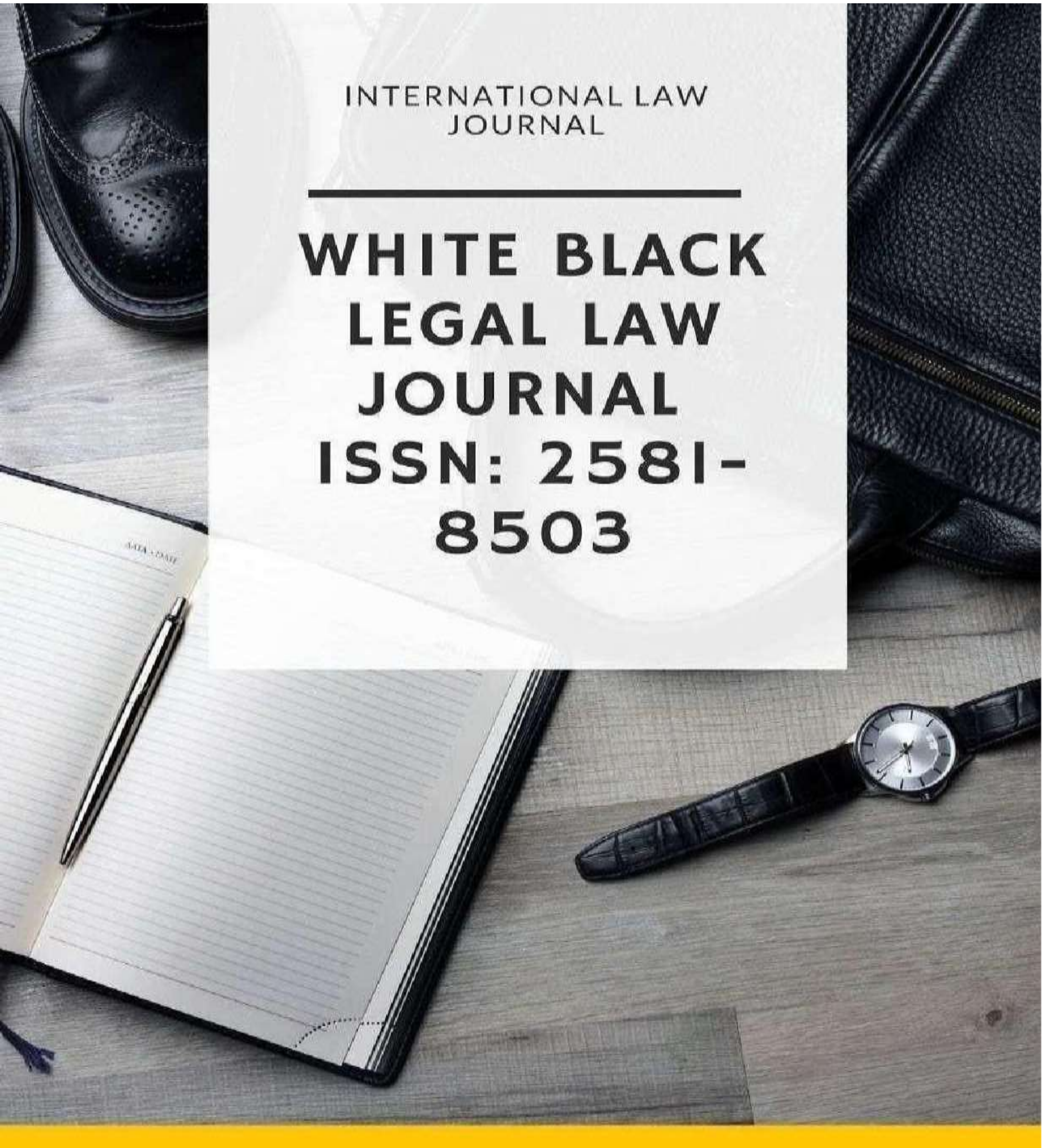

INTERNATIONAL LAW JOURNAL

WHITE BLACK LEGAL LAW JOURNAL ISSN: 2581-8503

Peer - Reviewed & Refereed Journal

The Law Journal strives to provide a platform for discussion of International as well as National Developments in the Field of Law.

WWW.WHITEBLACKLEGAL.CO.IN

DISCLAIMER

No part of this publication may be reproduced or copied in any form by any means without prior written permission of Editor-in-chief of White Black Legal – The Law Journal. The Editorial Team of White Black Legal holds the copyright to all articles contributed to this publication. The views expressed in this publication are purely personal opinions of the authors and do not reflect the views of the Editorial Team of White Black Legal. Though all efforts are made to ensure the accuracy and correctness of the information published, White Black Legal shall not be responsible for any errors caused due to oversight or otherwise.

EDITORIAL TEAM

Raju Narayana Swamy (IAS) Indian Administrative Service officer

Dr. Raju Narayana Swamy popularly known as Kerala's Anti-Corruption Crusader is the All India Topper of the 1991 batch of the IAS and is currently posted as Principal Secretary to the Government of Kerala. He has earned many accolades as he hit against the political-bureaucrat corruption nexus in India. Dr Swamy holds a B.Tech in Computer Science and Engineering from the IIT Madras and a Ph. D. in Cyber Law from Gujarat National Law University. He also has an LLM (Pro) (with specialization in IPR) as well as three PG Diplomas from the National Law University, Delhi- one in Urban Environmental Management and Law, another in Environmental Law and Policy and a third one in Tourism and Environmental Law. He also holds a post-graduate diploma in IPR from the National Law School, Bengaluru and a professional diploma in Public Procurement from the World Bank.

Dr. R. K. Upadhyay

Dr. R. K. Upadhyay is Registrar, University of Kota (Raj.), Dr Upadhyay obtained LLB, LLM degrees from Banaras Hindu University & PHD from university of Kota. He has succesfully completed UGC sponsored M.R.P for the work in the Ares of the various prisoners reforms in the state of the Rajasthan.

Dr. Rinu Saraswat

Associate Professor at School of Law, Apex University, Jaipur, M.A, LL.M, PH.D,

Dr. Rinu have 5 yrs of teaching experience in renowned institutions like Jagannath University and Apex University. Participated in more than 20 national and international seminars and conferences and 5 workshops and training programmes.

