ARTIFICIAL INTELLIGENCE AND CRIMINAL RESPONSIBILITY IN INDIA: ADDRESSING LEGAL AND REGULATORY GAPS

Diksha Dabas, National Law Institute University, Bhopal

ABSTRACT

The article under discussion is entitled Artificial Intelligence and Criminal Responsibility in India: Legal Challenges and Regulatory Gaps and provides a detailed analysis of how artificial intelligence (AI) intersects with criminal responsibility in the Indian jurisdictional setting within the framework of legal and regulatory issues. It begins with the definition of AI and a discussion of its expansion, functions, and emerging applications in India, with particular emphasis on the transformative role of AI in governance, business, and digital ecosystems. The discussion places AI within the conventional framework of criminal law, highlighting the foundational principles of *actus reus* and *mens rea* in attributing guilt and examining the doctrinal contradictions that arise when criminal conduct involves non-human, autonomous systems.

Through an examination of automation and algorithmic decision-making, the article demonstrates how AI systems may unintentionally or indirectly result in unlawful consequences, thereby challenging traditional theories of intent, causation, and liability. It further analyses the complex legal issues surrounding the attribution of criminal intent in AI-driven actions. Particular emphasis is placed on identifying potentially responsible actors--such as developers, users, deployers, or corporate entities--and evaluating the institutional gaps that exist within current accountability mechanisms.

The article also engages with evolving jurisprudence in India and comparable international developments, illustrating the continuing debate over liability when algorithmic systems are involved in harmful outcomes. In addressing regulatory concerns, the study highlights the absence of a comprehensive AI-specific legal framework in India and critically evaluates the limitations of relying solely on existing legal provisions and policy recommendations.

Ethical considerations, including transparency, accountability, and responsible governance, are also examined. The article ultimately proposes legislative and policy-based interventions aimed at strengthening India's

criminal law framework to respond effectively to the rapidly evolving challenges posed by artificial intelligence technologies.

Keywords: Artificial Intelligence; Criminal Responsibility; *Mens Rea*; Algorithmic Accountability; AI Governance.

I. INTRODUCTION

AI has quickly changed both governance, business, and digital environments across the globe. The past few years have seen the influence of AI on different industries in India increase rapidly. Although the developments offer efficiency and innovation, they are also associated with complicated legal issues of accountability arising when AI systems bring harm.[1]

Old school criminal law is founded on the leading principles of *actus reus* (the guilty act) and *mens rea* (the guilty mind).[2] These theories assume that there is a human agent who can be either intentionally guilty, knowingly guilty, or negligently guilty. Nevertheless, autonomous AI systems are not conscious but work in terms of algorithms, thus violating commonly accepted theories of criminal responsibility.[3]

This paper is an exploration of how artificial intelligence and criminal responsibility relate in India. It examines the theoretical issues that emerge due to algorithmic systems delivering detrimental results and assesses whether the current legal framework, especially the Bharatiya Nyaya Sanhita, 2023[4] can cope with the rising problems. The article also examines the regulatory loopholes, the development of new jurisprudence, and international response to AI before suggesting possible legislative and policy responses.

II. DEFINING ARTIFICIAL INTELLIGENCE: SCOPE AND CAPABILITIES

In order to understand how artificial intelligence (AI) and the issue of criminal responsibility intersect, it will be essential to have a solid definition of what AI is, what it is capable of, and how it is gradually becoming a part of the systems and models used in modern society. Artificial Intelligence is not a singular technology but a collection of computational capacities, such as machine learning and natural language processing, robotics and predictive analytics,

[1] NITI Aayog, *National Strategy for Artificial Intelligence: #AIforAll* 5–7 (2018).
[2] Andrew Ashworth & Jeremy Horder, *Principles of Criminal Law* 95–98 (7th ed. 2013).
[3] Gabriel Hallevy, *When Robots Kill: Artificial Intelligence Under Criminal Law* 15–18 (2013).
[4] Bharatiya Nyaya Sanhita, No. 45 of 2023, India Code (2023).

among others, that allow systems to execute tasks that could otherwise demand human intelligence.

The concept of AI has now become inherently defined as having the ability to make autonomous decisions, something that challenge the conventional legal provisions of liability and intent. As of March 2026, the scope of AI in India has witnessed remarkable growth, fueled by strategic government initiatives, international collaborations, and significant private sector investments.[5] The capabilities of modern AI systems have evolved from simple rule-based execution to complex, generative, and adaptive behaviours. The existing AI systems are capable of creating content, streamlining logistics, diagnosing health issues, and controlling crucial infrastructure with the involvement of little human control.

This operational independence is important in the evaluation of criminal responsibility; as AI systems gain the ability to act independently, the site of responsibility gets more mystified. In the Indian context, these capabilities are being harnessed to drive economic growth and innovation across diverse sectors such as healthcare, agriculture, education, manufacturing, and public administration.[6]

The extent of this integration is empirically evidenced in the growth of the sector in terms of finances. The rapid expansion of the Indian AI market from USD 3.1 billion in 2020 to USD 14.7 billion in 2025- demonstrates the deep integration of algorithmic systems into critical sectors.[7] As AI systems increasingly influence financial, healthcare, and administrative decisions, the potential for algorithm-driven harm expands correspondingly, thereby intensifying the urgency of developing a coherent legal framework for criminal accountability.

