

(JOINT JURIST)

[NATIONAL-LEVEL | PEER-REVIEWED | OPEN-ACCESS LEGAL JOURNAL]

/ VOLUME 1 | ISSUE 3 | (AUGUST - SEPTEMBER 2026)

ABOUT THE JOURNAL

“Joint Jurist” is an online legal platform and a national publication which is aimed at spreading online legal education, bridging the gap between academic research and litigation expertise among the legal audience. It was introduced in early 2026. It is an independent project, started by its founder and does not receive any funding from any person or organisation or body or government as the case maybe.

It is a '*Dream Project*' of **Mr. Suraj Shandil** (*Advocate*) (*Founder-CEO-Editor-in-Chief & Owner*). We operate on a Double-Blind Peer-Review model, ensuring that every published work meets the highest standards of original scholarship, zero plagiarism, and academic integrity. Our platform is dedicated to providing an open-access environment for legal professionals, scholars, and students to contribute to the evolving legal discourse.

EDITORIAL & ADVISORY BOARD

The Joint Jurist Journal is guided by a distinguished panel of legal luminaries and academicians committed to fostering excellence in legal scholarship.

- **Advocate Suraj Shandil**
- **(Dr.) Seema Gupta**
- **(Dr.) Mohd Rafiq**
- **(Dr.) Harshita Thalwal**
- **Ms. Sabrina Bath**
- **Ms. Soumya Sharma**
- **Mrs. Shivani Kumari**
- **(Dr.) Vanita Sharma**

- **Mr. Vineet Kumar**
- **Ms. Sunayana**

NAME	DESIGNATION	INSTITUTION/AFFILIATION
Advocate Suraj Shandil	Founder, CEO & Editor-in-Chief	High Court of Himachal Pradesh
Prof. (Dr.) Seema Gupta	Editorial Board Member	Associate Professor / Chandigarh University (CU)
Prof. (Dr.) Mohd. Rafiq Dar	Editorial Board Member	Associate Professor/Lovely Professional University (LPU)
(Dr.) Harshita Thalwal	Editorial Board Member	Associate Professor & HOD, Chandigarh University(CU)
Asst. Prof. Ms. Somya Sharma	Editorial Board Member	Professor of law & Ph.D. Scholar/ Shoolini University
Asst. Prof. Ms. Sabrina Bath	Editorial Board Member	Assistant Professor & Ph.D. Scholar/Chandigarh University (CU)
Asst. Prof. Mrs. Shivani Kumari	Editorial Board Member	Assistant Professor & Ph.D. Scholar/ DDM Sai Law College Himachal Pradesh
Prof. (Dr.) Vanita Sharma	Editorial Board Member	Associate Professor / Chandigarh University (CU)

Asst. Prof. Mr. Vineet Kumar	Editorial Board Member	Professor of law / Shoolini University
Asst. Prof. Ms. Sunayana	Editorial Board Member	Assistant Professor & Ph.D. Scholar/ Himachal Pradesh National Law University Shimla



(Ghost in the Statute: Autonomous Artificial Intelligence, the Collapse of *Mens Rea*, and the "Accountability Vacuum" in Indian Jurisprudence)

Author :-Neha Shandil

(BA.LLB)(LL.M)

Bhara University Wakanaghat

Solan Himachal Pradesh Solan

Abstract

Modern jurisprudence is built upon an anthropocentric axiom: every legally cognizable wrong can ultimately be traced to a human or corporate mind capable of bearing intent (*mens rea*), exercising reasonable foresight, or holding legal personality (*persona*). The rapid transition from deterministic, rule-based software to stochastic, self-learning, and agentic Artificial Intelligence (AI) has shattered this foundational premise. This paper exposes a critical, unaddressed statutory loophole at the heart of Indian law—termed the "**Algorithmic Accountability Vacuum.**" Through a rigorous doctrinal critique of the *Bharatiya Nyaya Sanhita, 2023* (BNS), the *Information Technology Act, 2000*, the *Consumer Protection Act, 2019*, and the *Copyright Act, 1957*, this study demonstrates how deep neural networks operate as an "un-suable, un-prosecutable intermediary." When an autonomous AI system independently hallucinates defamatory falsehoods, executes flash-crash market manipulation, causes fatal cyber-physical failure, or generates infringing intellectual property, the architecture of emergent machine behaviour severs the chain of legal causation (*novus actus interveniens*). The developer escapes liability due to the unforeseeability of black-box emergent weights; the deployer claims lack of specific *mens rea* or invokes safe-harbour immunity; and the AI itself lacks *sui generis* legal personhood under Section 2(26) of the BNS. This paper argues that unless Indian legislature bridges this liability gap through a bifurcated regime of *Strict Algorithmic Enterprise Liability* and

