directly against the manufacturer's cryptographically verified access matrix.

6.1 Tier 1: Full Telemetric Control (C_{full})

- **Applicability:** Traditional, centralized machine learning architectures where all training and inference data are aggregated into a single repository managed entirely by the manufacturer.
- **Control Vector:** The manufacturer retains complete telemetric visibility over input vectors, training telemetry, and output inferences.
- **Liability Application:**
  - Article 9 disclosure orders apply fully and without exception.
  - Failure to disclose telemetry legitimately triggers the Article 10(2)(a) presumption of defectiveness.
  - The Article 10(4) burden-of-proof reversals apply in full, aligning with strict liability principles. The manufacturer bears the burden of rebutting the presumption of defectiveness, reflecting its complete architectural control.

6.2 Tier 2: Asymmetric Federated Control ($C_{\text{federated}}$)

- **Applicability:** Federated Learning architectures utilized in high-risk environments (e.g., cross-institutional credit scoring), where the manufacturer controls the global model (W_{global}) and aggregation algorithms, but localized deployers control the raw training data and compute local updates (W_{local}).
- **Control Vector:** The manufacturer possesses zero visibility into raw data but maintains structural oversight via mandatory hardware-based cryptographic attestation and secure execution environments.
- **Liability Application:**
  - To prevent the unjust shifting of strict liability onto the manufacturer for deployer-induced errors, the application of Article 10(4) (presumption of defectiveness due to complexity) is conditionally suspended for the

manufacturer, contingent upon the provision of a formalized exculpatory mechanism.
- When a claimant establishes a *prima facie* case of algorithmic harm, the manufacturer can compel the relevant local deployer to supply a cryptographic attestation of data integrity and execution compliance.
- If the attestation proves the deployer operated outside defined parameters (e.g., executing a data poisoning attack, bypassing the TEE, failing to apply required noise for differential privacy), the manufacturer invokes the Article 11(1)(c) liability exemption. The law formally classifies the deployer's unauthorized intervention as a "post-deployment external alteration" outside the manufacturer's control, effectively shielding the central provider from strict liability.
- If the attestation is mathematically valid but the global model is still demonstrably defective, liability remains strictly with the manufacturer, as the defect fundamentally stems from the global aggregation algorithm or initial baseline model rather than edge-node tampering.

6.3 Tier 3: Fully Decentralized Environments ($C_{\text{decentralized}}$)

– **Applicability:** Permissionless, multi-agent autonomous systems, decentralized autonomous organizations (DAOs), or fully decentralized blockchain-based learning networks where no central orchestrator exists [7].
– **Control Vector:** The initial developer releases open-source code or a foundational architecture and lacks any ex-post intervention capacity, aggregation control, or attestation authority.
– **Liability Application:**
  - The taxonomy structurally precludes the application of Article 10(4) against the initial manufacturer. The extreme opacity and complexity are classified as an inherent technological reality of the decentralized protocol, not a condition sustained by the developer.
  - This grants immediate effect to the Article 11(1)(a) exemption (the manufacturer did not put the product into circulation in a controlled manner) or Article 11(1)(c), based on demonstrably zero architectural control. Liability shifts entirely to the end-users or deployers intentionally executing the autonomous nodes.

7 Operationalizing the Taxonomy: Legal and Contractual Mechanisms

Translating the theoretical Differentiated Liability Taxonomy into actionable legal strategy requires integrating these advanced technical mechanisms with existing commercial and tort law frameworks across the European Union.

Table 1. Matrix of Differentiated Liability and Role Mapping in AI Architectures

Liability Tier	AI Act / Revised PLD	Data Access	Cryptographic Attestation	Applicability of Art. 10(4)
Tier 1: C_{full} **Centralized ML**	Provider / Manufacturer (Central)	Full (Raw data, logs, telemetry)	Not Applicable (Direct access)	Full unmitigated application (Strict Liability)
Tier 2: $C_{\text{federated}}$ **Federated Learning**	Orchestrator: Provider / Manufacturer Edge Node: Deployer	Zero (Orchestrator blind to raw data, views only gradients)	Mandatory (TEEs, SGX, Remote Attestation)	Conditional (Suspended upon valid attestation of deployer fault)
Tier 3: $C_{\text{decentralized}}$ **Multi-Agent / DAO**	Decentralized Deployers / End-users	Absolute Zero (No centralized server)	N/A (No central verifying authority)	Structurally Precluded

7.1 Leveraging Article 11 Exemptions through Cryptography

The fundamental goal of the Tier 2 taxonomy is to allow manufacturers to legitimately and practically access the exemptions codified in Article 11 of the Revised PLD. Article 11(1)(c) explicitly provides that an economic operator is exempted from liability if they prove that the defectiveness that caused the damage did not exist at the time the product was placed on the market, or that the defectiveness came into being subsequently due to factors outside their control.

