CRIMINAL LIABILITY FOR AI HARM: POSITIONING INDIA'S BHARARTIYA NYAYA SANHITA WITHIN GLOBAL LEGAL TRENDS

***Mukul Lakra*[1]**

VOLUME 2, ISSUE 1 (JANUARY-JUNE 2026)

ABSTRACT

The burgeoning growth of artificial intelligence (AI) presents some novel problems regarding the criminal theory itself – that is, concerning the issues of attribution, culpability, and causation. Conventional theories of criminal law, based on individual human conduct and intentions, have become inadequate for dealing with harm caused through autonomous and semi-autonomous machines. This paper will evaluate whether the Bharatiya Nyaya Sanhita is equipped to tackle such problems or not, as compared to international perspectives on the same. Using a doctrinal and comparative research approach, the paper begins by examining the challenges posed by the use of AI technologies in terms of their opacity, unpredictability, and distributed decision-making processes.

Next, the paper examines the application of existing principles of liability according to the Bharatiya Nyaya Sanhita, emphasizing notable shortcomings in dealing with algorithmic harms. The comparative examination of the regulatory strategies adopted in important jurisdictions such as the EU, US, and China reveals an emerging trend toward hybrid liability schemes. From these discussions, the authors present a proposal for a new concept of criminal liability in India that takes into account the multiple actors involved, the design of the algorithm, and changes in the standard of foresight. They recommend several specific measures, such as creating offenses specific to AI, employing hybrid forms of liability, and building capacity in institutions and investigations. The paper concludes that although the Bharatiya Nyaya Sanhita offers a basic structure, its practical implementation for crimes involving AI will necessitate not only legal

[1] Mukul Lakra, *Research Scholar at University School of Law & Legal Studies, GGSIPU*

reform but also regulatory harmonization. Through the comparative approach taken by the study, it offers an important contribution to international debates on AI governance.

Keywords: Artificial Intelligence, Criminal Liability, Bharatiya Nyaya Sanhita, Comparative Criminal Law, Legal Reform

INTRODUCTION: AI HARM AND THE CRISIS OF CRIMINAL LIABILITY

The integration of artificial intelligence (AI) into various essential domains such as finance, healthcare, transport, and communications has led to a fundamental shift in the concept of risk and harm in modern societies. While previous technologies were merely tools that remained under direct human control, modern AI technologies have developed autonomous capabilities, learning abilities, and opacity in their operation. This has profound effects on criminal law, which has always been concerned with identifying human actions and intentions. However, when the risk of harm arises from an algorithmic process that cannot be predicted or understood, the principles of criminal law are challenged.[2]

Criminal responsibility is founded on the interaction between actus reus and mens rea, which requires the existence of a human being who has the capacity for both the act and the intent.[3] With AI systems, however, this relationship is complicated. In instances where the autonomous technology causes harm, whether due to its design, biased algorithms, or unexpected actions, the issue becomes how one can determine whether there is a guilty mind in relation to the crime. This could be the programmer, the installer, the user, and the organization itself that created the system. However, none may meet the conventional criteria for establishing guilt. Such a situation creates what is commonly referred to as an "accountability gap."[4]

It poses unique challenges in jurisdictions that are witnessing fast-paced changes in their laws. For example, in India, the new codification of its criminal laws via the Bharatiya Nyaya Sanhita (BNS) is a classic example of such doctrinal change. The BNS aims at bringing about an updated and

[2] Mireille Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* (Edward Elgar 2015) 102.
[3] Andrew Ashworth, *Principles of Criminal Law* (7th edn., Oxford University Press 2013) 85.
[4] Matthias C. Kettemann and Konrad Lachmayer (eds.), *Freedom of Expression in the Internet Age* (Springer 2019) 214.

efficient version of substantive criminal laws but still continues with traditional notions of mens rea and actus reus. The negligence and fraud under the BNS did not intend to include any sort of autonomy or partial autonomy for autonomous systems. Thus, how are we to treat the application of these laws to AI?[5]

Such issues are not only relevant to India. Around the world, legal frameworks have had to deal with such problems, although employing different regulatory measures. For example, the EU has chosen a risk-based regulatory model where ex ante regulation and compliance prevail, whereas the US has employed doctrines already present in its legislation along with sector-based regulation. Lastly, China has opted for a centralized regulatory approach, focusing on state supervision and control.[6] Nevertheless, the underlying point is similar, as there is a clear understanding of the limitations of criminal law when dealing with the challenges brought about by AI technology. Such international developments will help analyze India's situation better.