Scholars have increasingly questioned whether traditional legal doctrines are capable of regulating autonomous systems whose actions may not always be directly attributable to identifiable human intent.[8] This high-pace economic growth requires a corresponding development of regulatory and legal systems to deal with the eventual conflict between

[5] Ministry of Finance, Gov't of India, *Union Budget 2026–27: Budget Speech* (Feb. 2026).
[6] NITI Aayog, *Responsible AI for All: National Strategy for Artificial Intelligence* (2021).
[7] India Brand Equity Found. (IBEF), *Artificial Intelligence Industry in India – Market Size & Growth Analysis (2020–2025)*(2025).
[8] Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 Cal. L. Rev. 513 (2015); Gabriel Hallevy, *When Robots Kill: Artificial Intelligence Under Criminal Law* (2013); Andrew D. Selbst, *Negligence and AI's Human Users*, 100 B.U. L. Rev. 1315 (2020).

automated systems and the criminal law.

In early 2026, the Indian government reinforced its commitment to this technological trajectory. The Union Budget 2026–27 placed a strong emphasis on fostering AI and emerging technologies, underscoring the intent to prioritize their integration within national development strategies.[9] The scope of AI in India is further defined by its increasing engagement with the global technological order. International AI forums and multilateral engagements in early 2026 reflected India's growing importance in global AI governance debates.[10]

Nonetheless, the use of AI goes beyond economic creation; it also touches on the governance systems and transnational regulatory frameworks. India's aspiration to contribute to global AI governance frameworks illustrates that AI development is not merely technological but strategic in character.[11] In spite of such developments, the definition of the scope of AI should also consider the intrinsic risks and limitations which are linked with its usage. The rapid pace of AI evolution implies persistent challenges, including ethical concerns, infrastructural gaps, algorithmic bias, cybersecurity vulnerabilities, and skill shortages.[12]

Additionally, the rise of advanced generative models and autonomous systems introduces new variables into the legal analysis of criminal responsibility, particularly with respect to data misuse, misinformation, and large-scale automated harms.[13] These developments expand the role of artificial intelligence from a tool of domestic economic policy to a broader strategic and governance concern.

Consequently, understanding AI in the context of criminal responsibility requires navigating a complex terrain where technological autonomy intersects with economic ambition, regulatory evolution, and national interest.[14]

In summary, defining AI in the context of this report requires acknowledging its dual nature as both an economic engine and a source of novel legal challenges. The progression of the Indian AI market from 3.1 Billion USD in 2020 to 14.7 Billion USD in 2025 serves as a quantitative

[9] Ministry of Fin., Gov't of India, *Union Budget 2026–27: Budget Speech*, supra note 1.
[10] AI Impact Summit 2026, *Official Summary & Policy Outcomes Report*, supra note 2.
[11] World Econ. Forum, *Global AI Governance Developments & India's Strategic Positioning* (2026).
[12] NITI Aayog, *Responsible AI for All*, supra note 3.
[13] Ministry of Fin., Gov't of India, *Union Budget 2026–27: Budget Speech*, supra note 1.
[14] World Econ. Forum, *Global AI Governance Developments & India's Strategic Positioning*, supra note 7.

testament to the technology's expanding presence.[15] Together with the active focus of the government on integrating AI and its participation in the international arena, one may state that AI functions are becoming ingrained in the socio-economic infrastructure of India. This explosive growth, however, also creates regulatory loopholes in which the characteristic features of AI that include autonomy, lack of transparency, and unpredictability find their direct way to the classical principles of criminal responsibility.

III. CRIMINAL RESPONSIBILITY: LEGAL FRAMEWORK AND MENTAL STATES

The classic framework upon which criminal responsibility in India has traditionally based has been the two critical factors the occurrence of the physical component of the crime (*actus reus*) and the accompanying mentality or intent (*mens rea*). As of March 2026, this framework has undergone a historic transformation with the implementation of the **Bharatiya Nyaya Sanhita, 2023**, which officially replaced the **Indian Penal Code, 1860**, thereby altering the statutory structure governing criminal responsibility.[16] This transition marks a critical juncture in Indian jurisprudence, shifting the philosophical underpinning of the justice system from the punitive-centric colonial legacy established in 1860 to a modernized approach emphasizing rehabilitation and justice-oriented sentencing.[17]

The new legislative framework still enshrines essentially humanistic ideas of guilt, intention, culpability, and reformation ideas which assumes moral agency and cognitive consciousness.