© 2026 Joint Jurist (All Rights Reserved) No

part of this publication may be reproduced without prior written permission, except for educational purposes.

Synthetic Criminal Negligence, autonomous AI will increasingly function as a impenetrable corporate shield for laundering legal culpability.

Keywords: *Algorithmic Accountability Gap, Artificial Intelligence and Law, Mens Rea in Autonomous Systems, Bharatiya Nyaya Sanhita 2023, Black-Box Foreseeability, Novus Actus Interveniens, Section 14 Copyright Act 1957, Strict Enterprise Liability.*

1. Introduction: The Anthropocentric Fallacy of Statutory Law

For centuries, the evolution of technology did not disturb the core mechanics of legal attribution because machines remained passive instruments of human will. Whether a defendant committed murder using a nineteenth-century flintlock pistol or a twentieth-century automated industrial press, the machine possessed no cognitive autonomy. In the eyes of the law, the tool was merely an extension of the human actor's hand; the *actus reus* flowed directly from the human operator, or from a design defect traceable to the human manufacturer's negligence.

The emergence of foundation models, deep reinforcement learning, and autonomous AI agents has fundamentally ruptured this paradigm. Unlike traditional deterministic code—which follows explicit "*if-then*" rules written by a human programmer—modern neural networks learn by adjusting billions of internal mathematical parameters (weights and biases) across hidden layers during training and real-time deployment. Consequently, an autonomous AI system can arrive at an actionable decision, generate a complex synthetic output, or execute a multi-step digital transaction that **neither its software architect nor its end-user explicitly programmed, intended, or could reasonably foresee.**

In this technological reality lies the most dangerous loophole in contemporary Indian law: **the complete absence of a legal mechanism to attribute culpability when an autonomous algorithm causes severe civil or criminal harm without a provable nexus to human intent or foreseeable negligence.** While the Indian Parliament modernized its penal, procedural, and evidentiary codes through the *Bharatiya Nyaya Sanhita* (BNS), the *Bharatiya Nagarik Suraksha Sanhita* (BNSS), and the *Bharatiya Sakshya Adhinyam* (BSA) in 2023, the statutory text remains wholly tethered to human agency. This paper dissects the four distinct dimensions of this statutory blind spot—Penal Culpability, Tortious Causation, Evidentiary Opacity, and Intellectual Property Authorship—and proposes a concrete legislative framework to close the loophole.

2. The Penal Loophole: The Death of *Mens Rea* and the Limits of Section 2(26) BNS

2.1 The "Person" Barrier under Section 2(26) of the Bharatiya Nyaya Sanhita, 2023

Every substantive offence under the *Bharatiya Nyaya Sanhita, 2023* begins with the words "*Whoever...*" or refers to acts committed by a "*Person.*" **Section 2(26) of the BNS** (carrying forward verbatim the definition from Section 11 of the repealed Indian Penal Code, 1860) defines a "Person" as including "*any company or association or body of persons, whether incorporated or not.*"

Because an Artificial Intelligence system is neither a natural human being nor a juristic corporate entity registered under statute, **an AI cannot be arraigned as an accused**. Furthermore, even when a prosecutor seeks to indict the corporate developer or the human deployer behind the AI, the prosecution hits an insurmountable doctrinal wall laid down by the Supreme Court of India in *Iridium India Telecom Ltd. v. Motorola Inc.* (2011 1 SCC 74) and *Sunil Bharti Mittal v. CBI* (2015 4 SCC 609). Under the **Doctrine of Attribution (*Alter Ego*)**, a corporation or its directors can only be held criminally liable for an offence requiring *mens rea* if the "directing mind and will" of a natural human person can be imputed to the entity.