Given this background, the present paper attempts to critically analyze whether the existing doctrines of criminal liability under the BNS suffice for addressing criminal liability arising out of harms associated with AI. The paper will be premised on the assumption that although the doctrines of criminal liability are able to address some issues, they fail to address all the concerns associated with the harms connected with AI. The main research questions that will guide the analysis include: (i) the extent to which criminal liability under the BNS may be extended to cover AI harms; and (ii) how criminal law in India may respond to global developments.

From the methodological point of view, the article employs doctrinal and comparative methods of analysis. It involves an analysis of selected provisions of the BNS, together with the consideration of judicial concepts related to culpability and responsibility. In addition, a comparative analysis of selected international standards is conducted to detect any existing patterns and practices in this area. Notably, the focus here is more than merely descriptive in nature.

However, the subject matter of the article is restricted to the area of criminal liability only, and the article does not deal with any procedural or evidential aspects, unless they have a bearing on the

[5] K.D. Gaur, *Textbook on Indian Penal Code* (6th edn., Universal Law Publishing 2016) 57.
[6] Cary Coglianese and Alicia Lai, "Algorithmic Regulation: Technology, Governance, and Legal Design" (2022) 89 *Geo. Wash. L. Rev*. 101, 118.

question of attribution or liability. While the analysis may take into account global trends, it is primarily confined within the ambit of Indian law. The aim of this article is to examine the relationship between AI and criminal liability with the help of the BNS framework.

CONCEPTUAL FOUNDATIONS: AI, HARM, AND CRIMINAL RESPONSIBILITY

The discussion about criminal responsibility for harm caused by artificial intelligence should start with conceptual precision. Since there is no consistent legal definition of what "harm caused by artificial intelligence" means, it often creates uncertainty and lack of consistency in legal doctrine. At a minimum, harm caused by artificial intelligence can be defined as harm to one's physical well-being, finances, reputation, and psyche stemming from the functioning of an artificial intelligence technology, where the functioning includes some level of autonomy or decision making independent of any immediate human control.[7]

One of the hallmarks of AI technologies is their dependence on machine learning algorithms, which make it possible for them to adapt and improve upon the initial programing. Neural network machine learning algorithms can recognize certain patterns and make decisions that may be impossible for the programmers themselves to understand.[8] The inability to comprehend how an algorithm makes certain decisions—a problem known in machine learning circles as the "black box"—causes serious difficulties in legal analysis. For instance, in a traditional case of criminal law, there must be an intelligible relationship between the conduct and the harm that ensued.[9]

Autonomy, too, is closely connected to opacity. Although AI is not conscious and morally accountable, its functional autonomy can give rise to consequences that are similar to those of autonomous acts. For example, an AI program used for trading might act at such fast and large scales that it cannot be controlled in real time by humans. The result could be manipulation of the

[7] Ryan Calo, "Artificial Intelligence Policy: A Primer and Roadmap" (2017) 51 *U.C. Davis L. Rev*. 399, 404.
[8] Ian Goodfellow, Yoshua Bengio and Aaron Courville, *Deep Learning* (MIT Press 2016) 27.
[9] Frank Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information* (Harvard University Press 2015) 3.

market or monetary loss. Likewise, the AI algorithms responsible for content creation might spread defamation or disinformation without direct input from any human being when publishing it.[10]

These features imply that there must be an analysis of the principles on which criminal law rests. According to one of the main criteria of actus reus, the crime must comprise an act or failure to act caused by a person voluntarily.[11] In the case of AI, however, such an act may encompass many actions performed throughout the entire cycle of designing, programming, training, and using artificial intelligence. It also means that many people may be involved in the process at various stages.

