International Journal of
Law in Changing World

Volume 5 Issue 1 (2026) ISSN 2764-6068

Research article
JNL: https://ijlcw.emnuvens.com.br/revista
DOI: https://doi.org/10.54934/ijlcw.v5i1.188

CIVIL LIABILITY IN THE AGE OF ARTIFICIAL INTELLIGENCE: FROM TRADITIONAL TO MODERN THEORIES

Abbas Mirshekari
University of Tehran, Tehran, Iran

Jamshid Zargari
University of Applied Science and Technology, Tehran, Iran

Article Information:

Received
June 2, 2025
Reviewed & Revised
June 5, 2025
Accepted
June 6, 2025
Published
July 30, 2026

Keywords:

Artificial Intelligence,
civil liability,
collective compensation,
electronic personhood,
hybrid liability

ABSTRACT | 摘要 | RESUMEN

The integration of AI into social structures has exposed fundamental shortcomings in traditional civil liability theories, which struggle to address autonomous action, opaque processes, and fragmented accountability across multiple stakeholders. This study adopts a doctrinal and comparative-analytical method to evaluate modern liability frameworks—including electronic personhood, collective compensation, and hybrid liability—against the limitations of conventional approaches. The findings indicate that dynamic legal personhood, vicarious liability for stakeholders, objective risk thresholds, and expert oversight mechanisms offer coherent pathways to address unclear causation and non-transparent algorithms. The primary contribution is a conditional-pluralism framework that balances victim protection, fair risk allocation, and technological progress, providing practical tools for managing AI-related disputes in an evolving legal landscape.

FOR CITATION:

Mirshekari, A., & Zargari, J. (2026). Civil Liability in the Age of Artificial Intelligence: From Traditional to Modern Theories. International Journal of Law in Changing World, 5 (1), 123-144. DOI: https://doi.org/10.54934/ijlcw.v5i1.188

1. INTRODUCTION

As artificial Intelligence (AI) increasingly permeates everyday life—restructuring legal, economic, and social frameworks—the question of civil liability arising from AI-related harms has emerged as a pressing global concern. International efforts, notably the European Union's AI Act, underscore an urgent need for coherent legal frameworks capable of addressing civil liability challenges.[1] In the absence of such frameworks, affected parties risk suffering harm without effective recourse to compensation. [1]; [16]

Existing scholarship has made sustained efforts to align legal doctrines with the practical realities of AI. Early studies highlighted the inadequacies of traditional liability theories when applied to autonomous and complex systems. [2]; [29]; [41] More recent work signals a shift toward innovative theories that better account for AI's distinctive characteristics. [5]; [11] This growing body of research, often grounded in comparative analyses and case studies, has exposed significant gaps in current legal approaches, thereby calling for a fundamental re-examination of established legal concepts.

Despite these advances, the literature remains characterized by a critical lacuna: While existing studies have effectively diagnosed the shortcomings of traditional liability regimes and proposed isolated solutions, a comprehensive, integrated framework that systematically compares traditional and modern approaches and offers a context-sensitive synthesis is notably absent. Most prior research tends to advocate for a single modern theory in isolation or confines its analysis to specific jurisdictions or sectors. [19]; [30]

The present study directly addresses this gap by undertaking a doctrinal and comparative-analytical examination that goes beyond mere diagnosis. Its originality resides in three interrelated dimensions. First, unlike previous work that predominantly contrasts "traditional" and "modern" theories as binary opposites, this study investigates their potential complementarity under clearly specified conditions. Second, it proposes a conditional and pluralistic approach to civil liability in the AI era—one that rejects monistic solutions and instead maps distinct liability tools to concrete technical, contextual, and risk-related variables. Third, it places particular emphasis on the operational and institutional challenges that any reform proposal must confront in practice, thereby bridging the gap between theoretical sophistication and real-world implementability.

Accordingly, this paper provides a precise examination of the transition from traditional civil liability to modern theories, in response to the emerging demands of the AI era. It is structured around several focused research questions: Which civil liability theories have been, or could be, developed to

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 on artificial intelligence (Artificial Intelligence Act (AIA)) [2024] OJ L 202/1–xx.

https://doi.org/10.54934/ijlcw.v5i1.188

manage AI-generated risks, while ensuring fair victim compensation and supporting continued innovation? Why do traditional theories prove inadequate when applied to AI-related harm? What obstacles do modern theories face, and which practical solutions may strengthen them?

Through a systematic comparison of established and emerging theories, this study evaluates their respective strengths and weaknesses and offers concrete recommendations for legal policymaking. In doing so, it aspires not merely to add another critique of existing frameworks, but to make a pragmatic contribution to the development of adaptive legal responses, capable of evolving alongside AI capabilities

2. METHODOLOGY

This study adopts a qualitative, doctrinal, and comparative-analytical research design. Doctrinal legal research is particularly suitable for examining civil liability theories because it operates through the systematic analysis, critique, and synthesis of legal rules, principles, and doctrines.

First, traditional civil liability theories (fault, no-fault/strict, vicarious, and product liability) are reconstructed from primary legal sources as well as authoritative secondary sources (peer-reviewed articles and recognized legal commentaries). Second, a comparative-descriptive analysis is employed to examine how three modern frameworks (electronic personhood, collective compensation, and hybrid liability) have been articulated in recent legal scholarship, institutional proposals (particularly the European Parliament's 2017 Resolution on Robotics and the EU AI Act), and soft law instruments (e.g., OECD AI Principles). Third, the study synthesizes these strands into a conditional-pluralist model, mapping each theory to concrete technical, contextual, and risk-related variables.

The primary focus on European Union law (especially the AI Act, the proposed AI Liability Directive, and the Civil Law Rules on Robotics) reflects the EU's status as the first major jurisdiction to adopt comprehensive, risk-based AI regulation, thereby providing the most developed positive legal framework for comparative doctrinal analysis. US materials (e.g., Colorado SB 24-205, Executive Order 14110) and select international soft law instruments (OECD, UNESCO, ISO/IEC standards) are included where they illustrate alternative approaches or emerging consensus on key concepts. The selection is purposive, rather than random, prioritizing jurisdictions and instruments that have produced explicit doctrinal innovations directly relevant to AI liability. No empirical data from courts or surveys were collected.

As a library-based doctrinal study, this research does not generate new empirical evidence and therefore cannot empirically verify how frequently certain challenges occur in practice. Consequently, the

https://doi.org/10.54934/ijlcw.v5i1.188

suggestions offered in Section 5 are presented as normative, conceptually reasoned proposals rather than empirical solutions.