The definition of a "guilty mind" remains central to the attribution of criminal responsibility. Under the BNS, culpability continues to revolve around intention, knowledge, recklessness, and negligence.[18] The changes are based on the philosophy of reform thus present in the public discourse that surrounded the implementation of the BNS which is of modernization, simplification, and adherence to current standards of justice. A significant aspect of this reform is the formal recognition of community service as a sentencing alternative for specified offenses, signalling a shift toward corrective rather than exclusively carceral approaches.[19]

---

[15] India Brand Equity Found. (IBEF), *Artificial Intelligence Industry in India – Market Size & Growth Analysis*, supra note 4.
[16] *Bharatiya Nyaya Sanhita*, No. 45 of 2023 (India) (enforced Mar. 2, 2026).
[17] Press Info. Bureau, Gov't of India, *Implementation of New Criminal Laws* (Mar. 2026).
[18] *Bharatiya Nyaya Sanhita*, 2023, §§ 1–4 (India).
[19] *Bharatiya Nyaya Sanhita*, 2023, § 4(f) (India).

This however presents a conceptual paradox when used to Artificial Intelligence. Rehabilitation and community service assume the existence of a moral agent who can reflect, remorse and reintegrate functions that algorithmic systems, no matter how advanced they are, lack. Therefore, although the BNS is a progressive paradigm of human criminals, the regulative nature of the reformative orientation demonstrates a regulatory loophole in dealing with crimes that are either enabled or perpetrated by autonomous technologies.

Moreover, legal standards of ascription of mental capacity are still attached to biological and developmental indicators. Despite sweeping structural reforms, the minimum age of criminal responsibility remains seven years.[20] This age represents the assumption that a child under 18 years is psychologically immature to constitute an intent to commit a crime.

This has a jurisprudential issue, in relation to AI: to the extent that the law presupposes biological maturity and cognitive ability to assign guilt, how do we adapt this to synthetic systems that do not operate under conscious or moral awareness? The fact that the Indian criminal law continues to use age-based limits confirms that the criminal law is structurally geared towards human moral actors and it does not consider non-biological actors.

This model of mental states assessment is also formed by the advances of the mental health law. The **Mental Healthcare Act, 2017** introduced a rights-based model recognizing autonomy, advance directives, and institutional oversight in matters of mental health.[21] This method helps to hone judicial interpretation of the meaning of capacity and volition by defining within which a person can be criminal responsibility impaired because of an unsound mind.

Nevertheless, despite the increased protection against human cognitive vulnerability provided by courts and policymakers, these norms are also anthropocentric in nature. The law of sound mind and unsound mind is a legal concept that is based on a psychological premise on which human neurobiology is founded. The mental states that can be assessed according to such standards are not present in algorithmic systems.

Similarly, judicial and institutional efforts to promote mental well-being, such as guidelines encouraging psychological support mechanisms in educational and professional settings reflect

[20] *Bharatiya Nyaya Sanhita*, 2023, § 20 (India).
[21] *Mental Healthcare Act*, No. 10 of 2017 (India).

the expanding sensitivity of Indian law toward human mental health.[22] Liability is measured by the law, and algorithms are processed with probabilistic computation, without any experience.

Also, modern regulatory changes are reflected in the example of a wider readiness of the legislature to redefine the limits of criminal intent. Recent fiscal reforms have shifted certain non-violent economic infractions from criminal prosecution toward compliance-based enforcement frameworks.[23] This shift toward regulatory and not penal action implies the flexibility of the doctrine that can give an idea of how the harms associated with AI can be regulated. Rather than trying to establish an algorithmic *mens rea*, legislators can examine strict liability or compliance focused models of high-risk AI systems.

Overall, the legal environment by the year 2026 is characterized by radical change, as well as the persistence of humanist assumptions. The introduction of the Bharatiya Nyaya Sanhita represents a structural modernization of Indian criminal law, emphasizing proportionality, rehabilitation, and procedural clarity.[24] But the theoretical essence of criminal responsibility intention, cognitive awareness, biological maturity and moral responsibility still lie in the human psychology.

With the more in-depth adoption of AI technologies in the field of governance, commerce, and security, the distinction between a justice system that is based on human rehabilitation and the actual algorithmic autonomy poses an essential regulation challenge. The existing statutory system, though with its modernizing reforms, has not attempted to formulate a coherent theory of accountability of the non-human intelligent systems.

IV. ROLE OF ARTIFICIAL INTELLIGENCE IN CRIMINAL ACTS

The incorporation of the field of artificial intelligence into contemporary society has triggered the essential change in the very essence of the criminal activity, as the ways in which these mechanisms are deployed as tools of crime and, as living systems, affect the results. In the Indian context, the role of AI in criminal cases can be viewed as a continuum - on one end,

[22] Univ. Grants Comm'n, Gov't of India, *Guidelines on Mental Health & Well-Being in Higher Education Institutions*(2023).
[23] Ministry of Fin., Gov't of India, *Finance Act Amendments Relating to Decriminalisation of Economic Offences* (2023–2025).
[24] *Bharatiya Nyaya Sanhita*, No. 45 of 2023 (India).

there are already the automation of traditional crimes, and on the other end, there is algorithmic decision-making that makes it difficult to hold anyone accountable.