The issue of mens rea is another problem to consider. Usually, for a person to be criminally responsible, there must be an intent, knowledge, recklessness, or negligence in his actions.[12] The AI lacks consciousness; hence, it cannot be regarded as having any mental state. It is then necessary to ask whether mens rea can be attributed to the human agents who designed or utilized the AI. One plausible option is to attribute culpability to the designing and utilization stage when the risk of foreseeing harm was disregarded. Nonetheless, this approach becomes difficult if the harm arose from emergent properties that were unforeseeable.[13]

The next problem relates to defining the difference between fault and culpability. It must be noted that all the adverse consequences that may be produced by AI systems do not necessarily involve any negligent behavior or wrongful conduct. Certain cases might be related to statistical irregularities or unexpected interactions among different data sets. Nevertheless, criminal law does not provide for the punishment of any accidental occurrence unless it falls into one of the limited types of negligence.[14] The broadening of liability for all kinds of AI-caused damages will likely lead to the undermining of the moral basis of punishment, since the latter presupposes the presence of culpability.

[10] Woodrow Barfield and Ugo Pagallo (eds.), *Research Handbook on the Law of Artificial Intelligence* (Edward Elgar 2018) 45.
[11] Glanville Williams, *Textbook of Criminal Law* (2nd edn., Stevens & Sons 1983) 34.
[12] Andrew Ashworth, *Principles of Criminal Law* (7th edn., Oxford University Press 2013) 137.
[13] Thomas Burri, "The Politics of Robots: Regulation and the Future of Law" (2017) 48 *Common Market Law Review* 343, 356.
[14] H.L.A. Hart, *Punishment and Responsibility* (2nd edn., Oxford University Press 2008) 28.

Attribution in itself becomes a more difficult issue because of the involvement of multiple parties in the life cycle of an AI system. For example, developers, data suppliers, implementers, and users could each play a role in the formation of the circumstances that lead to harm. It is thus difficult to pinpoint responsibility using a purely individualist approach that is favored by criminal law.[15] Though doctrines of common intention and conspiracy provide some leeway, it is also not easy to apply them to the actions of multiple parties within the AI sphere.

Given these problems, academics have recently argued for the need to rethink criminal responsibility in the era of artificial intelligence. In their proposals, some authors suggest adapting current legal doctrines by employing analogical reasoning, whereas others support the creation of new theories of liability that are able to take into consideration the nature of algorithms.[16] One of the key issues in such academic debate is that of individual fault versus collective and/or risk-oriented forms of accountability. This issue has important consequences not only from a legal point of view but also from those of fairness and deterrence.

CRIMINAL LIABILITY UNDER THE BHARATIYA NYAYA SANHITA

The coming into force of the Bharatiya Nyaya Sanhita (BNS) of 2023 signifies an important step taken in the process of modernizing and rationalizing Indian substantive criminal law, rooted deeply in its colonial heritage. Nevertheless, even with all of the improvements brought about by the new code, the BNS is deeply entrenched in classical doctrines of criminal liability predicated on the presence of human volition, intention, and capacity for control. The following analysis will discuss the ability of existing provisions of the BNS to cope with harm caused by artificial intelligence (AI) technologies.

From a structural perspective, the BNS preserves the basic structure of criminal liability predicated on actus reus and mens rea, with crimes being classified depending on the type of harm caused as well as the culpability of the offender.[17] The various statutes concerning negligence, fraud, and crimes against the body and property serve as the foundation for the whole legal structure.

[15] Luciano Floridi et al., "AI4People—An Ethical Framework for a Good AI Society" (2018) 28 *Minds and Machines* 689, 695.
[16] Mireille Hildebrandt, *Smart Technologies and the End(s) of Law* (Edward Elgar 2015) 156
[17] K.D. Gaur, *Textbook on Indian Penal Code* (6th edition., Universal Law Publishing 2016) 49.

Although these laws are broad enough to cover almost any form of harmful behavior, they cannot be easily applied in cases where the perpetrator is an artificial intelligence. This is because there are problems when trying to fit algorithmic actions into human legal categories.

It might be possible to tackle the issue of harm caused by AI from a negligence perspective within the BNS framework. Historically, criminal negligence is described as any failure on the part of an individual to meet his or her duty of care, leading to the occurrence of some form of foreseen damage.[18] Where AI is concerned, negligence may include such acts as failure to properly evaluate, monitor, or improve the AI system, hence enabling harmful consequences to take place. An example in this regard would be deploying an autonomous system that lacks appropriate safeguards to ensure safety. However, applying negligence principles to AI can pose difficulties, especially regarding the issue of defining what is reasonable care in such cases.