3. TRADITIONAL THEORIES

Traditional civil liability theories were designed for human behaviors that have foreseeable and nearly similar consequences. Today, they must contend with AI's autonomous behavior, self-evolving mechanisms, and non-transparent decision-making. The present section explores how fault, no-fault, and product liability theories interact with AI technologies, and where these approaches appear likely to fall short in responding to the diverse harms that AI systems may produce.

3.1 Fault Liability

The first theory to be addressed is fault-based liability. This form of civil liability attributes responsibility to a party for damages resulting from negligent conduct that causes harm, requiring proof of fault, damage, and causation. [11]; [25] This framework is theoretically applicable to any actor in the AI lifecycle, including developers, manufacturers, or operators. However, it faces significant hurdles due to AI's inherent characteristics.

A primary obstacle is the so-called "black box" problem, which is particularly characteristic of machine-learning approaches based on deep neural networks, such as deep learning and transformer architectures. This opacity undermines the ability to trace outcomes back to specific human inputs or programming decisions, thereby complicating the establishment of causation and fault.[2] [30]; [35] The issue has motivated a distinct research agenda known as mechanistic interpretability, which seeks to reverse-engineer the learned algorithms and computational circuits within models. [6]

Moreover, AI's autonomous and self-learning capabilities further complicate fault attribution, as system behavior at the time of harm may diverge significantly from its initial design, rendering it difficult to assess what a reasonable person—such as the developer or operator—could have foreseen or prevented. [2]; [29]; [41] The notion of self-learning should be qualified by model architecture. Deterministic models follow fixed computations and produce identical outputs for identical inputs, enabling greater predictability. In contrast, probabilistic models—prevalent in deep learning and especially in generative transformer-based systems—generate outputs stochastically, leading to variability even under identical conditions. [50]

[2] G20 AI Principles [2019] adopted at the G20 Ministerial Meeting on Trade and Digital Economy, Tsukuba, Ibaraki Prefecture, Japan, 8–9 June 2019, Section 1.3; S.B. 24-205 Concerning Consumer Protections in Interactions with Artificial Intelligence Systems, 75th General Assembly, Regular Session (Colorado, USA) [2024], Section 6-1-1702(2); AIA, Article 78; Executive Order 14110: Safe, secure, and trustworthy development and use of artificial intelligence [2023] Federal Register 88(210)/75191, Section 10.

https://doi.org/10.54934/ijlcw.v5i1.188

The human-centric reasonable person standard, which evaluates conduct based on how a prudent and average individual would have acted under the same circumstances, struggles to accommodate non-human AI behavior and may require redefinition to reflect AI's unique operational modes.[3] [4] Finally, informational asymmetry—where AI stakeholders typically possess far more or better information than victims—combined with practical barriers to judicial redress, substantially intensifies the burden of proving fault and causation. [38]

3.2 No-Fault Liability

The second traditional theory covers a group of civil liability doctrines that remove fault as a necessary element to establish liability. Put simply, although these doctrines differ in theory and practice across jurisdictions, they generally impose liability on those who create or manage risks, regardless of fault.[4] [8]; [25]

3.2.1 Strict Liability

By following strict liability, stakeholders of high-risk AI systems cannot generally escape liability by merely invoking due diligence. [10]; [13] The European Union's AI regulatory framework reflects this approach[5] — it classifies AI systems according to risk level and requires strict compliance measures and ex-ante assessments. [19] That said, the exceptional character of these liability regimes limits their broad application to AI, since lawmakers usually confine them to clearly defined, high-risk activities. [16]

Notably, this doctrine also raises concerns about a chilling effect on innovation—the risk that fear of legal sanctions may discourage legitimate activity. Extensive liability exposure deters investment and experimentation for AI development. [43] Finally, while some regulatory efforts seek to define what constitutes a high-risk AI system, particularly in contexts involving critical decisions, serious difficulties remain.[6] Measuring AI-related risks and setting clear liability thresholds continues to pose challenges, largely because historical data is limited, and the long-term effects of AI remain only partially understood. [8]

[3] S.B. 24-205 (Colorado, USA), Sections 6-1-1702 & 1703; Resolution of the European Parliament of 20 October 2020 with Recommendations to the Commission on a Civil Liability Regime for Artificial Intelligence (Civil Liability for AI) (2020/2014(INL)) [2020] P9 TA (2020)0276, Article 8(2)(B).
[4] G20 AI Principles [2019], Sections 1.4 & 1.5.
[5] Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) [2021] COM (2021)206 final; AIA, Article 6 & Annex III; Civil Liability for AI, Article 4.
[6] S.B. 24-205 (Colorado, USA), Section 6-1-1701(9); AIA, Article 6.

https://doi.org/10.54934/ijlcw.v5i1.188

3.2.2 Absolute Liability

Building on the theory of liability without fault, absolute liability sets a more demanding framework. It holds individuals liable for harm caused by their AI systems, with no exceptions and no defenses. Unlike earlier doctrines, which allow defenses such as force majeure, absolute liability determines liability without regard to causation. [22]; [26]; [28] Given the autonomous and unpredictable behaviors of AI, this doctrine provides a strong incentive for stakeholders to comply with the highest safety and monitoring standards, because liability is established regardless of the precautions taken.

In the context of the EU AI Act, this strict liability approach finds partial alignment with the treatment of certain AI applications classified as prohibited practices,[7] which are deemed unlawful by default and banned outright from being placed on the market, put into service, or used within the EU.[8] In such cases, any deployment violating these prohibitions triggers inherent unlawfulness, leaving virtually no room for justification or defenses, thereby reinforcing a near-absolute form of accountability for harms arising from such uses.

This doctrine, however, deepens the concerns already linked to the previous doctrine, notably the risk of slowing innovation, and it also brings new difficulties with it. For example, liability without defenses may place an unfair burden on individuals who could not reasonably foresee or prevent the harm—especially where AI learning patterns behave unexpectedly, or where third-party interference, such as hacking, plays a role.

3.2.3 Vicarious Liability

Vicarious liability, another doctrine relevant to AI, extends the principles of liability without fault by imputing responsibility from an agent to a principal for wrongful acts committed within their relationship. [24] This doctrine serves to allocate liability to the party who is in the best position to control the conduct of the agent and bear the associated risks. However, in the context of contemporary agentic systems (commonly referred to as AI agents), the application of vicarious liability faces fundamental obstacles.