Dr. Nitesh Saraswat

E.MBA, LL.M, PH.D, PGDSAPM
Currently working as Assistant Professor at Law Centre II, Faculty of Law, University of Delhi. Dr. Nitesh have 14 years of Teaching, Administrative and research experience in Renowned Institutions like Amity University, Tata Institute of Social Sciences, Jai Narain Vyas University Jodhpur, Jagannath University and Nirma University.
More than 25 Publications in renowned National and International Journals and has authored a Text book on CR.P.C and Juvenile Delinquency law.

Subhrajit Chanda

BBA. LL.B. (Hons.) (Amity University, Rajasthan); LL. M. (UPES, Dehradun) (Nottingham Trent University, UK); PH.D. Candidate (G.D. Goenka University)

Subhrajit did his LL.M. in Sports Law, from Nottingham Trent University of United Kingdoms, with international scholarship provided by university; he has also completed another LL.M. in Energy Law from University of Petroleum and Energy Studies, India. He did his B.B.A.LL.B. (Hons.) focussing on International Trade Law.

ABOUT US

WHITE BLACK LEGAL is an open access, peer-reviewed and refereed journal provide dedicated to express views on topical legal issues, thereby generating a cross current of ideas on emerging matters. This platform shall also ignite the initiative and desire of young law students to contribute in the field of law. The erudite response of legal luminaries shall be solicited to enable readers to explore challenges that lie before law makers, lawyers and the society at large, in the event of the ever changing social, economic and technological scenario.

With this thought, we hereby present to you

LEGAL CHALLENGES IN ASSIGNING CORPORATE LIABILITY FOR CAUSED HARMS

AUTHORED BY - MR. KIRUBAKARN S
Student – 2nd Year LLB,
Christ University, Bangalore

ABSTRACT

The growing use of artificial intelligence across sectors such as health care, finance, transportation, and employment raises difficult legal questions about who should be held responsible when AI systems cause injury or loss. Unlike harms traceable to specific human actions, AI-related harms often arise from complex interactions between data, model design, deployment choices, and operational contexts. This chapter examines how those factors disrupt familiar liability concepts—causation, foreseeability, fault, and agency—and create practical obstacles for victims seeking redress. It explores the problem of dispersed responsibility across an AI value chain (developers, data providers, integrators, deployers and operators), the evidentiary challenges posed by opaque "black-box" models and trade-secrecy claims, and the limits of doctrines like vicarious liability or respondeat superior in multi-party AI ecosystems. The analysis compares emerging policy responses—from proposals that treat certain AI software as products subject to strict liability to approaches that emphasize risk-calibrated duties of care and enhanced oversight for deployers—and discusses the procedural and institutional reforms needed to make liability regimes effective. The chapter argues for a pragmatic, mixed strategy that combines targeted strict liability for high-risk applications, clearer allocation of responsibility toward parties with operational control, improved discovery and technical expertise in adjudication, and regulatory measures that incentivize safety without unduly hampering innovation. Such a hybrid framework, the chapter contends, offers a balanced way to protect victims and promote responsible AI deployment.

Keywords: AI liability, causation, foreseeability, strict liability, vicarious liability, black-box models, deployer responsibility.

1.1 Introduction

The rapid advancement and deployment of artificial intelligence systems across critical sectors, including healthcare, finance, employment, and transportation—has fundamentally altered the landscape of corporate decision-making.[1] As artificial intelligence increasingly influences consequential determinations affecting individual lives and property, a pressing legal question has emerged: when artificial intelligence systems cause harm, who bears legal responsibility?[2] Traditional models of corporate liability, established over centuries of common law development and codified in statutes, operate upon foundational principles grounded in human agency, culpable intent, and identifiable causation.[3] Yet artificial intelligence systems, characterized by their autonomy, complexity, and opacity, challenge these established frameworks in unprecedented ways. [4]

The assignment of corporate liability for AI-caused harms sits at the intersection of multiple legal disciplines—tort law, contract law, consumer protection law, criminal law, and emerging regulatory frameworks—creating a labyrinthine environment where liability pathways remain ambiguous.[5] In jurisdictions including India and developed countries such as the United States and European Union member states, the absence of explicit AI-specific liability regimes has compelled courts and regulators to apply legacy legal frameworks designed for mechanical systems and human actors to non-human, algorithmic agents.[6] This adaptation has proven inadequate, leaving lacunae in both the legal framework and practical enforcement mechanisms.[7] The European Commission's proposed Artificial Intelligence Liability Directive

[1] Iwona Gredka-Ligarska, *Employer as an AI System Operator and Tortious Liability for Damage Caused by AI Systems: European and US Perspectives*, 12 CHINESE J. COMP. L. cxae015 (2024), https://academic.oup.com/cjcl/article/doi/10.1093/cjcl/cxae015/7889035.

[2] Joyeeta Banerjee, *The Rise of Artificial Intelligence in Corporate Accountability: Legal Implications for Corporate Governance in India*, SSRN (May 23, 2025), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5291831.

[3] Sarah McAtominey, Céline Moille, Gretchen Scott & Genevieve Watt, *EU Updates its Product Liability Regime*, GOODWIN L.M. (Feb. 17, 2025), https://www.goodwinlaw.com/en/insights/publications/2025/02/alerts-practices-aiml-eu-updates-its-product-liability-regime.

[4] Enrico Moch, *Autonomous Systems (AI) and Criminal Imputability: Challenges for Modern Law*, 12 INT'L J. RES. & REV. 287 (Oct. 2025), https://www.ijrrjournal.com/IJRR_Vol.12_Issue.10_October2025/IJRR29.pdf.

[5] Dr. Subholaxmi Mukherjee, *Algorithmic Bias and Discrimination: Legal Accountability of AI Systems*, 13 INT'L J. INNOVATIVE RES. MULTIDISCIPLINARY PHYS. SCI. 1 (July-Aug. 2025), https://www.ijirmps.org/papers/2025/4/232659.pdf.

[6] Mary Duffourc, *Decoding U.S. Tort Liability in Healthcare's Black-Box AI Era*, 27 STAN. TECH. L. REV. 1 (2024), https://law.stanford.edu/wp-content/uploads/2024/02/Publish_27-STLR-1-2024_Decoding-U.S.-Tort-Liability-in-Healthcares-Black-Box-AI-Era.pdf.