There must thus be a line of difference drawn between:

1. AI as an *actus reus* automating tool; and
2. AI as a brain which affects or creates results.

A. Automation of Criminal Acts

The highest implication of AI in regard to law is the automation of criminal activities. In these situations, autonomous technologies systems enhance the intensity, velocity, and anonymity of conventional crimes. Large-scale phishing campaigns, AI-generated disinformation, automated financial fraud, and deepfake extortion exemplify this pattern.[25]

AI is a tool, a technologically advanced one, nevertheless. The criminal law doctrine has not been changed: the purpose of firing the system belongs to the person, and the AI is just a specific way of its implementation.

Nonetheless, automation presents a complexity of investigation. Deepfake technology, for instance, allows highly convincing fabricated audiovisual content to be disseminated at unprecedented speed.[26]

B. Inadvertent Automation and System Exploitation

Another more subtle loophole is the situation where autonomous systems contribute to crime unintentionally. A logistics platform that is operated by an AI, say can optimize supply chains without realizing that it is transporting contraband goods. In this kind of situation the system works as expected but does not put illegality into perspective.

This brings complicated questions:

- Was the developer anticipating abuse?

[25] *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules*, 2021, as amended 2024 (India).
[26] Ministry of Electronics & Info. Tech., Gov't of India, *Advisory on Deepfake & Synthetic Media Compliance* (2024).

- Did it have sufficient supervision?
- Is it negligence to fail to foresee that there will be exploitation?

Early experiences with generative AI revealed the phenomenon of "hallucinations," wherein systems produce inaccurate or fabricated outputs.[27] Although this unpredictability is usually viewed through the prism of reliability issues, it can also encourage fraud, misinformation, or other ill behaviours.

C. Autonomous Decision-Making and Algorithmic Agency

It becomes even more difficult when AI systems are given the freedom of choice. The algorithms trading platforms, predictive analytics engines, and self-driving cars are driven by incessant data examination and adjustive reasoning. In financial markets, for example, AI-driven trading systems may execute high-frequency transactions based on patterns that are opaque even to their operators.[28] In any place where markets are being manipulated or destabilized by such systems, the question is; Was the harm caused by a human actor, accidentally designed or a result of complexity in the system?

Despite being originally supportive, the fast development of agentic AI may indicate growing autonomy in the operations. With the systems shifting to execution as opposed to recommendation, the conventional doctrines of causation and intent are stretched.

D. Bias and Systemic Harm

AI decision-making also presents risks of systemic bias. Predictive policing algorithms, for instance, may disproportionately target specific communities based on historical data patterns.[29] Such systems, even without the explicit discriminatory motive, can lead to the results that implicate the constitutional protections and even criminal malfeasance on the part of the public officials.

Global regulatory developments, such as the European Union's AI Act, reflect recognition of

[27] OpenAI, *GPT-4 Technical Report* (2023); see also NITI Aayog, *Responsible AI for All*, supra note 3.
[28] Sec. & Exch. Bd. of India (SEBI), *Circular on Algorithmic Trading Controls* (2024).
[29] Rashida Richardson et al., Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, 94 N.Y.U. L. Rev. 192 (2019).

these risks through risk-based classification models.[30] India lacks a similar legislative model of high-risk AI systems in their system of criminal law.

To sum up, the application of AI on crimes has a spectrum. On one side, AI is a tool that enhances human will and on the other, autonomous systems produce results that are difficult to trace back to the traditional attribution model. The current system of Indian criminal law, which is structured around human agency, then has to face a fast-growing technological environment in which criminal harm can be automated, new and algorithmically mediated.

The difference between algorithmic autonomy and criminal responsibility is only going to widen unless there is the adaptation of doctrines.

V. LEGAL CHALLENGES IN ADDRESSING AI AND CRIMINAL RESPONSIBILITY

A. Attribution of Criminal Intent in AI-Driven Actions

One of the most challenging jurisprudential challenges facing the Indian criminal law in 2026 will be the attribution of criminal intent or *mens rea* to artificial intelligence-driven actions. In India, criminal responsibility has always been based on the accordance of a malaise (*actus reus*) and culpable mental condition (*mens rea*). This architecture makes the assumption of human consciousness - a person who has an intention, foresight, recklessness, or negligence. Even the most advanced systems of artificial intelligence lack consciousness, moral consciousness, and intent. They make their operations based on algorithmic optimization and probabilistic inference as opposed to volitional choice.

With autonomous AI system generating detrimental results, either in the form of monetary manipulation, misinformation, or unjustified decision-making, the law faces an ontological challenge. The system can produce outputs that it can take to be like purposeful behaviour but it does not intend in any legally cognisable form. Its results are unanticipated effects of information trends and coded goals. The lack of sentience makes the traditional question of a concept of a guilty mind conceptually incoherent.

[30] 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*).

This challenge only increases when it comes to complex machine learning systems the inner workings of which are inaccessible even to their creators. The black box phenomenon has been known to interfere with the traditional causation doctrine. There must be a causal and proximate relationship between the actions of a human being and the ensuing injury to be criminal.