An agentic system is today understood as an autonomous AI system capable of pursuing complex, long-term goals with minimal human supervision, by perceiving its environment, reasoning over plans, and executing actions through iterative perception-reasoning-action loops. [39] The functionality, predictability, and risk profile of these agentic systems are thus heavily shaped by 1) the underlying model's architecture and training data (influencing generality, capabilities, and biases), 2) available

[7] AIA, Article 5.
[8] AIA, Article 5(1)(a-h).

https://doi.org/10.54934/ijlcw.v5i1.188

integrations and tools (determining the scope of possible actions and environmental interactions), and 3) the specific deployment/use context (e.g., virtual software agents versus physical robotics, or enterprise versus consumer applications), which collectively affect controllability and the applicability of regulatory frameworks.

Crucially, the absence of legal personality for AI systems—even highly agentic ones—precludes their classification as agents in the strict legal sense required for vicarious liability. [30]; [44] Moreover, as these agentic systems exhibit increasing levels of goal-directed autonomy, proactivity, adaptability, and capacity for multi-step decision-making, the degree of direct human control becomes progressively attenuated. This erosion of effective control fundamentally blurs—and often severs—the traditional agency relationship between a human principal and the AI agent, thereby weakening (or eliminating) the doctrinal justifications for imposing strict vicarious liability on a human principal (developer, deployer, or user) for the AI system's acts or omissions. [5]; [10]; [13]; [45]

3.3 Product Liability

Product liability is the last of the traditional theories. It places liability on manufacturers and other supply chain actors for harm caused by defective products. [23] Notably, this theory differs from fault-based liability because it usually imposes liability without fault on producers—meaning, a person only needs to show that a defect in the product caused the harm.

By its nature, this doctrine suggests itself as a suitable framework for AI systems.[5] However, applying it to AI raises immediate difficulties tied to the definition of a product, especially whether software and algorithms qualify as such. [4] Many jurisdictions continue to wrestle with this issue and often classify AI software as a service. As a result, liability analysis shifts back toward fault-based liability. [5]; [7]

More troubling still, the continuous learning and autonomous evolution of AI systems appear to undermine the element of defect. An AI system's conduct at the moment harm occurs may not reflect any design flaw that existed at the point of manufacture. [3]; [16] This gap weakens the core premise of product liability and allows manufacturers to rely on the state-of-the-art defense. [36] Against this backdrop, the intricate, multi-stakeholder chains involved in AI development make matters worse. [5]

4. MODERN THEORIES

In the face of escalating complexities surrounding the establishment of robust legal frameworks for assigning responsibility and attributing liability in cases of harm caused by AI systems, modern theories have emerged as innovative solutions. These theories seek to bridge the gaps inherent in

https://doi.org/10.54934/ijlcw.v5i1.188

traditional liability theories. This section explores each theory, highlighting their conceptual foundations, practical implications, and distinct benefits.

4.1 Electronic Personhood

One of the most forward-thinking theories proposes granting legal personality to AI systems, a concept often denoted as electronic personhood. This approach envisions AI entities as independent actors capable of bearing direct liability for their actions, much like corporations. [11]; [45]; [46] At its essence, this theory posits that sufficiently advanced AI—those exhibiting capabilities such as machine learning, autonomous decision-making, and behaviors that could meet the cognitive benchmarks for civil liability—should be afforded a legal personality. This status would empower AI systems to assume responsibility for their conduct, even allowing the seamless application of established liability theories, especially vicarious liability. [12]; [27]; [42]

Foremost among advantages of this theory is its ability to resolve persistent difficulties in establishing causation and fault within the opaque architectures of AI systems, due to their non-transparent decision-making processes. By attributing personhood to AI, this framework simplifies the legal process for aggrieved parties, enabling more straightforward claims for compensation, without the need to disentangle the roles of multiple stakeholders. [4]

Moreover, a pragmatic view treats personhood not as an inherent, metaphysical property to be discovered, but as a flexible bundle of rights and responsibilities that society can tailor to specific needs. This allows the creation of bespoke solutions, such as making certain persistent and agentic AI systems directly sanctionable or contractually addressable, even in cases where tracing responsibility to human owners or developers becomes practically impossible. [32] This theory also fosters innovation by clarifying liability boundaries, which can encourage investment in AI technologies without the fear of unpredictable personal liabilities for human actors.

Institutionally, this idea has attracted significant attention, as evidenced by the European Parliament's 2017 deliberations on creating a specialized legal status for autonomous robots as electronic persons responsible for damages they cause.[9] Although this proposal has not yet translated into binding legislation [31], it represents a pivotal exploratory step toward embedding AI within comprehensive liability structures. Overall, the electronic personhood theory's primary strength lies in its potential to create a balanced ecosystem where AI's benefits are maximized, while safeguarding individual rights through direct liability and, over the longer term, supporting a coherent allocation of rights, duties, and societal responsibilities across both human and artificial entities.

[9] European Parliament Resolution of 16 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics (Civil Law Rules on Robotics) (2015/2103(INL)) [2017] OJ C 252/239, Paragraph 59(f).

https://doi.org/10.54934/ijlcw.v5i1.188

4.2 Collective Compensation

A second modern theory of civil liability, identified in legal literature, is the collective compensation theory. This theory supports a framework of shared liability among multiple stakeholders for compensating victims of AI–related harms.[10] [5] It developed as a practical response to the problem of AI's inherent opacity [19]; [30]; [48] — an opacity that makes it difficult, and often impossible, for plaintiffs to trace harm through traditional chains of causation. In many scenarios, harm arises from the cumulative interplay of contributions across these stages, making it exceedingly difficult to isolate a single causal actor or pinpoint individual fault under conventional doctrine. As a result, conventional linear doctrine of civil liability frequently fails.

Notably, the strength of the collective compensation theory lies in its no-fault character. Rather than focusing on prolonged investigations into individual fault, it places victim relief at the center of the legal response. By creating a dedicated compensation fund—financed through mechanisms such as levies on AI transactions or mandatory insurance obligations imposed on AI's stakeholders—the theory ensures timely and reliable compensation for injured parties, even when no single liable actor can be clearly identified. [14]; [18]

To implement such a mechanism feasibly in practice, several concrete execution steps would be required. These include: 1) defining the scope of covered AI systems and harms through clear regulatory criteria (e.g., high-risk applications, as per emerging frameworks); 2) establishing a centralized or decentralized fund administration body, potentially overseen by a public authority or industry consortium; 3) designing equitable contribution formulas based on factors such as the inherent hazard level, revenue from AI activities, compute usage, or market share in the value chain; 4) mandating participation via legislation or binding industry standards, with penalties for non-compliance; 5) creating efficient claims processes, including simplified evidentiary thresholds focused on demonstrating harm causation by the AI system, rather than fault; and 6) incorporating periodic audits, risk-based adjustments to contributions, and mechanisms for subrogation or recourse against grossly negligent actors to preserve incentives for safety.