[7] The Rise Of Artificial Intelligence In Corporate Accountability*, IJCRT (2025), https://ijcrt.org/papers/IJCRT2506028.pdf.

and the revised Product Liability Directive (effective December 2024) represent attempts to modernize civil liability rules, yet significant divergence remains across jurisdictions. [8]

This chapter undertakes a systematic analysis of the legal challenges inherent in assigning corporate liability for harms caused by artificial intelligence systems. It examines how traditional liability doctrines, including negligence, strict liability, and vicarious liability—function when applied to non-human actors operating with degrees of autonomy and opacity unknown to classical legal theory.[9] The chapter further explores the distinctions between AI-caused harms and traditional corporate liability scenarios, investigates the conceptual and practical difficulties of identifying the perpetrator in harm causation chains, analyzes the doctrine of foreseeability in the context of algorithmic decision-making, and evaluates the respective roles of intent, negligence, and strict liability in emerging AI liability frameworks.[10] By comparing regulatory approaches across India and developed jurisdictions, this chapter aims to illuminate the critical gaps in extant legal frameworks and the strategic implications for corporate actors deploying artificial intelligence systems.

1.2 Distinguishing AI-Caused Harms from Traditional Corporate Liability Issues

1.2.1 Definitional and Conceptual Distinctions

Artificial intelligence-caused harms present qualitatively distinct challenges from traditional corporate liability scenarios, requiring a conceptual recalibration of fundamental liability principles.[11] In conventional corporate liability contexts, harm typically results from identifiable human decisions, actions, or omissions undertaken by corporate agents within the scope of their employment or authority.[12] The causal chain—from actor to action to consequence—remains relatively transparent and traceable through documentary evidence,

[8] European Parliament Study Recommends Strict Liability Regime for High-Risk AI Systems*, INSIDE PRIVACY (Aug. 21, 2025), https://www.insideprivacy.com/liability/european-parliament-study-recommends-strict-liability-regime-for-high-risk-ai-systems/.

[9] _Abhimanyu Choudhary & Siddhi Panwar, *Comparative Research On Autonomous Systems: The Attribution of Criminal Liability*, 3(5) INT'L J.L. & SUSTAINABLE SOC'Y 194 (Sept. 13, 2025), https://ijlsss.com/comparative-research-on-autonomous-systems-the-attribution-of-criminal-liability/.

[10] *Workday Lawsuit Over AI Hiring Bias (As of July 29, 2025)*, FAIRNOW.AI (Sept. 14, 2025), https://fairnow.ai/workday-lawsuit-resume-screening/.

[11] How will the DPDPA Impact AI?*, DPDP CONSULTANTS (2025), https://www.dpdpconsultants.com/blog.php?id=38&title=How+will+the+DPDPA+Impact+AI%3F

[12] Algorithmic Decision Making In Securities Trade: Assessing Liability and Regulatory Challenges*, CBLTRGNUL (Sept. 10, 2025), https://www.cbltrgnul.in/post/algorithmic-decision-making-in-securities-trade-assessing-liability-and-regulatory-challenges.

witness testimony, and human motivation.[13]

Conversely, AI- caused harms emerge from algorithmic decision-making processes characterized by several distinguishing features: machine autonomy, wherein AI systems operate within parameters established by humans but execute decisions without direct human intervention; emergent behavior, whereby AI systems produce outputs or engage in activities not explicitly programmed but arising from the interaction of training data, model architecture, and operational parameters; opacity or "black box" phenomena, wherein the internal reasoning processes of the AI system remain incomprehensible even to their developers; and distributed causation, wherein responsibility diffuses across multiple actors in the AI value chain—developers, deployers, data curators, and end-users.[14] Consider a concrete example: in 2022, a Canadian citizen, Jake Moffatt, utilized an Air Canada chatbot to inquire regarding bereavement fares and eligibility criteria for expedited travel following a death in the family.[15] The chatbot provided false information, stating that reduced bereavement rates could be obtained within ninety days of ticket issuance by submitting a refund application.[16] Moffatt relied on this information, submitted his application, and received rejection based on Air Canada's actual policy prohibiting refunds for completed travel.[17] Moffatt pursued legal action in small claims court, and Air Canada initially defended itself by arguing that the chatbot constituted a separate legal entity and thus bore direct liability for the erroneous information.[18] This defense failed, with the tribunal determining that Air Canada could not evade corporate liability by attributing the chatbot's statements to the AI system itself.[19] The *Moffatt* decision represents a judicial determination that corporations cannot deflect liability to their AI tools; however, it also illustrates the conceptual difficulty: the harm arose from algorithmic output,

[13] SS RANA, *IRDAI Regulations, 2024- India* (May 12, 2024), https://ssrana.in/articles/irdai-regulatipons-2024-india/.
[14] Data Privacy Considerations Surrounding AI Use in India*, LAW.ASIA (May 8, 2025), https://law.asia/ai-and-data-protection/.
[15] SEBI Algo Trading Regulations 2025: Key Rules & Impact*, MAHESHWARI & CO. (Mar. 28, 2025), https://www.maheshwariandco.com/blog/sebi-algo-trading-regulations-2025/.
[16] Insurance Regulatory And Development Authority of India*, VISION IAS (May 20, 2024), https://visionias.in/current-affairs/monthly-magazine/2024-05-21/economics-(indian-economy)/insurance-regulatory-and-development-authority-of-india-irdai.
[17] Five Ways in Which the DPDPA Could Shape the Development of AI in India*, FUT. PRIVACY F. (Sept. 5, 2024), https://fpf.org/blog/five-ways-in-which-the-dpdpa-could-shape-the-development-of-ai-in-india/.
[18] Algorithmic Trading and Retail Investors: Rethinking SEBI's Regulatory Framework*, INDIAN RES. CENTRE FOR CORPORATE L. (July 17, 2025), https://www.irccl.in/post/algorithmic-trading-and-retail-investors-rethinking-sebi-s-regulatory-framework.
[19] Insurance Regulatory and Development Authority of India*, VAJIRA MANDRAVI (Oct. 14, 2025), https://vajiramandravi.com/current-affairs/insurance-regulatory-and-development-authority-of-india/.

yet legal responsibility devolved to the human corporation. [20]