Cheating and fraud, which are acts of deception, are another category of offenses that could serve as a potential means of addressing harm caused by some types of AI systems. It should be emphasized that AI technology could be used to achieve extremely effective acts of deception, misrepresentation, and manipulation, all of which require a minimum of supervision.[19] In these instances, one could argue that liability for the offense falls on the individual who uses the AI technology as an instrument of deceit. However, the problem arises when the act performed using AI technology occurs spontaneously without any desire or anticipation on the part of the person responsible.

In the same manner, rules pertaining to damage against persons and property can apply to instances where damage has been caused by artificial intelligence technology to persons or financial assets. For instance, an autonomous car that causes an accident and an automated stock trading application that causes monetary losses can potentially fall under an offense that already exists. But determining responsibility is still challenging when the harm comes from artificial intelligence.

[18] Ratanlal & Dhirajlal, *The Indian Penal Code* (34th edn., LexisNexis 2017) 312.
[19] Arvind Narrain and Apar Gupta, "Technology, Crime and the Law in India" (2020) 5 *Indian Journal of Law and Technology* 1, 9.

While in ordinary offenses, the problem of causation can easily be traced back to the person who performed the act, in the case of AI, there are many variables that make causation complicated.[20]

Another problem associated with the BNS is the lack of mention of specific liabilities arising from AI. While new regulatory systems in some other countries incorporate such aspects into legislation, the BNS ignores the possibility that autonomous systems could be involved in the perpetration of crimes. As a result, the issue would have to be resolved by drawing analogies, which, although flexible, could create problems when enforcing the law. There is a danger that the courts could either over-interpret some provisions and engage in over-criminalization, or take a conservative approach.

Interpretation of law by the Indian judiciary has always been an important aspect when making changes to the criminal laws according to the changing times. It is quite common for the Indian judiciary to make use of purposive interpretation to expand the scope of a particular section of a law according to the realities of the time.[21] In the context of artificial intelligence, however, the issues involved are more complex than ever before. Not only is it a matter of dealing with a new form of conduct, but also a new form of agency.

Another challenge pertains to determining the correct subject for liability. As with criminal law, the BNS has a predominantly individualistic approach to determining culpability, even if corporate liability is acknowledged in some instances. Nevertheless, the BNS falls short in accommodating the dispersed and collective process involved in developing and deploying AI systems. In such cases, developers, data scientists, corporations, and end users could be considered part of the processes leading to the operation of an AI system, but the BNS lacks clarity on how to allocate liabilities among these different subjects. There is a danger that liability may either be under-enforced, in which case none of these parties qualify as liable, or over-enforced, with liability being imposed on subjects that are marginally linked to the act.

Moreover, apart from the issues regarding doctrine, there also exist problems in terms of the practical aspect of pursuing a case for an AI crime. Proving causality and establishing the

[20] Glanville Williams, *Textbook of Criminal Law* (2nd edn., Stevens & Sons 1983) 38.
[21] *State of Maharashtra v. Mohd. Yakub* (1980) 3 SCC 57.

responsibility would require some level of technical know-how and access to the data within the system, both of which might not necessarily always be available. Moreover, the problem related to the opacity of AI makes proving causality particularly challenging due to the difficulty of tracing the decision-making process.

COMPARATIVE GLOBAL APPROACHES TO AI CRIMINAL LIABILITY

The problem of criminal responsibility for damages caused by artificial intelligence is not exclusive to India but is a common dilemma faced by the international community in adapting legal principles to fast-evolving technologies. Various countries around the world have taken varying stances on regulating AI and formulating laws based on their specific priorities and constitution, as well as the technology infrastructure that prevails in their respective societies. An examination of such policies helps understand how criminal law may respond to AI-related damages and serve as a point of comparison for India's Bharatiya Nyaya Sanhita.