Nonetheless, as AI systems will evolve and improve after their deployment, there may emerge some harmful consequences that were never explicitly programmed or could be foreseen. The causal chain that can be followed was previously linear and traceable and is now diffusion among layers of code and training data and autonomous optimization processes.

A malicious deed can be committed but the AI, as well as any human agent, will not meet the criterion of a criminal mind. The developer might not have anticipated the harm in question. The unlawful outcome might not have been the intention of the user. Under those conditions, criminal law faces the danger of addressing action that lacks culpability- a situation that will hamper deterrence and normative coherence.

India's contemporary regulatory responses, including obligations imposed on digital intermediaries to remove harmful AI-generated content within prescribed timeframes, address downstream effects but do not resolve this upstream conceptual challenge.[31] Although this can help mitigate harm, it does not provide the answer to the underlying question of how to attribute criminal intent to an operating actor that is not a human but an algorithm. The existence of such doctrinal friction is an indication that there has been a need to recalibrate theoretically Indian criminal jurisprudence.

B. Challenges in Identifying Responsible Parties: Developers, Users, and AI Systems

The responsibility distribution becomes even more complicated as AI systems are no longer viewed as helping devices but as semi-autonomous agents. The liability can theoretically be on the developers, deployers, users or corporate organisation. In actual sense, however, both types are challenging in different ways.

Developers often face the first scrutiny. Their case would normally be liable due to careless design, lack of putting up safety measures, or misuse, which was foreseeable. Nevertheless,

[31] *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules*, 2021, r. 3(1)(b), as amended 2024 (India).

criminal law does not suffice in defective engineering but culpable mental conditions. It is not often easy to prove that a developer had a criminal outcome in mind, or that they were negligently insensitive to it. In high-scale AI applications, accountability is further spread out to groups of engineers, data scientists, and other corporate managers. When an AI system evolves beyond its initial programming through self-learning processes, the link between original code and eventual harm becomes attenuated. The requirement of proximate causation becomes difficult to satisfy.[32]

Users and operators constitute the second potential locus of liability. Where AI is intentionally weaponized, for example, to conduct fraud or generate defamatory content existing criminal statutes remain fully applicable. In such instances, the AI operates merely as an instrument of human intent.[33] Nevertheless, when AI systems make independent decisions contrary to expectations of the users, this leads to complexity. When a system that has been installed on lawful grounds has an independent result of unlawful product, criminal intent attribution on the user will be a problem. Because generative systems are unpredictable, it is possible to defend based on lack of foresight. The user can say that the illegal act was not targeted or intended and can undermine the case of the prosecution.

The most disruptive potential is in reference to the system of AI. With AI models growing close to resemblance in human-like decision-making, there is some concern on whether they ought to assume some sort of independent accountability. Under Indian law, however, AI lacks legal personhood. It cannot be subjected to punishment, imprisonment, or moral condemnation.[34]

The law system is still anthropocentric, which is aimed at the agents of the human kind and, in some cases, corporate entities. A fundamental reorganization of the basic legal notions, such as personhood and punishment, would be needed to acknowledge AI as a distinct criminal responsibility bearer. This kind of recognition is controversial and speculative. Liability is in no single place and everywhere at the same time. Developers can make disclaimers, users can claim unpredictability, and the AI system cannot be prosecuted. The absence of definitive judicial precedents in India addressing AI-specific criminal liability underscores that this area remains in formative development.[35] As long as no clearer statutory or doctrinal principles are

[32] H.L.A. Hart, *Punishment and Responsibility* 28–53 (2d ed. 2008).
[33] *Bharatiya Nyaya Sanhita*, No. 45 of 2023 (India); *Information Technology Act*, No. 21 of 2000 (India).
[34] *Salomon v. A. Salomon & Co.*, [1897] A.C. 22 (H.L.).
[35] Ministry of Electronics & Info. Tech., Gov't of India, *Consultation Paper on Artificial Intelligence Governance Framework* (2025).

available, the allocation of criminal responsibility will be disputed.

VI. EVOLVING JURISPRUDENCE: PRECEDENTS AND EMERGING CASES

The jurisprudence of AI and criminal responsibility is still in its infancy in Indian jurisprudence. The judiciary has been cautious in situations where AI systems guide the court system. Rather than constructing new doctrines of liability, judicial responses have focused on safeguarding procedural integrity and ensuring that AI outputs do not compromise fairness.[36]

Globally, litigation involving AI has largely centered on civil disputes particularly copyright, data protection, and intellectual property.[37] Although these matters do not directly address criminal liability, they shape the broader discourse on AI accountability. A consistent judicial theme has been reluctance to grant AI independent legal standing or rights.[38] This reluctance is an indication of similar opposition to the application of AI systems with some form of independent criminal responsibility. Laws still bring AI outputs to bear on human agents or companies.

In India, the regulatory architecture has increasingly relied on compliance-based mechanisms. Instruments such as the Digital Personal Data Protection Act and amendments to Information Technology Rules impose financial penalties and content moderation obligations.[39] These are steps that focus on transparency and damage reduction. Nevertheless, they do not go the extra mile of redefining criminal responsibility. The emphasis is made on the enforcement of regulations instead of the doctrinal innovation in the penal law.