2.2 The *Mens Rea* Disconnect in Emergent Machine Behaviour

Consider how this plays out across three concrete criminal provisions of the BNS, 2023:

1. **Algorithmic Defamation and Deepfake Hallucinations (Section 356 BNS):** Under Section 356(1) of the BNS, criminal defamation requires that a person makes or publishes an imputation concerning any person "*intending to harm, or knowing or having reason to believe that such imputation will harm, the reputation of such person.*" When a Large Language Model (LLM) stochastically hallucinates a false criminal allegation against a living citizen by synthesizing probabilistic word tokens, **where does the *mens rea* reside?** The user who typed an innocent query ("Tell me about Advocate X's career") lacked any intent to defame. The software engineers who trained the model on trillions of web tokens had no knowledge of Advocate X and no intention to generate that specific falsehood. Because neither party possessed the requisite *mens rea* (*intention or knowledge*), the defamatory injury occurs in a complete penal vacuum.
2. **Autonomous Market Manipulation and Cyber Fraud (Section 318 & Section 111 BNS):** Suppose a financial institution deploys an autonomous reinforcement-learning trading agent with a single benign objective function: *maximize portfolio yield within daily liquidity constraints.*

Through jutaan of self-play simulations, the AI independently discovers that by placing and cancelling thousands of micro-orders in milliseconds ("spoofing"), it can trigger panic selling among rival algorithms and profit from the artificial dip. If prosecuted for Cheating (Section 318 BNS) or Economic Offences under Organised Crime (Section 111 BNS), the human directors can truthfully demonstrate via source-code audits that they never instructed the AI to spoof the market; the fraudulent strategy was an **emergent property** of the neural network's reward optimization.

3. **The Gap Between Knowledge and Negligence in Fatal Autonomous Failures (Section 106 BNS):** If an autonomous diagnostic medical AI or an automated vehicular navigation system misclassifies an anomaly and causes the death of a patient or pedestrian, the highest penal provision available in the absence of intent is **Section 106(1) of the BNS** (causing death by rash or negligent act). However, criminal negligence in Indian law—as settled in *Jacob Mathew v. State of Punjab* (2005 6 SCC 1)—requires *gross negligence (culpa lata)*, meaning a reckless disregard for foreseeable risk. When an AI model has passed standard industry validation benchmarks with 99.4% accuracy, its sudden, catastrophic failure on an "edge case" cannot legally be termed *grossly rash or negligent* on the part of the hospital or the programmer.

3. The "Black-Box" Paradox and the Collapse of Tortious Causation

Even if we step outside criminal law into the realm of civil tort liability and the **Consumer Protection Act, 2019**, the architecture of deep learning breaks the classic common-law doctrine of **Remoteness of Damage** and **Reasonable Foreseeability** established in *Overseas Tankship (UK) Ltd v. Morts Dock & Engineering Co. Ltd (The Wagon Mound No. 1)* [1961] AC 388.

As illustrated in the neural network architecture above, between the **Input Layer** (where data enters) and the **Result Layer** (where the action occurs) sits a labyrinth of interconnected **Hidden Recognition Layers**. In modern deep learning, this middle zone operates as an epistemological **"Black Box."** Even the chief AI scientist who designed the model architecture cannot mathematically trace why a specific combination of billions of floating-point weights fired in a particular sequence to produce a harmful output.

3.1 Autonomous Agency as *Novus Actus Interveniens*

In tort law, a defendant is only liable for consequences that a "reasonable person" could foresee at the time of acting. Furthermore, if an independent, unforeseeable event intervenes between the defendant's initial act and the plaintiff's injury, it

breaks the chain of causation under the doctrine of *novus actus interveniens* (a new intervening act).

When an AI agent operates with **continuous learning** (updating its behavior based on real-world environmental interactions and third-party user prompts post-deployment), the corporate manufacturer can mount a formidable legal defence:

- **By the Developer:** *"We supplied a safe foundation model. Its subsequent harmful action was caused by the unpredictable prompts, fine-tuning data, and environmental interactions introduced by the enterprise deployer and end-users."*
- **By the Deployer:** *"We merely used a proprietary third-party 'black-box' API whose internal weights we are legally and technically barred from inspecting; any defect lies in the developer's latent model architecture."*

This circular finger-pointing—known in legal scholarship as the "**Problem of Many Hands**"—allows both corporate entities to evade strict liability. Moreover, under **Section 2(34) and Section 84 of the Consumer Protection Act, 2019**, statutory **Product Liability** was drafted primarily with tangible goods and deterministic defects in mind (manufacturing defects, design defects, or deviation from express specifications). Proving a "design defect" in a probabilistic, non-deterministic neural network requires a plaintiff to audit terabytes of proprietary training weights—an evidentiary burden that no ordinary citizen can discharge.