In recent times, the European Union (EU) has come out as one of the leading jurisdictions in terms of the regulation of AI, adopting an ex-ante risk-based approach in regulation. The European Union does not make any changes in the existing criminal law in the initial stages of implementation but focuses on the development of the ex-ante obligations for high-risk AI, such as risk assessment and compliance obligations.[22] This method ensures that any harm is prevented prior to the occurrence by making it unnecessary for criminal sanctions to be imposed. Nonetheless, the European Union does not totally eliminate criminal liability but uses indirect methods to incorporate it into the process.[23] When there is a breach of regulations regarding the use of AI systems, then criminal liability becomes applicable, especially when negligence or recklessness is proven under the national laws of the member countries.

In contrast to the above, the United States has opted for the use of traditional doctrines to govern AI, with the country using a decentralized and sectoral approach in its AI regulation policy. Traditional crimes of fraud, negligence, and product liability have been used in addressing criminal liability relating to harms by AI technology, with no specific criminal laws governing AI being

[22] Karen Yeung, "A Study of the Implications of Advanced Digital Technologies for the Concept of Responsibility within a Human Rights Framework" (2018) European Commission Report, 12.
[23] Andrea Bertolini, "Artificial Intelligence and Civil Liability" (2020) 6 *European Parliament Study* 45, 61.

adopted.[24] Flexibility is one of the advantages that characterize this model, with courts capable of utilizing traditional doctrine in a more innovative way. However, one of the disadvantages of this model includes the lack of unity in regulating AI, with different sectors and states having diverging approaches. Lack of a common federal regulation of AI is another challenge with the adoption of this model, especially in the event that the issue crosses borders.

The Chinese approach differs vastly from the others because of its heavy emphasis on government involvement and centralized regulation. Within the Chinese legal framework, governance of AI is embedded within other forms of digital regulations, taking into consideration aspects of security and social stability, amongst others.[25] From the perspective of criminal law, there tends to be accountability placed on the people and companies responsible for the use and development of AI technology, especially when it poses a threat to national security. The state-led approach makes it possible to enforce regulations quickly without ambiguity; however, one of its major flaws is that it does not take into consideration the aspect of due process. Unlike in the EU and US, China prioritizes collectivism over individualism.

In addition to the previously mentioned important jurisdictions, there is also a general understanding that international cooperation needs to be enhanced in dealing with AI-associated crimes. The very nature of AI technology that is typically developed, applied, and used in one jurisdiction while having adverse effects in another jurisdiction presents difficulties from the perspective of territorial jurisdiction.[26] Existing frameworks such as MLATs and extradition treaties may prove insufficient in the face and complicated character of these crimes.[27] Consequently, a discussion on the possibility of developing international standards or even conventions related to AI governance in general and criminal law aspects, in particular, emerged.

The comparative analysis of the presented frameworks indicates certain trends in this regard. The first trend is the transition from punishment-oriented approaches to the use of preventive regulation with its focus on risk management and compliance. Another trend can be seen in the lack of fault-

[24] Ryan Calo, "Robotics and the Lessons of Cyberlaw" (2015) 103 *California Law Review* 513, 540.
[25] Rogier Creemers, "China's Social Credit System: An Evolving Practice of Control" (2018) SSRN Working Paper, 8.
[26] Dan Jerker B. Svantesson, *Solving the Internet Jurisdiction Puzzle* (Oxford University Press 2017) 67.
[27] Ugo Pagallo, *The Laws of Robots* (Springer 2013) 178.

based approach used in some cases and the development of certain hybrid frameworks with strict or risk-based liability elements. Finally, the third trend is associated with different approaches applied by various countries towards the role of the state in regulating and addressing autonomous vehicles. Different views on this matter emphasize the lack of universal model to address artificial intelligence harm.

In this respect, these international trends present both advantages and disadvantages to India. Firstly, the flexibility of the BNS system enables some aspects of the comparative models to be included in the system, especially concerning negligence and corporate liability. Secondly, the lack of a coherent regulatory regime pertaining to AI technologies constrains the application of the criminal law system to govern the issue effectively. In this context, the EU approach of regulating ex ante implies the significance of establishing relevant standards, and the U.S. example indicates the dangers that may arise out of such an approach. Finally, despite its success in several ways, the Chinese model might not be consistent with India's constitutional principles.