The absence of authoritative case law squarely addressing AI-induced criminal harm indicates a period of jurisprudential incubation.[40] Courts have not stated in details the criteria to determine intent or causation in AI-driven situations. As the AIs keep being adopted in the sphere of finance, governance, and the administration of communities, there is no doubt that the cases of the criminal liability clashing directly with it will increase. Once these cases are taken to even greater courts, they can lead to the clarification of the doctrines.

---

[36] Justice B.R. Gavai, Supreme Court of India, *Remarks on AI Use in Judicial Processes* (Public Address, 2023).
[37] *Thaler v. Perlmutter*, 43 F.4th 1207 (D.C. Cir. 2023).
[38] Id. at 1213.
[39] *Digital Personal Data Protection Act*, No. 22 of 2023 (India); *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules*, 2021, as amended 2024 (India).
[40] *Bharatiya Nyaya Sanhita*, No. 45 of 2023 (India); *Information Technology Act*, No. 21 of 2000 (India).

VII. REGULATORY GAPS AND POLICY IMPLICATIONS

India has a dynamic but yet an incomplete regulatory environment concerning AI-related risks. The country has shown positive action by addressing the issue with policy changes and amendments to digital governance regulations, but the country does not have a specific criminal law, focused on AI. It is a flexible strategy that does not have conceptual clarity. The use of retrofitting of old statutes creates doubt when it comes to autonomous system cases. Conventional criminal teachings assume foreseeability and intentions in man. When these doctrines are applied to algorithmic decision-making, they experience structural constraints.

Policy discussions in 2026 increasingly emphasize a principles-based governance model that integrates accountability and transparency into existing legal frameworks rather than constructing a standalone AI penal code.[41] Nevertheless, until more articulate explanations of how criminal responsibility is to be allocated in the situation involving autonomous systems, the regulation can only target symptoms, but not the causes.

The regulatory gap is particularly visible in high-risk contexts such as autonomous vehicles, algorithmic financial systems, and AI-driven misinformation campaigns.[42] In both of these areas, damage can manifest itself without the obvious intention of a human being. The lack of a customized system of criminal liability makes the courts dependent on creative interpretation instead of the clarity of the legislation.

VIII. ABSENCE OF AI-SPECIFIC LEGAL FRAMEWORK IN INDIA

There is a strong and decisive dichotomy in regulatory environment of artificial intelligence in India. On one hand, the State has actively promoted AI adoption across governance, industry, and public infrastructure.[43] On the other, the legal architecture governing the risks arising from AI particularly in relation to criminal responsibility remains fragmented and incomplete. As of early 2026, India continues to regulate AI primarily through the application of general statutes, supplemented by policy guidelines and ethical frameworks, rather than through a

[41] NITI Aayog, *Responsible AI for All*, supra note 3; AI Impact Summit 2026, supra note 2.
[42] Sec. & Exch. Bd. of India (SEBI), *Circular on Algorithmic Trading Controls* (2024); Ministry of Road Transp. & Highways, *Autonomous Vehicle Policy Discussion Paper* (2025).
[43] NITI Aayog, *National Strategy for Artificial Intelligence: #AIforAll* (2018).

comprehensive AI-specific legal regime.[44]

This structural gap is clearly apparent in a comparative analysis of the current legal tools. While statutes such as the Information Technology Act and general criminal laws continue to be applied to AI-enabled harms, there is no legislation that directly addresses the unique legal consequences of autonomous algorithmic decision-making.[45] Such a gap in the framework makes it harder to assign blame, as courts and implementation agencies have no obvious doctrinal reference to follow when AI systems are deployed in ways similar to those that are not controlled directly by humans.

Traditionally, the reaction of Indian regulating to AI has been one based on adaptation and not anticipation. Between 2020 and 2023, AI adoption expanded rapidly across sectors, yet this growth was not accompanied by the development of bespoke legislation tailored to AI-specific risks.[46] Existing legal provisions were also used in this period with the view to resolving new harms, which was a practical approach to the rate of technology change. Nevertheless, the absence of AI-related regulation at this stage of development was a major weakness, especially in the context of solving the problem of autonomous decision-making, lack of transparency, and decentralization of responsibility. Retrospective assessments conducted from 2024 onward confirm that while this adaptive approach mitigated immediate risks, it failed to pre-emptively address deeper structural challenges associated with AI governance.[47] Such a reactionary form of posture retrofitting of human behaviour to regulate autonomous systems cannot but restrain the legal system in terms of its ability to answer questions of essence about intent, causation and liability.