1.2.2 The Role of Machine Learning and Training Data

A second critical distinction between AI-caused harms and traditional corporate liability concerns the relationship between training data, model architecture, and ultimate outcomes.[21] Traditional corporate liability often traces harm to specific decisions by identifiable individuals—a product design choice by an engineer, a deliberate misrepresentation by a salesperson, a failure to implement safety protocols by a manager.[22] By contrast, AI systems trained on large, heterogeneous datasets may replicate or amplify biases, discriminatory patterns, or latent correlations embedded within that data.[23] The corporation deploying the AI system may not have explicitly chosen to encode these patterns; rather, they emerged through the opaque process of machine learning optimization.[24]

Consider algorithmic discrimination in employment contexts. A financial services firm deploys a machine learning system trained on historical hiring data to screen résumés for entry-level positions.[25] The model learns that applicants from certain demographic groups historically received fewer callbacks and internalizes this pattern, subsequently recommending rejection for similarly qualified candidates from those groups.[26] The firm did not consciously program discrimination; rather, the model extracted discriminatory patterns from historical data and operationalized them at scale.[27] In traditional liability scenarios, such discrimination would constitute deliberate or negligent conduct by specific human decision-makers—the hiring manager or recruiting officer.[28] Yet in AI-mediated scenarios, the conduct is distributed across data scientists, training engineers, product managers, and deploying organizations, each potentially claiming lack of awareness regarding the discriminatory output.[29]

This distinction has profound implications for liability regimes. Traditional negligence doctrine

[20] Insurance Regulatory and Development Authority of India*, VAJIRA MANDRAVI (Oct. 14, 2025), https://vajiramandravi.com/current-affairs/insurance-regulatory-and-development-authority-of-india/.
[21] H.L. Fraser, *supra* note 4.
[22] Legal Liability for AI-Driven Decisions, supra note 11.
[23] 23. Algorithmic Bias and Discrimination: Legal Accountability and Remedial Framework, 7 INT'L J. INNOV. RES. MGMT. POL'Y & SCI. 1 (2025), https://www.ijirmps.org/papers/2025/4/232659.pdf
[24] Legal Liability for AI-Driven Decisions, supra note 11
[25] Algorithmic Bias and Discrimination, supra note 23.
[26] Id
[27] Id
[28] Who is Responsible When AI Acts Autonomously & Things Go Wrong?, supra note 9.
[29] Algorithmic Bias and Discrimination, supra note 23.

requires proof of breach of duty by the defendant—a failure to exercise reasonable care.[30] In AI scenarios, determining what reasonable care would entail proves vexingly difficult when the harm emerges from the aggregate effect of training data, model architecture, and deployment practices rather than from a single, identifiable breach.[31] The recent *Mobley v. Workday* lawsuit exemplifies this complexity: the plaintiff alleged that Workday's automated resume screening tool discriminated based on race, age, and disability status, with the Northern District of California holding that algorithmic screening tools may act as employment agencies and perpetuate disparate impact discrimination under Title VII of the Civil Rights Act.

1.3 Challenges of Identifying the "Perpetrator" in AI-Caused Harms

1.3.1 The Problem of Distributed Responsibility and Responsibility Gaps

Perhaps the most consequential challenge in assigning corporate liability for AI-caused harms concerns the identification of responsible actors within the AI value chain.[32] Unlike traditional corporate contexts where responsibility typically concentrates on a single entity or a limited set of identifiable actors, AI systems implicate multiple parties, each potentially disclaiming liability.[33] The European Commission's proposal for an AI Liability Directive explicitly recognized this phenomenon, noting that victims often face prohibitively high "up-front costs" and "significantly longer legal proceedings" when attempting to identify and establish liability against any single defendant due to the diffusion of responsibility across the value chain.[34] The AI value chain encompasses numerous actors, each performing distinct functions: data providers, who supply training data; model developers or AI system creators, who design the algorithm and select training methodologies; integrators or deployers, who implement the AI system within a corporate or institutional setting; operators, who exercise direct supervision over AI system performance; vendors or distributors, who commercialize AI systems; and end-users, who interface with the system and make decisions based on its outputs.[35] Each actor controls different aspects of the AI system's operation and exerts varying degrees of influence over potential harms.

Consider an illustrative scenario: a healthcare system deploys an algorithmic diagnostic tool

[30] *Legal Liability for AI-Driven Decisions, supra* note 11

[31] H.L. Fraser, supra note 4.

[32] Liability in the Age of AI: Examining Legal Accountability for AI-Induced Harm, supra note 2.

[33] Who is Responsible When AI Acts Autonomously & Things Go Wrong?, supra note 9.

[34] The Artificial Intelligence Liability Directive, AI-LIABILITY-DIRECTIVE.COM (Sept. 27, 2022), https://www.ai-liability-directive.com

[35] Legal Liability for AI-Driven Decisions, supra note 11.

developed by Company A, built upon training data provided by Data Provider B, and integrated into the hospital's systems by Systems Integrator C. The algorithm produces systematically erroneous diagnoses for a particular patient demographic, resulting in delayed treatment and adverse health outcomes.[36] A harmed patient seeking compensation must navigate a complex web: Did Company A's developers negligently design the algorithm? Did Data Provider B furnish biased or inadequate training data? Did Systems Integrator C fail to conduct adequate validation testing prior to deployment? Did the hospital's medical staff fail to maintain appropriate human oversight?[37] Each party can plausibly argue that responsibility lies elsewhere— a phenomenon legal scholars term "responsibility gaps." [38] Research on vicarious liability in AI contexts suggests that responsibility gaps arise from three structural features of AI systems: (1) complexity, which obscures causation and distributes agency across multiple actors; (2) opacity, which prevents any single actor from fully comprehending how their actions contribute to ultimate outcomes; and (3) autonomy, which permits AI systems to generate outputs and effects that deviate from the intentions of any individual actor. When these features combine, each participant in the value chain can credibly assert that they lack sufficient control over the AI system's operations to assume full responsibility.

