Page 1424 - 1429

INTERNATIONAL JOURNAL OF LAW MANAGEMENT & HUMANITIES

[ISSN 2581-5369]

Volume 9 | Issue 4

2026

When the Machine Errs, Who Bleeds? Reconstructing Civil Liability for AI-Generated Harm in India

AYUSH KUMAR UPADHYAY*

ABSTRACT

AI systems are moving out of laboratories and into hospitals, courts and financial markets, yet Indian law remains largely silent on civil liability for the harm they cause. This article contends that the doctrinal gap is neither temporary nor curable by analogy to existing tort principles: it is a structural failure traceable to three intertwined features of advanced AI, namely opacity in decision-making, causal indeterminacy and distributed agency. It examines the inadequacies of negligence, of product liability and of the intermediary safe harbour in the context of AI harm, and proposes a three-tiered liability framework consisting of a rebuttable presumption of deployer liability, mandatory audit and disclosure obligations on developers, and a no-fault compensation fund for victims of high-risk autonomous systems, grounded in the constitutional obligations flowing from Article 21 of the Constitution of India. The proposal draws on the EU Artificial Intelligence Act and on the proposed AI Liability Directive, since withdrawn, and requires calibration for Indian conditions.

Keywords: *artificial intelligence, civil liability, tort law, negligence, product liability, absolute liability, article 21, eu ai act, ai liability directive, algorithmic accountability, india*

I. INTRODUCTION

The governance of artificial intelligence is among the most pressing legal questions of the twenty-first century.[1] The liability question it raises is easy to state and hard to answer: who is liable when an AI system causes harm? The chain of causation in AI-generated harm bears little resemblance to that in *Donoghue v. Stevenson*, where the manufacturer could be held liable to the ultimate consumer for bottling contaminated ginger beer.[2] This is not merely a theoretical puzzle; it has real victims. An uncalibrated clinical decision support system misdiagnoses cancer. A credit-scoring model refuses loans on racial grounds. A person is killed by an

* Author is a Student at Tata Institute of Social Sciences, Mumbai, Maharashtra, India.

[1] JACOB TURNER, ROBOT RULES: REGULATING ARTIFICIAL INTELLIGENCE 1-4 (2019).

[2] *Donoghue v. Stevenson*, [1932] AC 562 (HL).

 [ISSN 2581-5369]

autonomous navigation system. In none of these cases does existing Indian law offer a clear remedy.[3]

The argument proceeds in four parts. Part II examines the inadequacies of orthodox tort doctrine. Part III sets out the scope and the limits of the statutory options. Part IV proposes a three-tiered structure. Part V grounds the proposal constitutionally and concludes.

II. The Tort Doctrine is Failing the AI Victim

A. The Negligence Dead End

To establish a claim in negligence, the claimant must prove duty, breach, causation and damage.[4] Every element strains against the characteristics of contemporary machine-learning systems. Duty founders on proximity: a general-purpose large language model developer stands at a considerable remove from harm suffered by a third party through a downstream deployer's fine-tuned application. Breach is afflicted by opacity: neural networks are not merely complex but, as Pasquale puts it, 'black boxes' whose inner workings cannot readily be reconstructed from the outside.[5] Interpretability research suggests that even a model's creators may be unable to offer legally meaningful explanations of individual outputs.[6] Causation is the deepest problem. Where harm arises from an emergent interaction between a model, a dataset, a deployment context and user behaviour, no one of which would have produced it alone, the but-for test is not satisfied.[7]

B. Product Liability and Its Limits

One approach is to treat AI systems as products and their outputs as defects under the Consumer Protection Act 2019. Indian courts have already held that professional services fall within the ambit of consumer protection, and the academic literature suggests that AI models may be treated as products covered by defect liability.[8] The analogy cannot, however, resolve the post-sale learning problem: an AI system may be defect-free at the point of sale and become harmful through continued training or exposure to user data. Classical product liability doctrine offers a poor account of a product that alters its legally relevant properties after it has been sold.

---

[3] Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 Calif. L. Rev. 513, 540 (2015).
[4] Clerk & Lindsell on Torts para. 8-04 (Michael A. Jones & Anthony M. Dugdale eds., 23rd ed. 2020).
[5] Frank Pasquale, The Black Box Society 6-9 (2015).
[6] Finale Doshi-Velez & Been Kim, *Towards a Rigorous Science of Interpretable Machine Learning* 1 (arXiv:1702.08608, 2017).
[7] Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. Davis L. Rev. 399, 421 (2017).
[8] *M/s Spring Meadows Hospital v. Harjol Ahluwalia*, (1998) 4 SCC 39 (India).