After 2023, there was a discernible change in the regulatory philosophy of India. Rather than pursuing a singular, codified AI statute, policymakers increasingly embraced a principles-based governance model.[48] This approach, prominently articulated in deliberations surrounding the AI Summit 2026, emphasizes flexibility, ethical deployment, and voluntary compliance within existing legal structures.[49] The model is grounded in core principles of accountability,

[44] Ministry of Electronics & Information Technology (MeitY), *Consultation Paper on Artificial Intelligence Governance Framework* (2025).
[45] *Information Technology Act*, No. 21 of 2000 (India).
[46] NASSCOM, *AI Adoption in India: Market and Policy Trends* (2023).
[47] Organisation for Economic Co-operation & Dev. (OECD), *Artificial Intelligence in Society* (2019).
[48] NITI Aayog, *Responsible AI for All: Approach Document* (2021).
[49] Ministry of Electronics & Information Technology, *India and the Global Partnership on Artificial Intelligence (GPAI) Summit Report* (2023).

transparency, and human-centric design, reflecting an effort to balance innovation with societal safeguards. This tendency of ethical and principle-based governance as opposed to strict legislation is an indication of India that the country aims at developing a regulatory pattern that can fit the fast pace of technological change and may be applicable to other developing economies.

In this dynamic framework, the State has proposed specific regulatory interventions to particular risks of AI. Amendments to information technology regulations in 2024 mandated disclosures and visible markers for synthetic and manipulated content, reflecting an early attempt to address AI-enabled misinformation.[50] More recently, the introduction of the Specific Guidelines for AI (SGI) represents a significant regulatory development. As of March 2026, these guidelines impose enhanced compliance obligations on both domestic and multinational firms, requiring risk mitigation strategies, certification mechanisms, and adherence to defined ethical standards.[51]

All these progressive moves notwithstanding, the fundamental regulatory gap still continues to exist. India is yet to establish a specific legal framework that can help in adjudicating criminal accountability that may result due to AI-guided behaviour. The current quilt of modified laws, codes of ethics, and industry-related regulations does not offer conclusive responses to some basic questions of jurisprudence. When self-directed AI systems produce harm, which would otherwise be deemed a criminal act, e.g. fatal accidents of autonomous vehicles or massive fraud carried out by algorithmic financial networks, the law system faces a problem of accountability gap. Established doctrines of criminal liability, particularly *mens rea* and *actus reus*, are inherently premised on human cognition and volition.[52] Courts have consistently held that criminal responsibility generally requires a culpable mental state.[53] The principle of strict liability has also been invoked in situations where liability arises despite the absence of intention or negligence.[54] The application of these doctrines to the non-sentient systems necessitates legal fiction, which the current penal law is poorly placed to maintain.

[50] *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules*, G.S.R. 139(E), Gazette of India (Feb. 25, 2021) (as amended 2024).
[51] Ministry of Electronics & Information Technology (MeitY), *Advisory on Artificial Intelligence and Generative AI Platforms* (2024).
[52] K.D. Gaur, *Textbook on the Indian Penal Code* 112–15 (6th ed. 2016).
[53] *Nathulal v. State of Madhya Pradesh*, AIR 1966 SC 43.
[54] *Rylands v. Fletcher*, (1868) LR 3 HL 330.

Through this, the attribution of liabilities is still very disputed. The issue as to whether the developer, manufacturer, deployer or the end-user should be held responsible is a question of legal ambiguities and not of doctrine. Although the existing regulatory practices help in protecting consumers on the one hand and facilitating ethical AI design on the other, they still lack a predictable and enforceable system of criminalizing the culpability of autonomous systems. This has weakened the aims of deterrence and legal certainty and at the same time subjected innovators to legal uncertainty.

Corporate criminal liability principles provide a useful analogy when examining responsibility for AI-driven harms. In *Tesco Supermarkets Ltd v. Nattrass*,[55] the House of Lords articulated the identification doctrine, holding that the acts and mental state of senior management may be attributed to the corporation itself. This doctrine is frequently invoked to determine liability in cases where decision-making structures are diffused- an issue that similarly arises in AI-based systems where responsibility may lie with developers, deployers, or corporate entities controlling algorithmic operations.

To fill this void, India needs a proactive legal response that comes up with definite ideals on criminal responsibility of AI without compromising on innovation. This intervention should be able to balance humanistic principles behind criminal law and the imperative realities of autonomous systems. Until this reconciliation is achieved, the legal framework of AI in India will be structurally incomplete, trying desperately to keep pace with the pace of technological change and the needs of criminal justice.

IX. COMPARATIVE INTERNATIONAL APPROACHES TO AI LIABILITY

The issue of regulatory challenges of artificial intelligence does not belong to India alone. In various jurisdictions, policymakers and legal theorists have started considering ways to resolve the legal consequences of autonomous systems, in particular, accountability, transparency, and attribution of liability.

The European Union Artificial Intelligence Act is one of the most notable innovations that assumes a risk-based policy. The Act categorises AI systems based on the harm that they may cause to society and provides developers and deployers with different regulatory duties. There

[55] *Tesco Supermarkets Ltd. v. Nattrass*, [1972] AC 153 (HL).

are strict requirements of high-risk AI systems in terms of transparency, human control, and risk control. Although the EU framework primarily focuses on regulatory compliance rather than criminal punishment, it demonstrates how structured legal oversight can mitigate risks arising from autonomous decision-making.[56]

On the same note, the international organizations have developed ethical governance systems to direct the national regulatory approaches. The OECD Principles on Artificial Intelligence emphasize accountability, transparency, robustness, and human-centered design as foundational norms for responsible AI deployment.[57]

Another significant global instrument is the UNESCO Recommendation on the Ethics of Artificial Intelligence, which establishes a comprehensive ethical framework addressing human rights, algorithmic bias, and transparency in automated decision-making.[58] The recommendation stresses the importance of auditing and explainability in AI systems and acknowledges that mechanisms to hold accountable people to ensure that people trust the algorithms are needed.