Crucially, the proposed Tier 2 framework does not diminish consumer protection or alter the non-waivable nature of strict liability under Article 11 of the Revised PLD; rather, it provides a precise technical mechanism for *defect localization*. The remote attestation report generated by an Intel SGX or Gramine enclave serves as a tamper-evident, cryptographically signed ledger of the local model's execution state [2,29]. By presenting this attestation to a judicial authority, the manufacturer translates cryptographic mathematical certainty into legal burden-of-proof satisfaction. It undeniably proves that the statistical deviation or bias causing the claimant's financial harm was introduced post-deployment by the deployer's specific, localized data vectors, satisfying the high Article 11(1)(c) evidentiary threshold and correctly channeling delictual liability directly onto the deployer.

7.2 General Terms and Conditions and Rights of Recourse

While the Revised PLD governs non-contractual strict liability toward harmed consumers, commercial relations between the central orchestrator (manufacturer)

and local financial nodes (deployers) are dictated by contract. As Woisetschläger et al. note [37], ensuring compliance with both the AI Act and GDPR (Articles 26 and 28) requires the development of highly specific General Terms and Conditions binding both parties in the federated network. Under Article 14 of the Revised PLD, manufacturers retain an explicit statutory right of recourse against third parties whose actions contribute to or cause a defect. By integrating mandatory TEE execution and cryptographic attestation requirements into these bilateral commercial agreements, orchestrators establish an immediate, indisputable basis to exercise their right of recourse and claim full indemnification from deployers whose local data interventions trigger primary liability under Article 10(4).

8 Conclusion

Directive (EU) 2024/2853 represents a major and consequential expansion of European tort and technology law, bringing the complexities of "black-box" algorithmic decision-making firmly within the domain of strict product liability. Following the withdrawal of the highly specialized AI Liability Directive in early 2025, the Revised PLD stands as the primary regulation for victims of algorithmic harm, heavily reliant on procedural evidentiary presumptions that inherently, and intentionally, favor the claimant.

However, applying the Revised PLD's stringent disclosure mandates and burden-of-proof reversals to distributed, privacy-preserving architectures such as Federated Learning creates a significant structural incompatibility. By engineering systems to comply with the GDPR and the AI Act's rigorous mandates for data minimization and decentralized governance, FL manufacturers inherently surrender the telemetric visibility required to defend themselves against the Revised PLD's strict liability mechanisms and face liability for the deployment of privacy-enhancing technologies.

The implementation of the Differentiated Liability Taxonomy *de lege ferenda* resolves this critical gap. By entirely abandoning an outdated, binary understanding of architectural control, the taxonomy successfully integrates hardware-based Trusted Execution Environments and cryptographic attestation directly into the legal calculus of burden-of-proof. This synthesis of cryptographic verification and legal presumption ensures that consumers retain exacting, highly effective remedial mechanisms for algorithmic harm, while simultaneously preventing the disproportionate penalization of manufacturers deploying privacy-compliant, decentralized computational topologies. Calibrating legal accountability directly to quantifiable, cryptographically verifiable vectors of control represents the necessary, inevitable evolution of European liability law in the era of distributed AI. Future research should focus on the empirical verification of this taxonomy within the European Union's AI regulatory sandboxes. Testing these cryptographic attestation mechanisms in controlled, cross-border financial environments prior to the Revised PLD's transposition deadline in December 2026 will provide critical operational metrics for both regulators and industry stakeholders.