 [ISSN 2581-5369]

C. The Intermediary Escape

Section 79 of the Information Technology Act 2000 confers a safe harbour on intermediaries that do not exercise editorial control over third-party content.[9] As read down in *Shreya Singhal v. Union of India*, the policy is sound in so far as it avoids chilling internet infrastructure. Deployers of generative AI, however, are not mere conduits: they select models, shape outputs and direct commercialisation. Where control of that kind is exercised, extending intermediary immunity opens an accountability gap that neither the legislature nor the Supreme Court intended.[10]

III. Statutory Mapping and Its Shortcomings

India's approach to AI governance has been aspirational rather than regulatory. NITI Aayog's National Strategy (2018) and the IndiaAI Mission (2024) of the Ministry of Electronics and Information Technology emphasise investment and ethical principle but supply no enforceable liability rules.[11] The Digital Personal Data Protection Act 2023 imposes accountability obligations on data fiduciaries and requires purpose limitation,[12] but its architecture is oriented towards privacy rather than towards the physical or economic harm that flows from AI outputs. The Bureau of Indian Standards has issued voluntary guidance on responsible AI, which has not yet been absorbed into the standard of care applied by courts in a negligence analysis.

The lacuna is structural. Unlike the European Union, which has enacted the Artificial Intelligence Act 2024,[13] India has no risk-tiered statutory framework classifying AI applications by level of risk and requiring ex ante conformity assessment for high-risk applications. The consequence is that a victim harmed by a high-risk autonomous system, whether medical, judicial or infrastructural, enjoys no greater legal leverage than a victim harmed by a low-stakes recommendation algorithm. That is neither theoretically nor politically sound.

The underlying difficulty is institutional and epistemic. Courts are asked to pronounce on the safety of systems they cannot inspect, to apply standards of care that have been settled neither judicially nor legislatively, and to apportion loss among parties none of whom may have caused it entirely. Citron's concept of 'technological due process'[14] is particularly apposite here: where

---

[9] Information Technology Act, No. 21 of 2000, s. 79 (India); *see Shreya Singhal v. Union of India*, (2015) 5 SCC 1 (India) (reading down s. 79(3)(b)).

[10] Jack M. Balkin, *The Three Laws of Robotics in the Age of Big Data*, 78 OHIO ST. L.J. 1217, 1228 (2017).

[11] NITI AAYOG, *National Strategy for Artificial Intelligence* 14 (2018); MINISTRY OF ELECTRONICS & INFORMATION TECHNOLOGY, *IndiaAI Mission* (Gov't of India 2024).

[12] Digital Personal Data Protection Act, No. 22 of 2023, s. 8 (India).

[13] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), arts. 9-15, 2024 O.J. (L 2024/1689).

[14] Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1251-53 (2008).

 [ISSN 2581-5369]

an algorithmic decision deprives a person of life, liberty or property without an intelligible explanation, it violates Article 21 of the Constitution as read in *K.S. Puttaswamy v. Union of India*.[15]

IV. Three Tiers of Responsibility

A. Tier One: A Rebuttable Presumption of Deployer Liability

The first tier addresses the present accountability deficit by placing primary responsibility on the entity that deploys an AI system in a commercial setting. The justification is control: the deployer selects, configures and markets the system, and is best placed to absorb or to insure the risk.[16] This is a form of enterprise liability, akin to the strict liability attaching to abnormally dangerous activities under the *Rylands* principle.[17]

The presumption should, importantly, be rebuttable. A deployer who can show that the harm was caused solely by a latent defect in the developer's model, one that the deployer could not reasonably have been expected to discover through diligent conformity assessment, should be able to join the developer as a defendant. This apportions fault by reference to proximity while leaving the victim with a right of action against a single identifiable defendant. The proposed EU AI Liability Directive, withdrawn by the European Commission in 2025, would have introduced a comparable disclosure mechanism obliging operators of high-risk AI to evidence the behaviour of their systems; the procedural analogue remains worth adopting.[18]

B. Tier Two: Mandatory Audit and Disclosure Obligations

The second tier addresses the epistemic problem. Courts, victims and regulators presently have no reliable means of assessing how an AI system behaves. This article proposes that high-risk AI systems, defined by reference to the healthcare, criminal justice, financial services and critical infrastructure sectors, be required to maintain and to produce on request an audit trail covering (i) training data provenance and known biases; (ii) model architecture and the basis of key parameter choices; and (iii) post-deployment monitoring logs.[19]

The obligation is a doctrinal rendering of the black box critique.[20] It does not demand full explainability, which many neural architectures cannot deliver, but a minimum of transparency

[15] India Const. art. 21; *K.S. Puttaswamy v. Union of India*, (2017) 10 SCC 1, paras. 300-310 (India) (Chandrachud, J.).
[16] Sheila Jasanoff, The Ethics of Invention: Technology and the Human Future 211-13 (2016).
[17] *Rylands v. Fletcher*, (1868) LR 3 HL 330 (HL).
[18] 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), art. 3, COM (2022) 496 final (Sept. 28, 2022) (proposal withdrawn by the European Commission, 2025).
[19] Doshi-Velez & Kim, *supra* note 6, at 3-5.
[20] Pasquale, *supra* note 5, at 140-42.