In fact, it reflects principles of restorative justice and social solidarity by treating AI-related harm as a shared risk, rather than an isolated personal failure. The theory allocates the financial consequences of AI's risks across all those who benefit from the technology, which helps prevent any one stakeholder from absorbing disproportionate or unforeseeable losses. [16] Put simply, by channeling residual risks into collective mechanisms, the theory softens the deterrent effect of strict personal liability. [30]

[10] Civil Law Rules on Robotics, Paragraphs 57, 58 & 59; Civil Liability for AI, Article 22; Proposal for a Directive of the European Parliament and of the Council on Adapting Non-Contractual Civil Liability Rules to Artificial Intelligence (AI Liability Directive) [2022] COM(2022)496 final, Paragraph 31, Article 5.

https://doi.org/10.54934/ijlcw.v5i1.188

4.3 Hybrid Liability

Another modern theory is hybrid liability. It presents a practical and carefully reasoned step forward in AI governance to remedy the shortcomings of civil liability theories by combining established doctrines with AI's characteristics. [36] This theory involves AI applications, from simple tools to highly autonomous systems, with their different levels of risk. As a result, it proposes a classified regime that adjusts liability based on factors such as opacity, supervision, autonomy, and the kind of harm. [30]

Structured in at least three tiers, the theory aligns liability with both the characteristics of the AI system and its operational context. At the first tier, low-risk AI systems with high transparency and strong human monitoring are subject to fault-based liability, where claimants must prove fault, damage, and causation. [19] For instance, this tier could apply to routine consumer AI applications, such as basic recommendation algorithms in spam filters in email services (mirroring the minimal or no-risk category in the EU AI Act, which imposes virtually no mandatory obligations[11]). Moving upward, the second tier applies strict liability to more complex, opaque, or autonomous systems. Here, the framework eases the plaintiff's burden by requiring proof of causation alone rather than fault—an approach that appears especially suitable in black box features where examining internal processes is often hidden. [8] Examples include high-risk deployments such as AI systems used in medical diagnostics (aligning with the high-risk category under the EU AI Act, which demands rigorous conformity assessments, risk management, and transparency measures to protect health, safety, and fundamental rights[12]). At the highest level, the third tier relies on absolute liability for AI systems that pose high-stakes or systemic risks, using no-fault doctrines to secure swift compensation for victims. [16]; [30] This tier would cover scenarios involving systemic or catastrophic potential, such as advanced autonomous vehicles in widespread deployment, where failures could affect large populations or public safety on a broad scale (echoing elements of the highest scrutiny in risk-based frameworks like the EU AI Act, while noting that truly unacceptable risks are typically prohibited outright rather than merely assigned absolute liability[13]).

The hybrid-liability tiers add value beyond merely mirroring the EU AI Act's risk categories by addressing two gaps the Act leaves open: (1) dynamic oversight adjustment — allowing liability to shift downward where effective human control is demonstrated, even for high-risk systems; and (2) jurisdictional adaptability — enabling non-EU regulators to adopt liability tiers calibrated to local legal contexts, without adopting the entire EU risk taxonomy. Thus, hybrid liability operationalizes risk classification into enforceable, context-sensitive legal standards.

[11] AIA, Article 95.
[12] AIA, Articles 6, 8–51, Annex III & Annex I.
[13] AIA, Article 5.

https://doi.org/10.54934/ijlcw.v5i1.188

The strengths of this hybrid approach are notable. Its adaptability allows regulators to respond differently across the AI applications, from routine consumer to critical infrastructure, such as medical AI. By adjusting liability standards to match risk levels, the framework helps avoid the chilling effects that uniform and overly punitive rules can create, while still motivating developers to invest in safer and more transparent systems. [19]

5. CHALLENGES AND SUGGESTIONS

Driven by practice, rather than theory alone, the move from traditional civil liability theories to modern theories in the field of AI reflects clear weaknesses in older frameworks. Those theories were built for human behavior, not for AI that acts autonomously, operates opaquely, and interacts across complex networks. As a result, they often struggle with core issues—establishing causation in black box algorithms, allocating fault among dispersed stakeholders, and delivering timely redress to victims.

By contrast, modern theories suggest a more workable path forward. They respond directly to AI's features and appear better suited to balancing fairness with practical enforcement. Notably, these approaches support innovation while still preserving accountability, making it more likely that society can benefit from AI without weakening legal protection for those harmed. Despite the persuasive benefits of modern theories, each one of them faces serious practical difficulties that call for workable responses. This section examines these challenges in detail and suggests strategies for addressing them in a realistic and effective manner.

5.1 Dynamic Legal Personhood

At the core of the granting legal personhood to AI systems lies a fundamental constraint: AI systems cannot hold independent financial resources to compensate for the damage they cause. [4] To deal with this problem, scholars suggest tools such as mandatory insurance schemes or dedicated compensation funds, specifically established for AI-related harm. [5] However, mandatory insurance itself faces significant practical obstacles, particularly during the early phase of implementation. Insurers are likely to exhibit considerable hesitation in underwriting civil liability coverage for AI systems, primarily because of the acute difficulty in reliably quantifying and pricing the novel, rapidly evolving, and highly context-dependent risks involved.

Drawing inspiration from ancient Roman law, some have proposed the idea of a Digital Peculium—a limited financial pool assigned to an AI system, similar to the resources once allocated to slaves, allowing AI to bear liability up to a defined ceiling. [9] This proposal, while creative, appears difficult to apply in practice, since compensation caps may leave victims without adequate redress, especially where harm is wide or the AI system becomes insolvent. [34] From a normative view, granting

https://doi.org/10.54934/ijlcw.v5i1.188

rights to AI without attaching corresponding duties risks reducing legal personhood to a symbolic label—one that lacks real legal substance. [15]; [19]; [47] Notably, there is also a real danger that corporations or individuals might use AI's legal personhood as a shield to avoid liability. [17]; [31]

Furthermore, the growing tendency to anthropomorphize AI systems—attributing human-like emotions, intentions, moral agency, or sentience to them—exacerbates these concerns. [33] While such features may enhance user engagement, they frequently lead to the neglect of robust functional governance, particularly in light of documented harms such as psychological distress, encouragement of self-harm, and privacy violations. [37] In the context of legal personhood debates, excessive anthropomorphization risks misdirecting normative arguments toward granting personhood on illusory, human-like grounds rather than pragmatic accountability needs, potentially obscuring corporate liability and complicating the allocation of responsibility for AI-caused harms.