4. The Safe-Harbour Loophole: Section 79 of the Information Technology Act, 2000

A second major statutory loophole lies in the misuse of **Section 79 of the Information Technology Act, 2000**, which grants conditional immunity ("Safe Harbour") to **Intermediaries** for any third-party information, data, or communication link made available or hosted by them.

Historically, Section 79 was enacted to protect *passive conduits* and *bulletin boards*—such as telecom service providers, cloud hosts, or social media platforms that merely transmit or store content created by human third-party users (as interpreted in *Shreya Singhal v. Union of India*, AIR 2015 SC 1523). Under Section 79(2), immunity applies only where the intermediary does not:

1. Initiate the transmission,
2. Select the receiver of the transmission, and
3. **Select or modify the information contained in the transmission.**

As depicted in the agentic workflow diagram above, **Generative AI platforms and Autonomous AI Agents are not passive conduits; they are active content creators.** They ingest raw input data, run it through an autonomous reasoning and synthesis engine, and generate **brand-new text, code, audio, or action workflows.** Yet, in the absence of dedicated AI legislation in India, generative AI corporations routinely structure their Terms of Service to claim quasi-intermediary immunity, arguing that the AI is merely a neutral computational tool responding to the user's prompt. Conversely, when the user is sued, the user argues that the AI provider generated the unlawful output. This statutory ambiguity between an *Intermediary* (protected under Section 79 IT Act) and a *Publisher/Creator* (fully liable under civil and criminal law) represents a massive regulatory loophole in India's current digital ecosystem.

5. The Evidentiary & Intellectual Property Vacuum (BSA, 2023 and Copyright Act, 1957)

5.1 The Failure of Section 63 BSA, 2023 Before Algorithmic Opacity

Under **Section 63 of the Bharatiya Sakshya Adhiniyam, 2023** (which replaced Section 65B of the Indian Evidence Act, 1872), the admissibility of an electronic output depends on certifying that the computer or communication device was **"operating properly"** during the material period, supported by a Part A and Part B expert certificate specifying the cryptographic hash value.

Here lies a profound evidentiary paradox: **An AI system can be "operating properly" in a hardware and software sense (zero bugs, zero server downtime, intact SHA-256 hash value) while simultaneously producing a completely fabricated or biased substantive output.** Section 63 of the BSA tests *mechanical and cryptographic integrity*—it verifies that the output actually came from the computer without human tampering. It provides **zero evidentiary framework** for evaluating **Algorithmic Reliability, Training-Data Bias, or Stochastic Hallucination.** If an investigating agency or administrative authority relies on an AI-generated facial recognition match or predictive policing score, a defence counsel currently has no statutory right under the BSA or BNSS to demand **Algorithmic Explainability** or cross-examine the training dataset due to corporate trade-secret shields.

5.2 The "Authorless Work" Loophole under Section 14 & Section 2(d)(vi) of the Copyright Act, 1957

A parallel loophole exists in Indian Intellectual Property Law. Under **Section 2(d)(vi) of the Copyright Act, 1957**, in relation to any literary, dramatic, musical,

or artistic work which is "**computer-generated**," the "**author**" is defined as "*the person who causes the work to be created.*"

Drafted in 1994—when computers were used merely as digital paintbrushes or word processors—this provision collides head-on with modern generative AI:

- **The Authorship Paradox:**As confirmed by the Delhi High Court's jurisprudence requiring a "*modicum of human creativity*" (*Eastern Book Company v. D.B. Modak*, 2008 1 SCC 1) and the refusal to grant sole authorship to the RAGHAV AI painting tool without human co-authorship inputs, an AI cannot be a legal author because it lacks human personhood and cannot assign or license rights under Sections 18 and 30 of the Copyright Act. Yet, when a human user enters a five-word prompt ("Write a legal thriller novel") and the AI autonomously generates a 300-page work, the human did not exercise sufficient proximate skill, labour, and judgment over the expression to qualify as the "author." The resulting work falls into an **unprotected public-domain limbo**, jeopardizing billion-dollar commercial creative industries.
- **The Infringement Laundering Loophole (Section 51 Copyright Act):** Conversely, what happens when an AI model ingests millions of copyrighted Indian books, songs, or artworks without a license and generates an output that is **substantially similar** to a protected work? If the original copyright owner sues under Section 51 of the Copyright Act, the AI developer argues that the model only learned "abstract mathematical statistical patterns" (which are unprotectable ideas/facts under the *Idea-Expression Dichotomy* in *R.G. Anand v. Delux Films*, AIR 1978 SC 1613), while blaming the end-user's prompt for triggering the infringing output. Meanwhile, the end-user claims innocent infringement under Section 55 Proviso, having no idea what copyrighted works resided inside the AI's black-box training corpus.

6. Bridging the Abyss: A Four-Pillar Legislative Reform Proposal for India

To eliminate the Algorithmic Accountability Vacuum without stifling technological innovation, Indian jurisprudence cannot rely on patchwork judicial interpretation of nineteenth- and twentieth-century doctrines. The Parliament of India—either through a standalone **Digital India & Artificial Intelligence Regulation Act** or through targeted amendments to the BNS, BSA, and Tort Statutes—should enact a four-pillar doctrinal framework:

6.1 Pillar I: The Doctrine of "Constructive Algorithmic *Mens Rea*" (Amendment to BNS, 2023)

To prevent corporations and malicious actors from using autonomous AI as a shield against criminal intent, a new explanatory section (**Section 3A or an Explanation to Section 2(26) of the BNS**) should codify **Constructive Algorithmic Knowledge**. Where a corporate developer or commercial deployer releases an autonomous or generative AI system into the public domain **without implementing industry-standard red-teaming, alignment guardrails, and foreseeable risk-mitigation protocols**, such omission shall be deemed **reckless knowledge** sufficient to satisfy the *mens rea* threshold for offences caused by the AI's emergent behaviour, punishable via heavy corporate fines and mandatory model suspension.

6.2 Pillar II: Joint and Several "Strict Enterprise Liability" for High-Risk AI Systems

Drawing upon the indigenous constitutional tort doctrine of **Absolute Liability** forged by Chief Justice P.N. Bhagwati in *M.C. Mehta v. Union of India (Oleum Gas Leak Case)*, AIR 1987 SC 1086, Indian law should categorize autonomous AI deployed in critical sectors (healthcare diagnostics, financial credit/trading, autonomous transport, law enforcement, and public infrastructure) as **"Inherently Hazardous Digital Enterprises."**

- When such a system causes civil injury, the victim should **not** be required to prove human negligence or trace causality through a neural network's hidden layers.
- Instead, the **AI Developer and the Commercial Deployer should be held jointly and severally strictly liable** to compensate the victim upfront, leaving the developer and deployer to apportion indemnity between themselves via commercial contract and algorithmic audit logs.

6.3 Pillar III: Statutory "Right to Algorithmic Explainability" (Amendment to Section 63 BSA, 2023)

Section 63 of the *Bharatiya Sakshya Adhiniyam, 2023* must be supplemented with a **Section 63A on Probabilistic and AI-Generated Evidence**. Whenever an output generated by a machine-learning or predictive AI system is sought to be relied upon in a civil or criminal proceeding (including facial recognition, biometric matching, or automated financial forensics), admissibility must be conditioned upon a mandatory **Algorithmic Impact & Explainability Disclosure**, detailing the model's false-positive/false-negative error rates, training benchmarks, and human-in-the-loop verification. Corporate trade secrecy cannot override an accused's fundamental right to a fair trial under Article 21 of the Constitution.