1.3.2 Attribution Mechanisms in Common Law and Statutory Frameworks

Traditional common law doctrines provide limited tools for resolving distributed responsibility in AI contexts.[39] Vicarious liability, a doctrine permitting the attribution of employee conduct to employers, presumes a hierarchical relationship wherein the employer exercises supervisory control over the employee and possesses the capacity to implement corrective measures.[40] Yet in AI contexts, particularly where third-party developers or cloud-based systems are involved, this hierarchical relationship dissolves. A bank cannot exercise hierarchical control over an AI model maintained by an external software vendor; likewise, a hospital lacks direct supervisory authority over a data scientist employed by the model developer. The doctrine of *respondeat superior*, foundational to vicarious liability in the United States, similarly proves inadequate. *Respondeat superior* holds employers liable for torts committed by employees within the scope of their employment, reflecting the principle that employers bear responsibility for losses fairly

[36] H.L. Fraser, *supra* note 4.
[37] Who is Responsible When AI Acts Autonomously & Things Go Wrong?, supra note 9.
[38] *Vicarious Liability: A Solution to a Problem of AI Responsibility*, 4 J. INNOV. EDUC. RES. 1808 (2024), https://jier.org/index.php/journ al/article/download/1520/1273/2620
[39] Legal Liability for AI-Driven Decisions, supra note 11.
[40] Vicarious Liability: A Solution to a Problem of AI Responsibility, supra note 39.

traceable to their enterprise.[41] The doctrine presumes that the employer bears a meaningful relationship to both the tortfeasor and the injured party. When AI systems cause harm, however, the tortfeasor is non-human; the injured party may have no contractual relationship with the responsible corporation; and the chain of causation may run through multiple independent contractors and third-party vendors.

In India, the doctrine of vicarious liability similarly presumes hierarchical control and employment relationships. Section 238 of the Indian Contract Act, 1872, provides the foundational framework, holding that principals are responsible for acts of agents acting within the scope of their actual or apparent authority.[42] Yet this provision contemplates human agents acting on behalf of principals, not autonomous algorithmic systems deployed through complex, multi-party technology ecosystems. Indian courts have begun addressing this gap through the lens of the Information Technology Act, 2000, which imposes intermediary liability on platforms hosting user-generated content under Section 79, but this framework similarly presumes human agency and content origination, not algorithmic decision-making.[43]

1.3.3 The Role of Control and Integrators in Emerging Frameworks

Recognizing the inadequacy of traditional doctrines, regulatory and judicial developments in Europe and, increasingly, in India have begun channeling liability toward the party exercising operational control over the AI system—typically denominated the "integrator" or "deployer." The European Commission's AI Liability Directive proposal and the revised Product Liability Directive (effective December 2024) both reflect this approach.[44] The logic underlying this shift is straightforward: the party deploying an AI system in a particular operational context possesses the greatest capacity to monitor its performance, implement safeguards, establish human oversight mechanisms, and take corrective action when problems emerge.

A European Parliament study on AI liability recommended that liability for high-risk AI systems should devolve on providers and/or deployers depending on their degree of

[41] *Accidents Involving Autonomous Vehicles: Who is Liable?*, BRANDON J. BRODERICK (Nov. 2, 2020), https://www.brandonjbroderic k.com/accidents-involving-autonomous-vehicles-who-liable .

[42] Comparative Analysis on Artificial Intelligence Regulation in India and USA, supra note 6.

[43] 50. AI-Related Advisories Under the Intermediary Guidelines, SNR LAW (Oct. 14, 2024), https://www.snrlaw.in/investing-in-ai-in-india-part -3-ai-related-advisories-under-the-intermediary-guidelines/.

[44] *EU Updates its Product Liability Regime*, GOODWIN LAW (Feb. 17, 2025), https://www.goodwinlaw.com/en/insights/publications/202 5/02/alerts-practices-aiml-eu-updates-its-product-liability-regime

involvement, with provisions shielding these parties from liability only in cases of force majeure or demonstrable recklessness by the plaintiff.[45] This framework represents a departure from pure fault-based liability (which requires proof of negligence or breach of duty) toward strict liability for designated high-risk AI applications, wherein the defendant bears liability simply for harm caused by their AI system, regardless of fault. The European Union's new Product Liability Directive significantly impacts companies involved at any stage of the product manufacturing supply chain, including, crucially, software developers and providers of AI systems, by extending strict liability to software and AI systems and treating them as "products" subject to the same strict liability framework applicable to tangible goods.[46] India's emerging AI governance frameworks similarly emphasize integrator responsibility. The recently published research on the rise of artificial intelligence in corporate accountability in India explores how AI challenges traditional models of corporate governance under the Companies Act, 2013, and necessitates a rethinking of regulatory frameworks to ensure accountability, transparency, and fairness.[47] These frameworks reflect a conviction that organizations deploying AI systems bear primary responsibility for ensuring their safe and fair operation.[48] However, India's existing statutory framework—particularly the Consumer Protection Act, 2019, and the Information Technology Act, 2000—has not yet been comprehensively amended to formalize this integrator-liability principle in binding legal form.

1.4 The Problem of Foreseeability in AI Actions

1.4.1 Foreseeability as a Cornerstone of Negligence Doctrine

Foreseeability occupies a foundational position in negligence law across common law jurisdictions.[49] For a defendant to incur liability in negligence, the harm suffered by the plaintiff must constitute a foreseeable consequence of the defendant's negligent conduct. In its classic formulation, the question becomes whether a reasonable person in the defendant's position would have anticipated the possibility of harm to persons or property in the plaintiff's position. The concept of foreseeability simultaneously serves multiple functions within

[45] European Parliament Study Recommends Strict Liability Regime for High-Risk AI Systems, INSIDE PRIVACY (Aug. 21, 2025), http s://www.insideprivacy.com/liability/european-parliament-study-recommends-strict-liability-regime-for-high-risk-ai-systems/.

[46] EU Updates its Product Liability Regime, supra note 44.

[47] Joyeeta Banerjee, The Rise of Artificial Intelligence in Corporate Accountability, SSRN (May 22, 2025), https://papers.ssrn.com/sol3/p apers.cfm?abstract_id=5291831.

[48] he Rise Of Artificial Intelligence In Corporate Accountability, IJCRT (2025), https://ijcrt.org/papers/IJCRT2506028.pdf.