Responding to these concerns, one can suggest a model of dynamic legal personhood that adjusts the scope of legal status according to an AI system's autonomy, adaptability, decision-making capacity, and degree of human control. Under this framework, highly autonomous systems would carry liability of their own, while systems operating under wide human oversight would keep liability for human actors. By translating this idea into practice, global standards could be developed to assess AI autonomy, using criteria such as algorithmic complexity, decision-making independence, and the extent of human intervention. Although the question arises as to which institution would be competent to develop and implement such standards, existing and emerging international bodies could play leading or collaborative roles in this process. For instance, established organizations such as the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC)—through their joint technical committee on AI (ISO/IEC JTC 1/SC 42)—already produce widely recognized technical standards for AI management, risk assessment, and governance frameworks that could be extended or adapted to include autonomy evaluation metrics.[14] Complementary efforts by intergovernmental entities like the Organisation for Economic Co-operation and Development (with its updated AI Principles) and the United Nations Educational, Scientific and Cultural Organization (UNESCO)[15] provide normative foundations and multi-stakeholder processes suitable for building consensus on autonomy criteria. More recently proposed or evolving mechanisms, including networks of AI Safety Institutes, the United Nations' initiatives such as the Independent International Scientific Panel on Artificial Intelligence, or dedicated

[14] For example, ISO/IEC 22989:2022 Artificial intelligence concepts and terminology; ISO/IEC 23053:2022 Framework for AI systems using machine learning (ML); ISO/IEC 23894:2023 Artificial intelligence — Guidance on risk management; ISO/IEC 5338:2023 AI system life cycle processes; ISO/IEC TR 24027:2021 Bias in AI systems and AI aided decision making.

[15] UNESCO, Recommendation on the Ethics of Artificial Intelligence (the General Conference 23 November 2021, document SHS/BIO/REC-AIETHICS/2021).

https://doi.org/10.54934/ijlcw.v5i1.188

multi-stakeholder forums, could contribute to harmonized methodologies, ensuring both technical rigor and broad legitimacy across jurisdictions.

For AI systems granted complete legal personality and operating with little or no human supervision, national or international compensation funds would ensure effective victim compensation. At the same time, stakeholders involved with AI systems—where human supervision remains significant and legal personality is minimal—should remain liable under vicarious liability theory for the damage those systems cause.

5.2 Redesigning Collective Compensation

The collective compensation theory demands careful legal planning to prevent moral hazard, and to ensure it does not, in practice, protect negligent actors from liability. [36] The theory invites a setting in which corporations may neglect essential safety standards, trusting that liability will disperse across a collective fund. [49] This result weakens deterrence and erodes the principle of corrective justice—which expects wrongdoers to bear the costs of the harm they cause.[16]

Protecting against these risks requires a framework that firmly balances risk-sharing with individual liability. One practical approach would allow legislators to structure contribution schemes so they reflect not only a firm's market share or role in AI development, but also its safety record. As a result, actors that behave recklessly would carry heavier financial obligations, rather than hiding behind the collective.

Safety principles can be drawn from established sources, such as the EU AI Act, which adopts a risk-based approach emphasizing human oversight, technical robustness and safety, transparency, privacy and data governance, fairness, societal well-being, and accountability.[17] Alongside this, lawmakers could codify binding safety standards—drawing on internationally recognized frameworks, such as ISO/IEC 42001,[18] which focus on organizational governance, risk assessment, responsibility allocation, human oversight, technical robustness, incident management, and continuous improvement—rather than attempting to impose a single, rigidly fixed technical standard.

Such standards provide practical, evaluative tools that remain adaptable to the diverse forms of AI systems (varying in model types, training strategies, data regimes, and integration architectures) and their task- and context-specific effectiveness. An overly prescriptive, uniform technical standard risks becoming quickly outdated or stifling innovation, whereas safety-oriented standards offer certainty at the organizational level and can be legally operationalized—for instance, when determining whether an AI

[16] An age-old maxim says, "everyone is responsible, nobody is responsible."
[17] AIA, Recital 27, Articles 9, 10, 13, 14, 15, 16, & 43.
[18] International Organization for Standardization, ISO/IEC 42001:2023 Artificial Intelligence — Management System (2023).

https://doi.org/10.54934/ijlcw.v5i1.188

operator has exercised due diligence, implemented sound governance, or adopted reasonable safety measures in a given use case. Additional guidance may also be informed by resources from organizations such as the OECD [40] and non-profits, including the Future of Life Institute [20], whose AI Safety Index evaluations highlight varying levels of safety practices across leading developers and underscore the value of independent assessments.

Importantly, the collective mechanism should not operate as an absolute liability shield. Where clear and convincing evidence demonstrates gross negligence, willful misconduct, or severe breach of mandatory safety obligations by a specific actor, the regime must preserve the possibility of supplementary or escalated individual accountability. This may take the form of an increased contribution coefficient applied retroactively to that participant, exclusion from certain protections of the fund, or — in particularly egregious cases — the activation of parallel direct civil liability outside the collective framework. Such a dual-track structure safeguards the deterrent function of liability law and prevents strategic reliance on the fund as a means of evading personal responsibility.

Notably, positive incentives may also play a role. Offering tax advantages or reduced fund contributions to companies that submit to regular third-party safety audits likely encourages liable conduct without stifling innovation. By embedding such safeguards, the collective compensation theory shifts from a narrow compensatory tool into a regulatory-liability model.

5.3 Thresholds Based on Objective Metrics

The hybrid liability theory also faces a practical issue that, in fact, appears manageable if addressed with care. That is, setting clear thresholds for moving between liability tiers—starting with low-risk systems under strong human oversight, extending to complex systems governed by strict liability, and reaching critical systems subject to absolute liability. These boundaries often remain unclear and depend on subjective judgments, which can draw courts into lengthy disputes and, as a result, delay compensation for victims.