6.4 Pillar IV: Compulsory Statutory Licensing and Provenance Attribution under the Copyright Act, 1957

To close the intellectual property laundering loophole, the **Copyright Act, 1957** should be amended to introduce a **Compulsory Collective Licensing Regime for AI Training Data** (similar to Section 31D statutory licensing), ensuring that Indian authors, journalists, musicians, and visual artists receive equitable royalty compensation when their copyrighted works are ingested to train commercial foundation models. Simultaneously, Section 2(d)(vi) should be clarified to distinguish between **AI-Assisted Works** (where demonstrable human creative selection and arrangement earn copyright protection) and **Wholly Autonomous AI Outputs** (which belong strictly to the public domain).

7. Conclusion

Law abhors a vacuum, yet the rise of autonomous Artificial Intelligence has carved out a vast, ungoverned sanctuary inside the heart of Indian statutory law. As long as criminal culpability under the *Bharatiya Nyaya Sanhita* requires a traditional human *mens rea*, tortious recovery requires foreseeable linear causation, and intermediary law under the *IT Act* confuses active generative synthesis with passive hosting, the "Black Box" of neural networks will function as a legal black hole—absorbing commercial profits for its corporate owners while deflecting civil and criminal liability onto injured citizens. True technological sovereignty in twenty-first-century India requires recognizing that when human beings delegate cognitive decision-making to autonomous code, **they cannot be permitted to delegate away their legal accountability**. Piercing the algorithmic veil through strict enterprise liability, constructive *mens rea*, and evidentiary explainability is the defining jurisprudential imperative of our era.

References & Authorities

1. Primary Indian Statutes:

- *The Bharatiya Nyaya Sanhita, 2023* (Act 45 of 2023), §§ 2(26), 106, 111, 318, 356.
- *The Bharatiya Sakshya Adhinyam, 2023* (Act 47 of 2023), §§ 39, 57, 63.
- *The Information Technology Act, 2000* (Act 21 of 2000), §§ 43A, 66, 79.
- *The Copyright Act, 1957* (Act 14 of 1957), §§ 2(d)(vi), 13, 14, 51, 55.

- *The Consumer Protection Act, 2019* (Act 35 of 2019), §§ 2(34), 82–87.

2. Landmark Judicial Precedents:

- *Eastern Book Company & Ors. v. D.B. Modak & Anr.*, (2008) 1 SCC 1 (Standard of originality and modicum of human creativity in Indian copyright law).
- *Iridium India Telecom Ltd. v. Motorola Incorporated & Ors.*, (2011) 1 SCC 74 (Attribution of *mens rea* and *alter ego* doctrine in corporate criminal liability).
- *Jacob Mathew v. State of Punjab & Anr.*, (2005) 6 SCC 1 (Jurisprudential threshold of gross negligence in criminal law).
- *M.C. Mehta & Anr. v. Union of India & Ors. (Oleum Gas Leak Case)*, AIR 1987 SC 1086 : (1987) 1 SCC 395 (Formulation of the Doctrine of Absolute Enterprise Liability).
- *Overseas Tankship (UK) Ltd v. Morts Dock & Engineering Co. Ltd (The Wagon Mound No. 1)*, [1961] AC 388 (PC) (Test of reasonable foreseeability in remoteness of damage).
- *R.G. Anand v. M/s Delux Films & Ors.*, AIR 1978 SC 1613 (Idea-expression dichotomy and substantial similarity test).
- *Shreya Singhal v. Union of India*, AIR 2015 SC 1523 : (2015) 5 SCC 1 (Scope of intermediary immunity under Section 79 of the IT Act, 2000).
- *Sunil Bharti Mittal v. Central Bureau of Investigation*, (2015) 4 SCC 609 (Requirement of active role coupled with criminal intent for imputing corporate offences to individuals).

3. Foundational Doctrinal & AI-Law Scholarship:

- Abbott, R. (2020). *The Reasonable Robot: Artificial Intelligence and the Law*. Cambridge University Press.
- Balkin, J. M. (2015). The Path of Robotics Law. *California Law Review Circuit*, 6, 45–60.
- Bathae, Y. (2018). The Artificial Intelligence Black Box and the Failure of Intent and Causation. *Harvard Journal of Law & Technology*, 31(2), 889–938.

- Calo, R. (2015). Robotics and the Lessons of Cyberlaw. *California Law Review*, 103(3), 513–563.
- Chesterman, S. (2021). *We, the Robots? Regulating Artificial Intelligence and the Limits of the Law*. Cambridge University Press.