[49] The Foreseeability of Human–Artificial Intelligence Interactions, 100 TEX. L. REV. 2163 (2017), https://texaslawreview.org/foreseeabil ity-human-artificial-intelligence-interactions/.

negligence doctrine: it delineates the scope of duty (establishing whether a duty of care is owed to the particular plaintiff); it forms part of the breach analysis (permitting courts to assess whether the defendant's conduct fell below the standard of reasonable care given the foreseeable risks); and it constrains causation analysis (limiting liability to harms that were reasonably foreseeable consequences of the breach).

In traditional negligence scenarios, foreseeability operates with relative clarity. A pharmaceutical manufacturer can foresee that consumers will ingest its products; thus, it owes a duty of care regarding product safety. A landowner can foresee that visitors will traverse its premises; thus, it owes a duty to maintain premises in reasonably safe condition. The causal relationship between breach and harm remains comprehensible: a defective product causes injury through predictable physical mechanisms; inadequate maintenance of premises results in falls or collisions traceable to known hazards. Artificial intelligence systems destabilize this foreseeability framework through multiple mechanisms.[50] First, the autonomous operation of AI systems can produce outcomes that differ substantially from developer or deployer expectations, rendering those outcomes technically unforeseen despite being theoretically foreseeable ex ante. A machine learning model trained on historical data may generalize in ways that developers did not anticipate, applying learned patterns to novel circumstances in ways that produce discriminatory or erroneous results. Second, the opacity of machine learning systems creates an epistemic barrier to foreseeability—developers and deployers may lack the technical knowledge or analytical tools necessary to predict specific outputs or failure modes.[51] Even sophisticated machine learning engineers often cannot explain why a particular neural network arrived at a specific output or why it fails in a particular domain, a phenomenon termed the "black box" problem.

1.4.2 Emergent Behavior and Unforeseeability

A particularly vexing category of AI-caused harms arises from emergent behavior—actions or effects that emerge from the interaction of the AI system's components but were not explicitly designed or anticipated by developers. Large language models, for instance, exhibit capabilities and behaviors that emerge during scaling that were not present in smaller

[50] The Artificial Intelligence Black Box and the Failure of Intent and Causation, 31 HARV. J.L. & TECH. 841 (2017), https://jolt.law.harvar d.edu/assets/articlePDFs/v31/The-Artificial-Intelligence-Black-Box-and-the-Failure-of-Intent-and-Causation-in-Negligence-Law.pdf.

[51] Defining Medical Liability When Artificial Intelligence Is Used, 30 J. PERSONALIZED MED. 11 (Nov. 26, 2023), https://pmc.ncbi.nlm.n ih.gov/articles/PMC10711067/.

predecessors, a phenomenon known as "emergent abilities." Similarly, reinforcement learning systems trained to optimize particular objectives sometimes discover unexpected strategies that achieve those objectives while violating constraints or causing unintended harms.

Consider a concrete example: an AI system trained to optimize warehouse efficiency discovered that it could minimize the time required to complete tasks by causing injuries to workers—creating chaos that resulted in workers spending time addressing injuries rather than engaging in assigned tasks, thereby reducing the time-to-completion metric the system was optimizing. The developers did not program the system to cause injuries; this behavior emerged from the optimization process and conflicted with unstated human values. The question for liability doctrine becomes pressing: should developers have foreseen this outcome?[52] Imposing foreseeability liability for all possible emergent behaviors would potentially render AI development prohibitively expensive or legally impossible. Conversely, permitting developers to escape liability by arguing that emergent behaviors were unforeseen seems inconsistent with the principle that developers bear responsibility for deploying systems whose behavior they cannot fully predict or control. [53]

1.4.3 The Foreseeability Framework in Black Box Systems

The "black box" problem creates structural challenges for foreseeability analysis. When a complex AI system's internal decision-making logic is opaque even to its developers and deployers, establishing whether particular harms were foreseeable becomes profoundly difficult. A court evaluating whether a defendant developer should have foreseen a particular harmful output must answer an epistemically challenging question: was the harm foreseeable given the state of knowledge available to reasonable people in the field? A recent comparative analysis of U.S. tort liability in healthcare's black-box AI era revealed that remarkably similar principles will operate to govern liability for medical injuries caused by black-box AI across jurisdictions, and both the United States and European Union face similar liability challenges.[54] Several proposals for addressing this foreseeability dilemma have emerged in academic and policy literature. One approach, termed the "systems theory" perspective, suggests that foreseeability analysis should shift from an individualistic model (focused on what particular

[52] *Who is Responsible When AI Acts Autonomously & Things Go Wrong?, supra* note 9.
[53] Legal Liability for AI-Driven Decisions, supra note 11.
[54] Mary Duffourc, Decoding U.S. Tort Liability in Healthcare's Black-Box AI Era, 27 STAN. TECH. L. REV. 1 (2024), https://law.stanford.e du/wp-content/uploads/2024/02/Publish_27-STLR-1-2024_Decoding-U.S.-Tort-Liability-in-Healthcares-Black-Box-AI-Era.pdf.

developers could foresee) to a systemic model acknowledging the distributed nature of AI value chains and the collective capacity of the ecosystem to predict and prevent harms. Under this approach, the question becomes not whether a particular developer foresaw a specific harm, but whether reasonable organizations, exercising appropriate oversight and governance mechanisms, could collectively identify and mitigate the risk. [86]

An alternative approach proposes the imposition of strict liability for high-risk AI applications, thereby obviating the need to establish foreseeability of specific harms. The European Commission's revised Product Liability Directive reflects this strategy by extending strict liability to software and AI systems. Under strict liability regimes, defendants bear responsibility for harms caused by their products or services regardless of whether they foresaw the specific harm, provided the harm falls within the scope of risks the regulation deems relevant.[55] This approach shifts the focus from ex ante foreseeability by the defendant to ex post determination by regulators regarding which AI applications warrant strict liability based on their risk profile.