References

1. Abbasi Tadi, A., Alhadidi, D., Rueda, L.: Trustformer: A trusted federated transformer. arXiv preprint arXiv:2501.11706 (2025), https://arxiv.org/html/2501.11706v1
2. Ali, M.H., Mohammed Aqib, A., Syed Muneer, H., et al.: Privacy-preserving federated learning with verifiable fairness guarantees. arXiv preprint arXiv:2601.12447 (2025), https://arxiv.org/html/2601.12447v1
3. Aljanabi, M., Mohammed, S.Y., Omran, A.H.: Detecting data poisoning attacks in federated learning for healthcare applications using deep learning. Iraqi Journal for Computer Science and Mathematics **4**(4), 18 (2023). https://doi.org/10.52866/ijcsm.2023.04.04.018, https://ijcsm.researchcommons.org/ijcsm/vol4/iss4/18
4. AmCham EU: Joint statement: Withdrawal of the artificial intelligence liability directive (2025), https://www.amchameu.eu/position-papers/joint-statement-artificial-intelligence-liability-directive-aild
5. Amed, S.: Fsl-bdp: Federated survival learning with bayesian differential privacy for credit risk modeling. arXiv preprint arXiv:2601.11134 (2026), https://arxiv.org/html/2601.11134v1
6. Bai, J., Chowdhuryy, M.H.I., Li, J., et al.: Phantom: Privacy-preserving deep neural network model obfuscation in heterogeneous tee and gpu system. In: 34th USENIX Security Symposium (2025), https://www.usenix.org/system/files/usenixsecurity25-bai-juyang.pdf
7. Budi, K., Tubagus, A.R.: The deterrence dilemma: Assessing criminal liability standards for emerging digital offenses. RLR (2025), https://journal.sabtida.com/index.php/rlr/article/download/86/49
8. Butt, T.A., Iqbal, M.: Governing what the eu ai act excludes: Accountability for autonomous ai agents in smart city critical infrastructure. arXiv preprint arXiv:2605.01091 (2026), https://arxiv.org/pdf/2605.01091
9. Chen, J., Estrada, D.H., Guan, H.: Bank credit default risk assessment model based on federated learning. Informatica **48**(8), 133–142 (2024). https://doi.org/10.31449/inf.v48i8.5601, https://www.informatica.si/index.php/informatica/article/view/5601
10. CSIS: Protecting data privacy as a baseline for responsible ai (2024), https://www.csis.org/analysis/protecting-data-privacy-baseline-responsible-ai
11. European Commission: Proposal for a directive on adapting non-contractual civil liability rules to artificial intelligence (ai liability directive) - 52022pc0496 (2022), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52022PC0496
12. European Parliament: Ai liability directive | legislative train schedule (2025), https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-ai-liability-directive
13. European Parliament: Artificial intelligence and civil liability (2025), https://www.europarl.europa.eu/RegData/etudes/STUD/2025/776426/IUST_STU(2025)776426_EN.pdf
14. European Parliament and Council: Directive (eu) 2024/2853 of the european parliament and of the council (revised pld) (2024), https://www.lexaris.de/library/tableofcontents/6811882
15. European Parliament and Council: Regulation (eu) 2024/1689 of the european parliament and of the council (ai act) (2024), https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689

16. European Union: Annex iii: High-risk ai systems (2024), https://artificialintelligenceact.eu/annex/3/
17. Faure, M.G.: Liability rules for ai-related harm (2022), https://cris.maastrichtuniversity.nl/ws/portalfiles/portal/108561069/Faure_2022_Liability_Rules_for_AI_Related_Harm.pdf
18. Galanis, N.: Defending against data poisoning attacks in federated learning via user elimination. arXiv preprint arXiv:2404.12778 (2024), https://arxiv.org/html/2404.12778v1
19. Gomez, R.: Tee-based key-value stores : a survey. arXiv preprint arXiv:2501.03118 (2025), https://arxiv.org/html/2501.03118v1
20. Guo, J., Paverd, A., Vaswani, K., Pietzuch, P.: Verifiablefl: Providing verifiable claims to federated learning using exclaves. arXiv preprint arXiv:2412.10537 (2026), https://arxiv.org/pdf/2412.10537
21. Houston, C., Kennedy, Hilal, A., Momeni, M.: The role of federated learning in improving financial security: A survey. arXiv preprint arXiv:2510.14991 (2025), https://arxiv.org/html/2510.14991v1
22. Kaarat, A., Batthula, V.J.R., Segall, R.S.: Unified federated ai framework for credit scoring: For privacy, fairness, and scalability. International Journal of Artificial Intelligence (AI) in Business and Management **2**(1), 1–31 (2026). https://doi.org/10.4018/IJAIBM.398859, https://www.igi-global.com/article/unified-federated-ai-framework-for-credit-scoring/398859
23. Karatas, C., Karadogan, H., Ertug, A.Y., et al.: Federated learning architecture: Data privacy and system security approaches. arXiv preprint arXiv:2607.09391 (2026), https://arxiv.org/html/2607.09391v1
24. Kloza, D., Drechsler, L., Fernandes, E., et al.: If it ain't broke, don't fix it? ten improvements for the upcoming tenth anniversary of the general data protection regulation. Computer Law & Security Review **60** (2026), https://lirias.kuleuven.be/retrieve/e008ceed-526f-40c0-9034-a228398f085f
25. KU Leuven: Citip working paper - ku leuven (2023), https://www.law.kuleuven.be/citip/en/docs/books/citip-working-paper-ai-liability.pdf
26. Li, S.: Google scholar profile (2024), https://scholar.google.com.tw/citations?user=mTOl4eEAAAAJ&hl=vi
27. Li, S., Faure, M.G.: The revised product liability directive: A law and economics analysis. Journal of European Tort Law **15**(2), 140–171 (2024). https://doi.org/10.1515/jetl-2024-0010, https://pure.eur.nl/files/162305980/10.1515_jetl-2024-0010.pdf
28. Nannini, L.: When less regulation means more complexity: The eu ai liability directive withdrawal. AAAI AIES (2025), https://ojs.aaai.org/index.php/AIES/article/view/36679/38817
29. Nguyen, T.L., De Oliveira, M.T., Braeken, A., Ding, A.Y., Pham, Q.V.: Toward verifiable federated unlearning: Framework, challenges, and the road ahead. IEEE Internet Computing **30**(1), 59–70 (2026). https://doi.org/10.1109/MIC.2026.3656638, https://www.computer.org/csdl/magazine/ic/2026/01/11359714/2dsHn71Y6qc
30. Rimkutė, D.: The new eu product liability directive: Doctrinal analysis. Access to Justice in Eastern Europe **8**(Spec), 74–97 (2025). https://doi.org/10.33327/AJEE-18-8.S-c000160, https://ajee-journal.com/upload/attaches/att_1767169778.pdf
31. Rimkutė, D.: The regulatory function of the new product liability directive. European Journal of Risk Regulation (2026), https://www.cambridge.org/core/jou