By adopting a uniform method based on both quantitative and qualitative criteria, lawmakers can reduce this uncertainty. In particular, thresholds could incorporate assessments of three core dimensions frequently highlighted in AI safety frameworks: 1) the degree of autonomy (i.e., the system's capacity for independent, multi-step action with limited or no human intervention), 2) task competence (i.e., demonstrated performance level across designated functions or domains), and 3) scope of generality (i.e., breadth of applicability, adaptability to untrained tasks, and transfer learning ability).[19] Such

[19] The Society of Automotive Engineers (SAE) outlines six levels of driving automation, from Level 0 (fully manual) to Level 5 (fully autonomous). These levels have been adopted by the U.S. Department of Transportation.; ISO/IEC. (2023). ISO/IEC 23894: Artificial intelligence — Guidance on risk management. International Organization for Standardization/International Electrotechnical Commission.

multidimensional evaluation aligns closely with emerging proposals that differentiate risk tiers according to combinations of strength in autonomy, competence, and generality, thereby enabling more precise and defensible categorization of systems, from narrowly scoped tools to highly capable, open-ended ones. [21] To keep these standards current, an interdisciplinary expert committee made up of legal scholars and AI engineers could periodically review and revise the thresholds, ensuring they reflect technological developments and newly identified risks.

To mitigate the substantial risk of strategic under-classification — whereby developers might deliberately design or present systems in ways that minimize apparent risk exposure and thereby reduce liability obligations — independent third-party conformity assessment should be mandatory for all systems proposed for mid- or high-liability tiers. Such assessments, conducted by accredited bodies, would prevent self-serving misclassification and preserve incentives for advancing genuinely capable systems in high-stakes domains. Without robust safeguards against this form of regulatory gaming, developers may gravitate toward conservative, low-autonomy designs that trigger lighter liability regimes, even when technically superior and socially valuable applications exist in medicine, autonomous transport, or critical infrastructure. This distortion could suppress innovation precisely in those areas where AI has the greatest potential to deliver societal benefit, ultimately skewing market outcomes toward safer-but-less-impactful technologies and undermining the overall objectives of balanced risk allocation and technological progress.

5.4 Conditional Pluralism in Civil Liability Theories for AI

The foregoing analysis might suggest that it ultimately prefers hybrid liability as the most balanced and practicable framework. However, just as no single traditional doctrine can comprehensively address all AI-related harm scenarios, none of the modern proposals should be treated as the exclusive or universally superior solution.

Following a pluralist yet disciplined methodology, the present study submits that each theory contains partial but genuine truths, and offers valuable tools for distinct classes of problems. Theories are not final destinations, but instrumental means toward the overriding telos of civil liability in the AI era: achieving justice between victim protection, fair risk allocation among stakeholders, and the socially desirable advancement of beneficial technology.

Accordingly, the preferable path forward is not to select one modern theory and discard the others, but to construct a conditional, context-sensitive liability map that activates different elements, or combinations of elements, depending on clearly specified situational variables. The most relevant variables include at minimum: 1) degree of system autonomy (narrow-task tool vs. long-horizon agentic system), 2) opacity of the decision-making process, 3) severity and scale of potential harm (individual vs.

https://doi.org/10.54934/ijlcw.v5i1.188

catastrophic), 4) availability and effectiveness of human oversight, 5) structure of the value chain (single developer vs. multi-layered supply chain involving foundation model providers, fine-tuners, deployers, cloud infrastructure, end-users), and 6) existence and enforceability of mandatory insurance or compensation funds.

Under this conditional-pluralist framework: 1) Electronic personhood (possibly dynamic / limited) gains strongest normative weight in long-lived, highly autonomous, minimally supervised systems — especially open-source or decentralized agentic AI — where tracing back effective human control becomes practically impossible, and a stable accountability locus is urgently needed; 2) Collective no-fault compensation mechanisms should be activated as the default or residual layer whenever multi-stakeholder opacity prevents reliable individual causation tracing, or when swift victim relief is regarded as more important than pinpointing fault; and 3) Hybrid liability constitutes the backbone regime for the largest share of currently foreseeable high-risk applications because it offers proportionality and preserves deterrence incentives while easing the victim's burden in opaque settings.

Such conditional pluralism avoids both monistic dogmatism and undisciplined relativism. It requires lawmakers and courts to transparently and revisibly articulate the precise triggering conditions and priority rules among the available tools. Doing so demands intellectual humility, analytical rigor, and moral responsibility. By embracing this disciplined pluralist stance, the legal system can better track technological reality, respond adaptively to new capabilities and risks, and remain oriented toward the overriding goal of justice rather than the idolization of any single doctrinal construct.

6. CONCLUSION

The rapid integration of AI into society signals a clear shift—one that brings serious legal concerns with it. Civil liability's traditional theories, built around human conduct, are poorly suited to AI systems that act autonomously, operate opaquely, and function within networks of actors. These theories often fail to trace causation inside black box algorithms, assign fault among scattered stakeholders, or deliver timely remedies to injured parties. As a result, gaps emerge in liability, victim compensation, and the careful balance between encouraging innovation and protecting safety. Over time, these weaknesses risk eroding public trust in AI technologies, which likely slows adoption and chills the innovation many sectors rely on. Moving away from outdated liability theories towards AI-specific theories is therefore necessary—these newer approaches seek to align liability with AI's distinctive features while supporting technological development.

https://doi.org/10.54934/ijlcw.v5i1.188

Modern theories—electronic personhood, collective compensation, and hybrid liability—present practical and forward-looking responses to the problems of earlier theories. Yet, none of them should be elevated to the status of a sole, universally applicable paradigm. Instead, the most defensible position is a conditional, pluralist framework that accords each approach its appropriate domain and weight according to context-specific criteria (autonomy, opacity, harm scale, oversight feasibility, value-chain complexity). Such disciplined pluralism — while intellectually more demanding — better honors the complexity of AI-mediated harm, avoids the hubris of monistic solutions, and keeps the ultimate telos of civil liability firmly in view.

At the core, developing such a conditional-pluralist liability framework through global cooperation, shared standards, and interdisciplinary research stands as a legal necessity. By refining and selectively combining dynamic legal personhood, collective compensation, and hybrid liability under clearly stated conditions and priority rules, legal systems can adapt alongside AI, rather than lag behind it. This approach offers a realistically flexible yet normatively responsible path forward—one where responsibility keeps pace with technology. Supported by sustained theorizing and intellectual investment, such a framework allows technology and justice to grow together, protecting individual rights while opening space for AI's considerable potential.

https://doi.org/10.54934/ijlcw.v5i1.188

REFERENCES

[1] Al-Dulaimi, A. O. M., & Mohammed, M. A. A. W. (2025). Legal responsibility for errors caused by artificial intelligence (AI) in the public sector. International Journal of Law and Management, 1–28.

[2] Asaro, P. M. (2012). A Body to Kick, but Still No Soul to Damn : Legal Perspectives on Robotics. In P. Lin, K. Abney, & G. A. Bekey (Eds.), Robot Ethics: The Ethical and Social Implications of Robotics (Vol. 900, Number 400, pp. 169–186). MIT Press.