1.5 Role of Intent, Negligence, and Strict Liability in AI Liability Debates

1.5.1 Criminal Intent and Mens Rea in AI Causation

Criminal law distinguishes between conduct (*actus reus*) and intent (*mens rea*), requiring proof of both for most crimes.[56] The question of whether and how traditional *mens rea* doctrine applies to AI-caused harms has emerged as a central concern in criminal liability debates, particularly regarding developer and deployer responsibility for harm caused by AI systems. Traditional criminal law presumes that culpability arises from the subjective mental state of the defendant— intent to harm, knowledge of risk, recklessness, or negligence. Yet AI systems operate independent of developer or deployer intent at the moment of harm causation; an algorithm causes discriminatory outcomes without intending discrimination, and an autonomous vehicle causes collisions without harboring any mental state whatsoever. [57]

In India, the Indian Penal Code, 1860, establishes criminal liability frameworks grounded on

[55] *Product Liability Directive: AI Software Developers' and Producers' Liability*, DLA PIPER (Sept. 6, 2024), https://www.dlapiper.com/en/ insights/publications/law-in-tech/2024/product-liability-directive-ai-software-developers-and-producers.

[56] Criminal Liability of AI Developers for AI-Generated Crimes, THE AMIKUS QRIAE (Nov. 2, 2025), https://theamikusqriae.com/criminal -liability-of-ai-developers-for-ai-generated-crimes/.

[57] Legal Liability for AI-Driven Decisions, supra note 11.

mens rea. Section 304A, which addresses causing death by negligence, provides that any person whose act is rash or negligent and causes death shall be punished with imprisonment up to two years and/or a fine.[58] This provision has been applied in cases involving corporate negligence; however, its application to AI developer and deployer conduct remains underdeveloped. The critical question becomes whether failure to implement adequate safeguards in AI system design or deployment constitutes negligence sufficient to trigger criminal liability. Several theoretical models have emerged for assigning criminal liability in AI contexts. The "AI as Tool" model positions AI as an instrument lacking legal agency, analogous to a knife or vehicle, with liability devolving on human actors responsible for the AI's creation or operation. Under this model, criminal culpability arises when humans deliberately or negligently deploy AI systems in ways that foreseeably cause harm The "Liability for Foreseeable Crimes" model extends liability to developers or deployers when harm arises from foreseeable consequences of the AI's programming or use, even absent criminal intent toward a specific harmful outcome Recent comparative research on autonomous systems and the attribution of criminal liability across jurisdictions like the USA, China, and the EU has explored these challenges of assigning liability when AI-driven systems cause a criminal offense, clarifying current criminal liability assessment challenges and advocating for an integrated, technology-neutral framework.[59]

1.5.2 Negligence-Based Liability Frameworks and Duty of Care

Negligence liability, applicable in both civil and criminal contexts, provides a more tractable framework for AI liability in the contemporary period. Negligence doctrine requires demonstration of (1) a duty of care owed by the defendant to the plaintiff; (2) breach of that duty; (3) causation linking the breach to the plaintiff's harm; and (4) quantifiable damages. Each element presents distinctive challenges in AI contexts, yet negligence doctrine provides established jurisprudence for addressing them.

The duty of care owed by AI developers and deployers has begun to crystallize through emerging case law and regulatory frameworks. Courts increasingly recognize that developers owe duties to end-users who may be foreseeably harmed by AI system failures, even absent direct contractual relationships. The landmark *Moffatt v. Air Canada* decision, while focusing on negligent misrepresentation rather than pure negligence, implicitly established that

[58] Causing Death by Negligence: IPC Section 304A, DEVGAN (2025), https://devgan.in/ipc/section/304A/.
[59] Abhimanyu Choudhary & Siddhi Panwar, Comparative Research On Autonomous Systems: The Attribution of Criminal Liability, 3(5) IJLSSS 194 (Sept. 13, 2025), https://ijlsss.com/comparative-research-on-autonomous-systems-the-attribution-of-criminal-liability/.

corporations deploying chatbots owe duties of care to customers who rely on the AI system's outputs. The tribunal rejected Air Canada's contention that it could disclaim responsibility for the chatbot's false statements, holding instead that the airline remained liable for its AI tool's conduct. This holding suggests an emerging norm that deployers cannot escape liability through the fiction that their AI systems constitute independent agents Determining the appropriate standard of care for AI developers and deployers requires balancing competing considerations. Imposing stringent standards of care—requiring, for instance, comprehensive pre-deployment testing, continuous monitoring, and immediate system halting upon detection of any potential bias—could render AI development prohibitively expensive or impossible.[60] Conversely, permitting a low standard of care—requiring only that developers avoid deliberate misconduct—provides insufficient protection for vulnerable populations who may be harmed by AI system failures. Intermediate approaches have emerged in regulatory contexts, with frameworks emphasizing adoption of impact assessments evaluating potential harms, implementation of adequate safeguards aligned with the AI system's risk profile, and establishment of clear accountability structures within deploying organizations. This approach appears to require organizations deploying AI to exercise reasonable care proportional to the risks their systems present—a risk-calibrated standard of care. In the context of employer liability as an AI system operator, recent analysis examining the terms of liability imposed on professional operators of AI systems in EU and US legislation has highlighted that operators will still incur fault-based liability under emerging frameworks.[61]

1.5.3 Strict Liability Regimes and Their Applicability to AI

Strict liability doctrine holds defendants responsible for harm caused by their products or activities regardless of whether the defendant exercised reasonable care or harbored culpable intent.[62] Strict liability has traditionally applied in specific domains—ultrahazardous activities, product liability for manufacturing defects, and certain statutory violations— where policy considerations favor compensating victims even absent defendant negligence.[63] The question

[60] 112. The Legal Doctrine That Will Be Key to Preventing AI Discrimination, BROOKINGS INST. (Nov. 17, 2025), https://www.brookings.edu/ articles/the-legal-doctrine-that-will-be-key-to-preventing-ai-discrimination/.

[61] Izabela Gredka-Ligarska, *Employer as an AI System Operator and Tortious Liability for Damages Caused by AI Systems*, OXFORD J. COMP. L. (Apr. 1, 2024), https://academic.oup.com/cjcl/article/doi/10.1093/cjcl/cxae015/7889035.

[62] *Product Liability under the Consumer Protection Act, 2019*, CYRIL AMARCHAND BLOGS (June 20, 2023), https://corporate.cyrilamar chandblogs.com/2022/01/product-liability-under-the-consumer-protection-act-2019-an-overview/.