 [ISSN 2581-5369]

sufficient to make legal process possible. The doctrine of *res ipsa loquitur* can be adapted to the same end: where an AI system causes harm in circumstances that would not ordinarily arise in the absence of fault, and where evidence of the system's behaviour lies peculiarly within the developer's control, a rebuttable inference of defective design should arise if the audit record is not produced. Parliament could confer on an AI regulator investigative powers comparable to those the Competition Commission of India exercises under the Competition Act 2002.

C. Tier Three: A No-Fault Compensation Fund

The third tier recognises that the two-party tort model is structurally inadequate for a class of cases in which harm is systemic, diffuse and causally over-determined. Where a pattern of harm implicates several AI systems at once, as with algorithmic discrimination in lending or in content moderation, the injury to a particular plaintiff cannot be traced to any single deployer. A no-fault compensation fund financed in proportion to revenue offers a means of compensating victims without requiring proof of individual causation.[21]

Such a regime is not unprecedented in Indian law. The no-fault liability schemes under the Employees' Compensation Act 1923, the Public Liability Insurance Act 1991 and the Motor Vehicles Act 1988 are each designed to meet situations of causal indeterminacy or informational asymmetry that would otherwise defeat recovery.[22] The absolute liability of hazardous enterprises laid down by the Supreme Court in *M.C. Mehta v. Union of India* rests on a constitutional foundation. Extending that reasoning to AI-generated harm is a logical step rather than a leap.[23]

V. Constitutional Grounding and Conclusion

The proposed framework is not merely a pragmatic response to a regulatory void; the Constitution requires it. In *K.S. Puttaswamy*, the Supreme Court held that the right to informational self-determination forms part of the right to life and personal liberty guaranteed by Article 21.[24] Where a decision materially affects life, liberty or livelihood, where the basis of that decision is unintelligible, and where no means of challenge exists, the state bears not merely a permissive power to legislate but an affirmative duty to ensure that the subject of the automated decision has a meaningful remedy.[25]

---

[21] Algorithmic Accountability Act of 2022, H.R. 6580, 117th Cong. s. 3 (2022); *cf.* Algorithmic Accountability Act of 2019, H.R. 2231, 116th Cong. (2019).

[22] Viktor Mayer-Schonberger & Thomas Ramge, Reinventing Capitalism in the Age of Big Data 122 (2018).

[23] *M.C. Mehta v. Union of India*, (1987) 1 SCC 395 (India); *Union Carbide Corp. v. Union of India*, (1989) 1 SCC 674 (India).

[24] *K.S. Puttaswamy*, *supra* note 17, para. 127 (Chandrachud, J.); India Const. art. 21.

[25] Joseph Raz, The Morality of Freedom 166-68 (1986).

 [ISSN 2581-5369]

The three-tier approach discharges that obligation through deployer liability as the default (so that victims are not burdened with identifying who was at fault), disclosure (so that legal process becomes epistemically possible), and a compensation fund (so that diffuse systemic harm is not left uncompensated for want of individual causation). The tiers draw on the tripartite structure of the EU AI Act while accounting for India's own regulatory architecture and for the predominance of small and medium enterprises in the Indian technology sector, which compliance obligations designed for hyperscale platforms would otherwise deter.

It may be objected that prescriptive liability rules would dampen AI innovation at a moment when India is seeking to become a global powerhouse in the field. The objection conflates certainty with permissiveness. The present regime, in which the effective rule is one of no liability, does not encourage responsible innovation; it encourages the externalisation of risk. An insurance market is a precondition of the sustainable commercial deployment of high-risk AI, and a workable allocation of liability is a precondition of an insurance market. As the Srikrishna Committee observed, accountability and innovation are not mutually exclusive values.[26]

India stands at a constitutional and technological juncture. The materials for a principled AI liability framework already exist in tort law, in constitutional law and in comparative regulatory experience. What is wanting is the political will to assemble them before the harms of AI make doctrinal order impossible to achieve. The framework outlined here is contestable and open to revision, but it is architecturally sound and rooted in the basic principle that those who deploy autonomous systems capable of causing harm are answerable to those they harm.

*****

[26] COMMITTEE OF EXPERTS UNDER THE CHAIRMANSHIP OF JUSTICE B.N. SRIKRISHNA, *A Free and Fair Digital Economy: Protecting Privacy, Empowering Indians* (Ministry of Electronics & Information Technology, 2018).

 [ISSN 2581-5369]