[3] Barfield, W. (2018). Liability for autonomous and artificially intelligent robots. Paladyn, 9(1), 193–203. https://doi.org/10.1515/PJBR-2018-0018/MACHINEREADABLECITATION/RIS

[4] Barfield, W., & Pagallo, U. (2020). Advanced introduction to law and artificial intelligence. Edward Elgar Publishing.

[5] Benhamou, Y., & Ferland, J. (2020). Artificial Intelligence & Damages: Assessing Liability and Calculating the Damages. Leading Legal Disruption: Artificial Intelligence and a Toolkit for Lawyers and the Law, Forthcoming, 1–20. https://doi.org/10.2307/J.CTT207G8BT.22

[6] Bereska, L., & Gavves, E. (2024). Mechanistic Interpretability for AI Safety - A Review. ArXiv, 1–55.

[7] Boch, A., Hohma, E., & Trauth, R. (2022). Towards an Accountability Framework for AI: Ethical and Legal Considerations. Institute for Ethics in Artificial Intelligence, 1–12. https://doi.org/10.13140/RG.2.2.10231.50086

[8] Bottomley, D., & Thaldar, D. (2023). Liability for harm caused by AI in healthcare: an overview of the core legal concepts. Frontiers in Pharmacology, 14, 1–9. https://doi.org/10.3389/FPHAR.2023.1297353/BIBTEX

[9] Brennan, D. (2024). Proposals for the Rights and Liabilities of Artificial Intelligence through Electronic Personhood: Conscious Machines, Hermit Crabs, and the Distant Past. SSRN, 1–49.

[10] Bruyne, J. De, & Tanghe, J. (2017). Liability for damage caused by autonomous vehicles : a Belgian perspective. Journal of European Tort Law, 324.

[11] Bruyne, J. De, & Vanleenhove, C. (2021). Artificial Intelligence and the Law. Intersentia.

[12] Calverley, D. J. (2008). Imagining a non-biological machine as a legal person. AI and Society, 22(4), 523–537. https://doi.org/10.1007/S00146-007-0092-7

[13] Čerka, P., Grigiene, J., & Sirbikyte, G. (2015). Liability for damages caused by artificial intelligence. Computer Law & Security Review, 31(3), 376–389. https://doi.org/10.1016/J.CLSR.2015.03.008

[14] Corrales Compagnucci, M., Fenwick, M., & Forgó, N. (2018). Robotics, AI and the future of law. Springer.

[15] Cui, Y. (2024). Blue Book on AI and Rule of Law in the World (2021) (Artificial Intelligence and the Rule of Law). Springer.

[16] Custers, B., & Fosch-Villaronga, E. (2022). Law and Artificial Law And Artificial Intelligence: Regulating AI And Applying AI In Legal Practice. Springer.

[17] De Bruyne, J., Van Gool, E., & Gils, T. (2021). The Cambridge Handbook of Private Law and Artificial Intelligence Edited by The Cambridge Handbook of Private Law and Artificial Intelligence. Cambridge University Press.

[18] Determann, L. (2024). Determann's field guide to artificial intelligence law : international corporate compliance. Edward Elgar Publishing Limited.

[19] DiMatteo, L. A., Poncibò, C., & Cannarsa, M. (Editors). (2022). The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics . Cambridge University Press .

[20] Future of Life Institute. (2025a). AI Safety Index. https://futureoflife.org/ai-safety-index-summer-2025/

[21] Future of Life Institute. (2025b). Safety Standards Delivering Controllable and Beneficial AI Tools.

[22] Garner, B. A. (2004). Black's Law Dictionary (Eighth). Thomson West.

[23] Geistfeld, M. (2012). Products liability law. Wolter Kluwer Law & Business.

[24] Gray, A. (2018). Vicarious Liability. Hart Publishing. https://doi.org/10.5040/9781509920266

[25] Gray, A. (2021). The evolution from strict liability to fault in the law of torts. In The Evolution from Strict Liability to Fault in the Law of Torts. Bloomsbury Publishing Plc. https://doi.org/10.5040/9781509941025

[26] Gupta, A. (2025). Framework for Addressing Liability and Accountability Challenges due to Artificial Intelligence Agents. Indian Journal of Integrated Research in Law, 7, 910–926.

[27] Hall, J. S. (2007). Beyond AI : creating the conscience of the machine. Prometheus Books.

[28] Jefferies, C., & Klar, L. N. (2023). Tort Law (7th ed.). Carswell.

[29] Junod, V. (2019). Liability for damages caused by AI in medicine : Progress Needed. Journée de Responsabilité Civile 2018, Collection Genevoise, 119–149.

[30] Kerrigan, C. (2022). Artificial intelligence : law and regulation. Edward Elgar Publishing Limited.

[31] Khan, A. (2023). Evolution of Artificial Intelligence: A Transformative Role in the Judicial System. SSRN Electronic Journal. https://doi.org/10.2139/SSRN.4548724

[32] Leibo, J. Z., Vezhnevets, A. S., Cunningham, W. A., & Bileschi, S. M. (2025). A Pragmatic View of AI Personhood. ArXiv, 1–31.

[33] Li, M., & Suh, A. (2021). Machinelike or humanlike? A literature review of anthropomorphism in AI-enabled technology. In 54th Hawaii International Conference on System Sciences, 4053–4062.

[34] Lior, A. (2020). AI Entities as AI Agents: Artificial Intelligence Liability and the AI Respondeat Superior Analogy. Mitchell Hamline Law Review, 46(5).

[35] Liu, H. W., Lin, C. F., & Chen, Y. J. (2019). Beyond State v Loomis: artificial intelligence, government algorithmization and accountability. International Journal of Law and Information Technology, 27(2), 122–141. https://doi.org/10.1093/IJLIT/EAZ001

[36] Lohsse, S., Schulze, R., & Staudenmayer, D. (2019). Liability for Artificial Intelligence and the Internet of Things. In S. Lohsse, R. Schulze, & D. Staudenmayer (Eds.), Liability for Artificial Intelligence and the Internet of Things. Bloomsbury Publishing. https://doi.org/10.5040/9781509925841

[37] Maeda, T., & Quan-Haase, A. (2024). When human-AI interactions become parasocial: Agency and anthropomorphism in affective design. Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency, 1068–1077.

[38] Mantelero, A. (2018). AI and Big Data: A blueprint for a human rights, social and ethical impact assessment. Computer Law & Security Review, 34(4), 754–772. https://doi.org/10.1016/J.CLSR.2018.05.017

[39] Nisa, U., Shirazi, M., Saip, M. A., & Pozi, M. S. M. (2025). Agentic AI: The age of reasoning—A review. Journal of Automation and Intelligence, 1–21.