Finally, the comparative analysis highlights the necessity for a balanced approach towards criminal liability with respect to AI in law. Although there is no ready-made solution to incorporate into the legal regime of India, important insights gained from analyzing foreign models could assist in formulating a new framework that will take into account the needs for innovation, effectiveness, and fairness in criminal liability in an age of technological progress.

RECONCEPTUALIZING CRIMINAL LIABILITY UNDER INDIAN LAW

It can thus be seen that the current approach towards criminal liability, even when reformulated through the Bharatiya Nyaya Sanhita, faces inherent obstacles when applied to cases of injury caused by AI. This is not due to the lack of a legal basis but because of the inherent principles of personal responsibility and cause and effect within the legal system. The solution to dealing with such injuries in India's criminal law will thus need to incorporate both an understanding of the problem and a novel approach to liability.

One such approach could begin with expanding the interpretation of doctrinal principles under the current legal system. Indian courts have often been found willing to pursue purposive and

contextual interpretations of statutes, especially as social and technological advances demand it.[28] This flexibility can be used to broaden the application of well-known principles, like negligence and recklessness, to include risks related to artificial intelligence. For example, the principle of "reasonable care" can be modified to include responsibilities pertaining to AI, including testing algorithms, avoiding biases, and ongoing monitoring. While this process does not necessarily necessitate any new legislation at once, it does rely on judges' understanding of AI technology and its capacity for doing harm.

Interpretive broadening will not be enough in instances when harm results from a complicated series of events that cannot be attributed to one particular flaw in the process. In such circumstances, it is necessary to factor in some aspects of design-based and due diligence liability into the criminal law framework. The reason is that responsibility can also be placed for failure on the part of the developer of an AI system in terms of the way the system is designed, trained, or governed.[29] For instance, a developer who uses an AI system in ways that can foreseeably lead to harm without placing any checks on its use can be considered liable.

Another aspect of reconceptualization in relation to the above point is that of multi-party responsibility. As pointed out above, AI technologies cannot be attributed to any one person alone; instead, they are developed as a result of collaborative efforts among developers, data suppliers, companies, and users. The traditional individualistic approach of criminal law is inadequate to address the above scenario.[30] While the Indian legal system offers certain provisions for joint responsibility in criminal law, such as common intention and abetment, it is assumed that some form of cooperation exists between the parties. This assumption is likely to be invalid when dealing with the complexities associated with AI technologies.

In addition, there is the issue that should be addressed regarding the balance between innovation and liability. The liability regime that goes beyond what is necessary could potentially hinder technological innovation and the deployment of technology that could benefit from the use of AI in such areas as health care and public administration. On the other hand, lax liability standards

[28] *State of Maharashtra v. Mohd. Yakub* (1980) 3 SCC 57.
[29] Ugo Pagallo, *The Laws of Robots: Crimes, Contracts, and Torts* (Springer 2013) 121.
[30] Luciano Floridi et al., "AI4People—An Ethical Framework for a Good AI Society" (2018) 28 *Minds and Machines* 689, 696.

could leave people exposed to risks without sufficient legal protection.[31] The problem is that a delicate balance should be struck by retaining the functions of law, without hindering innovation.

Another significant aspect relates to the idea of foreseeability in relation to establishing criminal liability. For classical offenses of negligence, liability depends on the foresight of risk by the accused as well as the taking of measures that would prevent such risk. In the case of AI, however, the notion of foreseeability becomes problematic due to some level of unpredictability in the operation of these machines.[32] Instead of doing away with foreseeability completely, it would be better to look into the issue through the lens of current technology. Here, the question is whether a similarly situated but technically knowledgeable person should have foreseen the risk and prevented it from happening.

Similarly, institutional capacity also plays an essential role in any attempt at reconceptualization. Not only does a well-defined theory of criminal liability need to be articulated; there also needs to be sufficient institutional capacity to enforce that definition. In other words, it is necessary for legal actors to have some knowledge about algorithms and how they work. Otherwise, it will prove impossible for these actors to be able to properly implement any liability scheme.[33] Of course, these issues relate more to institutional capacity than substantive law, per se.

Lastly, any attempt at reconsidering criminal responsibility in light of the developments associated with AI will need to be accompanied by a wider discussion on what goals should be assigned to criminal law in such conditions. Criminal law has been performing several roles, such as deterrence, punishment, and society's condemnation of the criminal conduct of an offender. However, under certain circumstances related to AI, the goals of criminal law may have to be adjusted. The focus, therefore, could move towards prevention and regulation as opposed to punishment when harm is caused due to systematic failure and not individual fault.