rnals/european-journal-of-risk-regulation/article/regulatory-function-of-the-new-product-liability-directive/913EE57212715C4E5F9EEFA0A05A2CCE
32. Sarabdeen, J.: Ai product liability under eu and canadian laws. Frontiers in Artificial Intelligence **9**, 1709945 (2026). https://doi.org/10.3389/frai.2026.1709945, https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2026.1709945/full
33. Shahsavari, Y., Baseri, Y., Hafid, A., et al.: Integration of federated learning and blockchain in health care: Tutorial on medical data, architectures, privacy, security, and regulatory compliance. Journal of Medical Internet Research **28**, e80178 (2026). https://doi.org/10.2196/80178, https://www.jmir.org/2026/1/e80178, pMID: 42296530, PMCID: 13268642
34. Singh, S.K., Roy, J., So, M.: Zero-trust agentic federated learning for secure iiot defense systems. arXiv preprint arXiv:2512.23809 (2025), https://arxiv.org/html/2512.23809v1
35. Valadares, D.C.G., Perkusich, A., Martins, et al.: Privacy-preserving blockchain technologies. PMC / Sensors **23**(16), 7206 (2023). https://doi.org/10.3390/s23167206, https://pmc.ncbi.nlm.nih.gov/articles/PMC10457747/, pMID: 37631709, PMCID: PMC10457747
36. Wang, Z.: Confidential federated learning with homomorphic encryption. DiVA Portal (2024), https://www.diva-portal.org/smash/get/diva2:1835970/FULLTEXT01.pdf
37. Woisetschläger, H., Mertel, S., Krönke, et al.: Federated learning and ai regulation in the european union: Who is responsible? – an interdisciplinary analysis. arXiv preprint arXiv:2407.08105 (2024), https://www.researchgate.net/publication/382178426_Federated_Learning_and_AI_Regulation_in_the_European_Union_Who_is_liable_An_Interdisciplinary_Analysis
38. Yang, F., Abedin, M.Z., Hajek, P.: An explainable federated learning and blockchain based secure credit modeling method. European Journal of Operational Research **317**(2), 449–467 (2024). https://doi.org/10.1016/j.ejor.2023.08.040, https://www.sciencedirect.com/science/article/abs/pii/S0377221723006677
39. Yaramolu, L.S.K.G.: Privacy-driven federated ai in financial fraud detection. WJAETS (2025), https://wjaets.com/sites/default/files/fulltext_pdf/WJAETS-2025-0493.pdf
40. Yuhala, P.: Enhancing Security and Performance in Trusted Execution Environments. Ph.D. thesis, University of Neuchatel (2024), https://yuhala.github.io/assets/pdf/thesis-peterson-yuhala.pdf
41. Zhang, C., Jin, H., Shi, S., Yu, et al.: Enabling trustworthy federated learning via remote attestation for mitigating byzantine threats. arXiv preprint arXiv:2509.00634 (2025), https://arxiv.org/abs/2509.00634
42. Zhang, L.: When compliance becomes exclusion: Proportionality, liability legitimacy, and the governance of third-party ai fine-tuning (2026), https://aixiv.science/pdf/aixiv.260707.000002
43. Zhang, L., Duan, B., Li, J., Ma, Z., Cao, X.: A tee-based federated privacy protection method: Proposal and implementation. Applied Sciences **14**(8), 3533 (2024). https://doi.org/10.3390/app14083533, https://www.mdpi.com/2076-3417/14/8/3533