[40] Organisation for Economic Co-operation and Development. (2024). OECD Principles on Artificial Intelligence. https://www.oecd.org/en/topics/ai-principles.html

[41] Pagallo, U. (2012). Three roads to complexity, AI and the law of robots: On crimes, contracts, and torts. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 7639 LNAI, 48–60. https://doi.org/10.1007/978-3-642-35731-2_3

[42] Pagallo, U. (2013). The Laws of Robots. Springer Netherlands.

[43] Pagallo, U. (2017). From Automation to Autonomous Systems: A Legal Phenomenology with Problems of Accountability. In C. Sierra (Ed.), International Joint Conferences on Artificial Intelligence Organization (IJCAI-17).

[44] Rachum-Twaig, O. (2019). Whose Robot Is It Anyway?: Liability for Artificial-Intelligence-Based Robots. University of Illinois Law Review, 2020, 1–40.

[45] Sullivan, H. R., & Schweikart, S. J. (2019). Are Current Tort Liability Doctrines Adequate for Addressing Injury Caused by AI? AMA Journal of Ethics, 21(2), 160–166. https://doi.org/10.1001/AMAJETHICS.2019.160

[46] Teubner, G. (2006). Rights of Non-humans? Electronic Agents and Animals as New Actors in Politics and Law. Journal of Law and Society, 33(4), 497–521. https://doi.org/10.1111/J.1467-6478.2006.00368.X

[47] Yalman, B. (2024). Electronic Personhood: A Compact Analysis of Legal Personality for Artificial Intelligence. Ex/Ante, 2024(1), 3–13. https://doi.org/10.3256/978-3-03929-059-8_02

[48] Yas, N., Qaruty, R. Al, Hadi, S. A., & AlAdeedi, A. (2023). Civil Liability and Damage Arising from Artificial Intelligence. Migration Letters, 20(5), 430–446. https://doi.org/10.59670/ML.V20I5.3554

[49] Yi, C. (2022). Artificial intelligence and international law. Springer.

[50] Youvan, D. (2024). Cognition, Complexity, and Emergence: Exploring the Boundary Between Deterministic and Probabilistic Systems. Researchgate, 1–33.

ABOUT THE AUTHORS

Abbas Mirshekari – PhD. Associate Professor, Private and Islamic Law, University of Tehran, Tehran, Iran. Deputy in the Private and Islamic Law Department at the University of Tehran. Research record reflects expertise in various areas of law, including private law, civil liability law, contract law, civil registration law, comparative law, human rights, and Personal law. He has published over 50 scientific research articles and authored 7 research and practical books in these fields.
e-mail: mirshekariabbas1@ut.ac.ir
ORCID ID: https://orcid.org/0000-0002-4927-1720

Jamshid Zargari – Master of Laws from the University of Judicial Sciences and Administrative Services, Tehran, Iran. Visiting Lecturer at University of Applied Science and Technology, Tehran, Iran. Attorney at Law in the Iran Central Bar Association, Tehran, Iran. Contributions to the field extend to editorial roles, including serving as an editor for legal publications and a reviewer for legal journals.
e-mail: jamshidzargari94@gmail.com
ORCID ID: https://orcid.org/0000-0002-8004-2969

ABOUT THIS ARTICLE

Conflict of Interests: The authors declare no conflicting interests

CRediT Authorship Contribution Statement:

Abbas Mirshekari contributed to the supervision, project administration, and review & editing of the manuscript.
Jamshid Zargari was responsible for the conceptualization, methodology, investigation, writing of the original draft, and validation of the work.

Data Availability: This manuscript is a doctrinal and comparative-analytical study based exclusively on publicly available legal sources, including statutes, case law, official reports, and peer-reviewed academic literature. All sources cited in this paper are listed in the References section.

AI-use Disclosure: During the preparation of this manuscript, the authors used DeepL solely for language polishing and grammatical refinement. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the final version of the manuscript. No AI tools were used for idea generation, data analysis, or substantive intellectual work.

LA RESPONSABILIDAD CIVIL EN LA ERA DE LA INTELIGENCIA ARTIFICIAL: DE LAS TEORÍAS TRADICIONALES A LAS MODERNAS

RESUMEN

La integración de la inteligencia artificial (IA) en las estructuras sociales ha puesto de manifiesto deficiencias fundamentales en las teorías tradicionales de responsabilidad civil, las cuales presentan dificultades para abordar la actuación autónoma, los procesos opacos y la responsabilidad fragmentada entre múltiples partes interesadas. Este estudio emplea un método doctrinal y analítico-comparativo para evaluar los marcos modernos de responsabilidad —tales como la personalidad electrónica, la indemnización colectiva y la responsabilidad híbrida— frente a las limitaciones de los enfoques convencionales. Los resultados indican que la personalidad jurídica dinámica, la responsabilidad por actos de terceros (responsabilidad vicaria) atribuible a las partes interesadas, los umbrales de riesgo objetivo y los mecanismos de supervisión experta ofrecen vías coherentes para abordar la causalidad difusa y la falta de transparencia de los algoritmos. La principal contribución es un marco de pluralismo condicional que equilibra la protección de la víctima, la distribución equitativa del riesgo y el progreso tecnológico, proporcionando herramientas prácticas para gestionar los litigios relacionados con la IA en un panorama jurídico en constante evolución.

Palabras clave: Inteligencia artificial, responsabilidad civil, indemnización colectiva, personalidad electrónica, responsabilidad híbrida

人工智能时代的民事责任：从传统理论到现代理论

摘要

人工智能融入社会结构，暴露了传统民事责任理论的根本缺陷，使其难以应对自主行为、不透明的运作过程以及涉及多方主体的责任分散问题。本研究采用法理分析与比较分析相结合的方法，针对传统责任模式的局限性，评估了包括“电子人格”、集体赔偿及混合责任在内的现代责任框架。研究结果表明，动态法律人格、利益相关者的替代责任、客观风险阈值以及专家监督机制，为解决因果关系不明和算法不透明问题提供了切实可行的路径。本研究的主要贡献在于构建了一个“有条件的多元主义”框架，该框架在受害者保护、公平的风险分担与技术进步之间实现了平衡，为在不断演变的法律环境中处理人工智能相关纠纷提供了实用工具。

关键词：人工智能，民事责任，集体赔偿，电子人格，混合责任

https://doi.org/10.54934/ijlcw.v5i1.188