REFORM PROPOSALS: TOWARDS AN AI-RESPONSIVE CRIMINAL LAW FRAMEWORK

[31] Cary Coglianese and Alicia Lai, "Algorithmic Regulation: Technology, Governance, and Legal Design" (2022) 89 *Geo. Wash. L. Rev.* 101, 130.
[32] Andrew Ashworth, *Principles of Criminal Law* (7th edn., Oxford University Press 2013) 168
[33] Frank Pasquale, *The Black Box Society* (Harvard University Press 2015) 189.

It is evident from the above analysis that although existing doctrines of the Bharatiya Nyaya Sanhita offer some form of justification for dealing with injuries caused by AI technology, they lack the necessary tools and mechanisms required to fully tackle the challenges posed by such injuries. This calls for more than mere interpretation; what is required now is a series of reforms that can effectively reformulate criminal law to fit the new realities of algorithmic decisions.

Among the essential aspects that need reform is the creation of offenses and other legal mechanisms tailored to dealing with AI-related harms. Due to the lack of clear definitions of such harm within current statutory law, any efforts towards punishing such offenders would lead to inconsistencies. Instead of trying to fit AI-induced harm into offenses whose definition does not necessarily cover such actions, it would be much better if certain types of AI-induced harm were created.[34] In this case, the offenses can focus specifically on harms whose current legal treatment is unsuitable.

On the other hand, the drafting of such measures must take into consideration that there is no room for over-criminalization in this respect. Criminal law, as a coercive mechanism, should always be subject to the tenets of proportionality and culpability.[35] Consequently, AI-related offenses can have graded liability standards, where deliberate use of AI systems, reckless use of such systems, and even negligence in using AI can have varying levels of sanctions imposed on offenders. In this manner, the legal system will maintain its moral integrity when it comes to the imposition of criminal sanctions.

The second reform strategy involves adopting a hybrid liability system which is an integration of the fault-based and risk-based liability models. It has been mentioned before that the fault-based model might not always work effectively when dealing with harm emanating from unpredictable behavior of artificial intelligence machines. However, the pure application of strict liability rules is likely to punish individuals without necessarily finding them morally responsible for their actions.[36] The problem can be addressed by adopting a hybrid model whereby liability is tied to the non-compliance of established standards of care or duties. The hybrid system entails a set of

[34] Ugo Pagallo, *The Laws of Robots: Crimes, Contracts, and Torts* (Springer 2013) 134.
[35] H.L.A. Hart, *Punishment and Responsibility* (2nd edn., Oxford University Press 2008) 53.
[36] A.P. Simester and G.R. Sullivan, *Criminal Law: Theory and Doctrine* (6th edition, Hart Publishing 2016) 92.

obligations that must be fulfilled by persons participating in the development and implementation of the technology.

Connected to this is the requirement to develop investigative and evidentiary methods for dealing with crimes involving AI technology. It is crucial to note that prosecuting cases that involve AI technology requires evidence to prove causality and identify actors who committed the crime. Unfortunately, current investigative procedures lack the capacity to meet these requirements.[37] The time has come for experts to develop specific forensic methods, such as auditing algorithms, tracking data movement, and reconstruction of decision-making processes. Besides, it may become necessary to review the evidentiary requirements, particularly regarding the admissibility of algorithms and expert testimony in analyzing AI technology.

Institutional reforms are also equally important in ensuring that any legal reforms achieve their intended goals. In order to ensure that there is proper enforcement of the criminal liability associated with AI technologies, it will require concerted action by many institutions, from the police to the judiciary.[38] Building capacity through training and establishing specific units responsible for handling cyber and AI-related crimes are key to ensuring that the actors in the legal process are technically equipped to handle the tasks involved.

Another critical aspect of reform involves the creation of a wider policy framework for the regulation of AI in India. Criminal laws, by their very nature, function as a form of reaction to wrongdoing after it has taken place. For this reason, it must be supplemented by precautionary measures that will help minimize the possibility of any harm being caused in the first place.[39] This involves creating a set of rules regarding how AI is designed and implemented, and ensuring that there is adequate supervision to ensure that such rules are followed.