Collectively, these global efforts demonstrate a new global order of reality: successful AI regulation should be based on both moral protection and enforceable rules. Although India has been adopting a number of policy frameworks, which are consonant with these principles, it has not translated these principles into a full-scale statutory regime that covers criminal responsibility over AI-driven damages.

X. ETHICAL CONSIDERATIONS: ACCOUNTABILITY AND TRANSPARENCY IN AI SYSTEMS

Ethical governance has been one of the key pillars of the AI strategy in India. The emphasis on accountability and transparency reflects recognition that legal enforceability depends on technical traceability.[59] Attribution of responsibility is virtually impossible without mechanisms of explainability, audit logs and documentation standards.

[56] 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).
[57] Organisation for Economic Co-operation and Development (OECD), OECD Principles on Artificial Intelligence(2019).
[58] UNESCO, Recommendation on the Ethics of Artificial Intelligence (2021).
[59] Organisation for Economic Co-operation & Dev. (OECD), *OECD Principles on Artificial Intelligence* (2019).

Transparency obligations, including disclosure of AI-generated content and platform responsibilities, enhance public trust and investigative capacity.[60] It makes investigation easier, but it fails to answer the philosophical dilemma of whether algorithmic autonomy or moral culpability is more relevant. India's alignment with international ethical standards underscores its commitment to responsible AI deployment.[61] However, enforcement and judicial interpretation shall become the final test of these structures. The moral obligations are required to translate into legally recognizable obligations in case they are to bear any impact on the criminal adjudication.

XI. STRENGTHENING THE LEGISLATIVE RESPONSE

In order to increase the activity aimed at optimizing the laws and regulations, one should close the existing gap between the compliance system on the platform and individualized criminal responsibility. Implementation of regulations must not be content moderation but be more reflective of apportionment of liability in using high risk AI. It should be a sensitive methodology. Much criminalization can deter innovation, and no regulation can promote impunity. This is to create a balance of responsibility in that regard, that in instances where the deployment is malice there should be penal action taken whereas in instances where the deployment is made in good faith protection should be granted to the innovation.

The Indian criminal law must strike a balance between the humanistic origins of the criminal law and the reality of the algorithmic agency. Laws will continue to rise in opposition to the technology development and the development of the law until the principles of intent, causation, and foreseeability are modified to match autonomous systems.

XII. CONCLUSION

The quick penetration of artificial intelligence into government, business, and social infrastructure has essentially changed the character of decision-making in the modern society. But the structural principles of Indian criminal law are still entrenched in the belief that malevolent behavior is instigated by human beings who already have the intent, knowledge, or carelessness. The assumption discussed in this article is more and more challenging to maintain

[60] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 (*Artificial Intelligence Act*).
[61] UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (2021).

in the case when autonomous algorithmic systems can generate results which can cause considerable harm without direct human intervention. The current system of criminal law based on the concepts of *mens rea* and *actus reus* are inadequately adapted to situations where the blame might be distributed among complicated technologies in which developers, deployers, and users are involved.

The lack of a unified legal framework to deal with the damages brought about by AI has left a legal void in which control over responsibility becomes indeterminate. The existing dependence on generic criminal law and information technology policies is not a full-fledged solution, because they were developed to govern human actions and behaviours, but not autonomous computer-based functions.

To solve these issues, it is necessary to build a systematic structure of the law based on the facts of artificial intelligence. To begin with, the responsibility in relation to AI systems must be distributed between developers, deployers and users by determining a clear liability architecture according to the extent of their control over a system. Developers and deployers can be responsible respectively of poor design or careless program implementation, and deployment failures by deployers who do not take reasonable precautions during operational deployment. The second one is that when dealing with high-risk AI systems that can be used in sensitive areas like healthcare, finance, and transportation, one can assume a strict liability model to provide sufficient protection against disastrous damages.[62] This would acknowledge that some technologies have systemically risky characteristics which need greater legal regulation. Third, it might be useful to create a dedicated AI regulatory agency that will monitor, certify, and investigate the harms associated with algorithms in India. The institution would be able to align technical skills and legal enforcement tools, thus enhancing accountability as well as responsible innovation.

After all, the problem of artificial intelligence is not only a technological issue but also a legal one. The Indian criminal law should advance in a way that its human base is adjusted to the new reality of the autonomous decision-making systems. It will be crucial to design a consistent doctrine of criminal responsibility in relation to AI to make sure that the technological advancement is not set ahead of the legal system to safeguard the interests of the society.

[62] Gabriel Hallevy, *When Robots Kill: Artificial Intelligence Under Criminal Law* 177–201 (2013).