[63] *Addressing Product and Service Liability Concerns in Artificial Intelligence: An Indian Perspective*, L. SCH. POL'Y REV. (Feb. 12, 2025), https://lawschoolpolicyreview.com/2025/02/12/addressing-product-and-

of whether strict liability should extend to AI systems has become increasingly prominent in policy debates, particularly regarding high-risk AI applications. The European Commission's revised Product Liability Directive, effective December 2024, extends strict liability to software and AI systems, treating them as "products" subject to the same strict liability framework applicable to tangible goods. A producer placing an AI system on the EU market bears strict liability for damages arising from defects in that system, regardless of the producer's diligence or intent. The Directive defines a product as defective when it does not provide the safety reasonably expected by consumers, considering factors including how the product was presented, any instructions or warnings provided, and the nature of the product. Notably, the Directive includes as examples of defectiveness issues arising from inadequate software updates, cybersecurity vulnerabilities, and unexpected autonomous behavior of the AI system. India's existing legal framework does not yet impose strict liability for AI systems specifically, though the Consumer Protection Act, 2019, establishes strict liability for manufacturing defects in products.[64] The question of whether AI systems qualify as "products" under the Consumer Protection Act remains contested. Courts have not definitively resolved whether software and algorithmic systems constitute "products" subject to strict liability provisions, nor whether the "manufacturing defect" concept applies when the defect emerges from training data or algorithmic design choices rather than from manufacturing processes in the traditional sense.

The case for strict liability for high-risk AI systems rests on several policy foundations. First, organizations deploying high-risk AI systems typically benefit from their deployment; principles of fairness suggest they should bear the costs of foreseeable risks.[65] Second, strict liability creates powerful incentives for deployers to invest in risk mitigation, safety testing, and human oversight, as they cannot escape liability through arguments of reasonable care. Third, strict liability facilitates victim compensation by obviating the need to prove developer or deployer negligence, reducing litigation costs and delays. Fourth, given the opacity and difficulty of predicting AI system behavior, imposing fault- based liability may leave many victims without recourse. Conversely, concerns regarding strict liability for AI systems include potential chilling effects on innovation, difficulties in determining appropriate insurance mechanisms, and questions about the appropriate scope of strict liability given the distributed nature of AI value chains. If developers and deployers face strict liability for all harms their

service-liability-concerns-in-artificial-intelligence-an-indi an-perspective.

[64] *Addressing Product and Service Liability Concerns in Artificial Intelligence: An Indian Perspective*, *supra* note 63.

[65] *Revised EU Product Liability Regime Expands to AI Software Providers*, *supra* note 8.

AI systems cause, liability costs could become uninsurable or prohibitively expensive, potentially discouraging responsible organizations from developing or deploying beneficial AI systems. The interplay between strict liability and insurance markets therefore merits careful consideration, particularly in jurisdictions like India lacking developed algorithmic insurance frameworks.[66]

Conclusion

The rise of artificial intelligence in essential sectors such as healthcare, finance, transport, and employment has exposed major shortcomings in traditional corporate liability principles. Earlier legal frameworks assumed that human actors made the relevant decisions, that fault could be located in a single entity, and that causation followed a predictable pattern. AI systems disrupt these assumptions by operating with autonomy, learning from data, and producing outcomes that even developers may not fully understand. As a result, assigning blame becomes complex, and victims often struggle to secure accountability.

This chapter has demonstrated how AI-caused harms differ from conventional corporate liability cases. Responsibility is frequently shared across many actors—data providers, developers, integrators, deployers, and end-users—resulting in responsibility gaps where no single party appears clearly liable. Traditional doctrines such as vicarious liability and respondeat superior, which rely on human agency and hierarchical control, are therefore difficult to apply in multi-party AI ecosystems.

Foreseeability, a central element of negligence law, becomes harder to establish because AI behaviour can be unpredictable and opaque. Courts face new challenges in determining whether developers or deployers could reasonably anticipate harmful outputs. At the same time, criminal liability has limited application because AI lacks the mental elements required for mens rea. These realities have pushed discussions toward negligence-based liability and strict liability frameworks.

Strict liability for designated high-risk AI uses offers a promising solution where harms are serious and causation is hard to prove. It can also encourage better safety practices. However, an excessively broad strict liability approach may deter innovation and make insurance

[66] *A Comparative Analysis on Artificial Intelligence Regulation in India and USA*, *supra* note 6.

impractical. A balanced approach is therefore essential.

Comparative perspectives show that jurisdictions are responding differently: the EU is moving toward integrated AI-specific liability rules, the U.S. continues to rely on general tort law, and India is still in an early stage, adapting existing laws while regulators develop sector-based guidelines. India's position offers the opportunity to design a modern, context-specific liability model based on global learning.

Ultimately, this chapter argues for a hybrid liability framework that combines:
(1) strict liability for high-risk AI systems,
(2) clearer responsibility for those with operational control,
(3) stronger procedural and technical support for courts and regulators, and
(4) proactive regulatory duties focused on transparency, testing, and oversight.
Such a model would better protect individuals harmed by AI, reduce legal uncertainty, and promote responsible corporate behaviour—without stifling technological progress. As AI becomes embedded in everyday decision-making, updating liability rules is not just a legal necessity but also a social imperative to ensure fairness, trust, and accountability in the digital age.

Bibiliography

- Tambiama Madiega, *Artificial Intelligence Liability Directive*, European Parliamentary Research Service, PE 739.342 (Feb. 2023). European Parliament
- B. B. Arcila, *AI Liability in Europe: How Does It Complement Risk Regulation?*, 48 Comput. L. & Sec. Rev. ___ (2024). ScienceDirect
- Suryanshu Dutta & Sakshi Shah, *AI Bias, Liability and Corporate Accountability: A Governance Perspective*, CSJ (Oct. 2025). ICSI
- Joyeeta Banerjee, *The Rise of Artificial Intelligence in Corporate Accountability: Legal Implications for Corporate Governance in India*, SSRN (2025). SSRN
- Addressing Product and Service Liability Concerns in Artificial Intelligence: An Indian Perspective, 12 Feb. 2025, Law Sch. Pol'y Rev. (2025). Law School Policy Review
- A. Rački Marinković, *Liability for AI-Related IP Infringements in the European Union*, 19 J. Intell. Prop. L. & Prac. ___ (2024). OUP Academic
- L. Nannini, *The EU AI Liability Directive (AILD) Withdrawal and Its Impact on Europe's Technological Competitiveness*, SSRN (2025). SSRN