Lastly, any attempt at reform must also incorporate the international dimension of AI governance. As AI operates internationally, domestic regulatory approaches cannot be pursued in a vacuum. It is imperative that India engages itself in international negotiations towards the creation of

[37] Frank Pasquale, *The Black Box Society* (Harvard University Press 2015) 201.
[38] Cary Coglianese and David Lehr, "Regulating by Robot: Administrative Decision Making in the Machine-Learning Era" (2017) 105 *Georgetown Law Journal* 1147, 1195.
[39] Karen Yeung, "Algorithmic Regulation: A Critical Interrogation" (2018) 12 *Regulation & Governance* 505, 512.

harmonious standards and mechanisms to deal with AI-induced harm.[40] This would involve, among other things, making domestic laws in line with evolving international standards, bilateral agreements, and participation in international principle-making relating to AI liability.

CONCLUSION

The fast developments in the field of artificial intelligence have revealed the inherent paradox in the realm of criminal law – an area of law originally concerned with regulating human behavior is now faced with harms arising out of processes carried out by semi-autonomous technologies. It is clear from the discussion above that even though the Bharatiya Nyaya Sanhita presents a well-structured doctrine based on traditional liability doctrines, it may not suffice to meet all the new challenges posed by artificial intelligence related offenses.

It was evident from the analysis above that provisions available in the BNS, to some extent, can be adapted into covering cases involving harms caused by AI through theories of negligence, fraud, and corporate liability. Nevertheless, this adaptation is bound to have limitations and will mostly be founded on analogies that might be lacking coherence. There being no specific provisions for AI crimes means that there is the potential of under-criminalizing and, at the same time, over-extending liability. Under-criminalizing refers to a situation where an activity that constitutes a crime does not get criminalized whereas over-extension means that there is imposition of liability without culpability.

Through the comparative analysis carried out in this article, it emerges that India is not unique in its experience with the problem at hand. Countries around the world have implemented different approaches to solving the matter, including the EU's approach based on regulation, the American legal doctrines of strict liability and negligence, and the centralized Chinese enforcement system. These different approaches may vary in their design and content, but they all indicate the same thing - that the conventional criminal law cannot be applied when it comes to harm resulting from AI without further regulation.

[40] Dan Jerker B. Svantesson, *Solving the Internet Jurisdiction Puzzle* (Oxford University Press 2017) 189.

Based on these findings, this paper suggests a new way to conceive of criminal liability within the Indian legal framework, one that requires moving away from rigid doctrinal classifications and towards a more adaptable and context-specific approach to criminal liability. This involves a broader interpretation of extant principles, adopting design and due diligence standards, as well as acknowledging the distribution of responsibility within artificial intelligence technologies. In no way does this suggest a departure from the foundational principles of criminal law, only an adaptation thereof.

The recommendations mentioned above in the previous section intend to implement these theoretical considerations. It is argued that the creation of dedicated offenses relating to artificial intelligence, the use of a hybrid model of responsibility, and enhanced capacity in terms of investigation and institutions are vital measures towards the construction of an appropriate criminal law regime for artificial intelligence. However, what should not be overlooked is that it is equally significant to incorporate the criminal law regime into an environment of artificial intelligence governance, which will focus on prevention and international cooperation.

In essence, therefore, the question is not whether there should be a confrontation between the criminal law and AI-related harms, but rather what sort of engagement would be appropriate in this regard. India's criminal legal system, by virtue of the Bharatiya Nyaya Sanhita, has the basic ingredients needed for such an engagement. Yet, making this possible demands a reconsideration of conventional wisdom and active engagement in the global dialogue regarding AI.

Given the constant evolution of AI technology and its ever-increasing presence in various spheres of human endeavor, the problems raised in the present paper are likely to become even more significant in the future. It remains to be seen how successful the development of criminal law will prove to be in addressing those problems through a combination of doctrinal coherence and institutional flexibility. From this point of view, India finds itself at a crossroads: it can choose to adopt the former strategy based on minor modifications to the existing regulatory framework or opt for the latter by designing a legal system more fit for the digital era.