Artificial Intelligence, Accountability and Criminal Responsibility: Regulatory Challenges for India's Digital Economy

MR. AMIT KUMAR, MR. AMRITANSHU SHEKHAR,
DR. ANAND SINGH PRAKASH, MR. ADITYA SHEKHAR,
DR. HEMANTA KUMAR PANDA, MS. ELISHA LAKRA,
DR. MAHAVEER PRASAD MALI

Abstract: *With the increasing use of artificial intelligence in sectors ranging from self-driving cars, health care diagnosis, financial markets, and even lethal autonomous weapons, there is a noticeable inadequacy in the current criminal liability framework. With the traditional doctrines of criminal liability based on actus reus and mens rea, coupled with the principle of individual responsibility, the current legal regime fails to provide answers to the challenges posed by autonomous machines which function with varying degrees of human involvement. The present paper will focus on the intricate issues pertaining to the criminal liability of AI-powered actions in the Indian legal regime. The paper then goes on to differentiate between those instances where AI is merely used as a tool, and responsibility is attributed to human agents like programmers, manufacturers, and users, and those cases where autonomous AI agents come into the picture, rendering attribution difficult owing to the "black box" problem and issues of culpability of intent. A comparison of legal personhood of AI in different jurisdictions is undertaken, with special focus on the proposal for electronic personhood in the European Union. A discussion of the difficulties faced in each domain, namely, autonomous vehicles, health care, white-collar crimes, and content creation, is followed by some suggestions for reforming Indian criminal law, which include the amendment of the Indian Penal Code and the Information Technology Act.*

Keywords: Artificial Intelligence, Criminal Liability, Autonomous Systems, Mens Rea, Actus Reus, Legal Personality, Indian Penal Code, Algorithmic Accountability, Strict Liability, AI Regulation, India.

Mr. Amit Kumar, Assistant Professor of Law, The ICFAI University, Jharkhand, email id - amitsingh49@gmail.com
Mr. Amritanshu Shekhar, Assistant Professor of Law, The ICFAI University, Jharkhand, email id - amritanshu211@gmail.com
Dr. Anand Singh Prakash, Assistant Professor of Law, The ICFAI University, Jharkhand, email id - anandair03@gmail.com
Mr. Aditya Shekhar, Assistant Professor of Law, The ICFAI University, Jharkhand, email id - aadityasfj@gmail.com
Dr. Hemanta Kumar Panda, Professor, The ICFAI University, Jharkhand, Email id - hemant.panda2012@gmail.com
Ms. Elisha Lakra, Assistant Professor of Law, The ICFAI University, Jharkhand, email id - elishalakra0@gmail.com
Dr. Mahaveer Prasad Mali, Assistant Professor of Law, Nims School of Law, Jaipur, email id – dr.mahaveerprasad@nimsuniversity.org

Introduction

The exponential growth in artificial intelligence technology has brought about a new age of unparalleled technological potential, transforming various sectors, governance, and even our personal lives. Self-driving cars traverse our streets; software conducts high-frequency trades in financial markets; artificial intelligence is being used for diagnosis in the medical field; and artificial intelligence is generating content that cannot be differentiated from human-generated content. Along with all of these incredible developments, there exists a pressing need within the legal world that the existing criminal law system cannot answer. Who is criminally responsible when the harm results from an autonomous machine? The Indian Penal Code of 1860 and the entire framework of criminal law are based on certain basic principles where a person acts in a manner that requires mens rea and actus reus.[1] However, in the context of artificial intelligence, especially highly autonomous AI, this premise breaks down. In the event that a self-driving car injures a pedestrian, a high-frequency trading software crashes a stock exchange, or a diagnostic artificial intelligence incorrectly diagnoses a patient leading to their untimely death, the causality chain is complex. However, the programmer who authored the code was unaware of the exact result, the producer may have been compliant with industry standards, the user may have trusted the program in good faith, and the AI itself cannot legally be treated as a criminal due to lack of human traits required for criminality. In other words, there may arise instances where harm occurs but there exists no person to hold responsible for it, thus leaving a vacuum in the justice system. Therefore, the time has arrived when it becomes imperative to reassess the criminal liability regime in the wake of increasing use of AI. India being a highly digitized country with ambitious plans regarding incorporation of artificial intelligence in its various fields, needs to deal with such concerns. The present paper tries to critically analyze the legal issues arising out of application of artificial intelligence to the doctrine of criminal liability, especially in the Indian context. First, it gives a brief explanation of what artificial intelligence means from a legal perspective and how its varying degrees of automation may pose different liabilities. Later, it examines the basic tenets of criminal law and their application to the acts of artificial intelligence. While one may consider cases where artificial intelligence acts merely as a tool and thus liability can be attributed to humans in terms of negligence, vicarious liability, and even strict liability, this paper attempts to deal with cases in which artificial intelligence works as an autonomous agent. It is not only necessary to examine the notion of legal personality for artificial intelligence in itself but to also consider the notion of a responsible agent. In addition, a comparative approach is taken whereby some of the approaches being taken by other jurisdictions, including the EU, UK, and US, are considered as well. The specific problems associated with artificial intelligence in terms of criminal liability within the fields of autonomous driving, healthcare, finance, content creation, and lethal autonomous weapons are discussed here. Finally, the paper will discuss the changes that should be made to the Indian criminal law to accommodate artificial intelligence.

[1] Kirpichnikov, Danila, et al. "Criminal liability of the artificial intelligence." *E3S web of conferences*. Vol. 159. EDP Sciences, 2020.

Understanding Artificial Intelligence in the Legal Context

It is necessary to discuss the meaning of AI in a legal context and why the technical specifics of the technology create particular issues in liability before moving on to the discussion about how criminal law may be applied to AI. From a legal perspective, artificial intelligence is considered to be a system that demonstrates intelligent behavior through the analysis of its environment and execution of specific actions towards achieving certain goals. The technical differentiation of AI systems is important for the analysis. The first one is between weak AI and strong AI. While weak AI, or simulated thinking, is defined as intelligent behavior with no consciousness and subjectivity behind the process, in strong AI, or actual thinking, there should be the same consciousness and subjective behavior as demonstrated by humans. Thus, while a chatbot is capable of engaging in a conversation but cannot understand the nature of this interaction, strong AI would be the ability to do the same but with the same subjectivity as humans. There is no existing system that exhibits consciousness, self-awareness, or intention. This is a crucial differentiation as the legal framework of mens rea necessitates an entity that is conscious and capable of intention, which can be said is something that a weak AI is incapable of.[2]

Another useful categorisation is the differentiation between narrow AI and general AI. Narrow AI, also known as specialised AI, refers to an AI designed for a specific task or set of tasks. Examples of these include IBM's chess AI Deep Blue, spam detectors, face detection, and medical diagnosis tools. In contrast to strong AI, a narrow AI is only able to function effectively within its specialised field of work. On the other hand, AGI, or Artificial General Intelligence, will have the capacity to comprehend, acquire, and use information to do a variety of activities, similar to or even superior to the way humans do them. To date, AGI remains purely theoretical as it has yet to be accomplished. From the perspective of criminal law, however, it is narrow AI that dominates the conversation since it is the one currently doing damage.[3]

In addition to the above categorizations, the degree of autonomy of the AI is likely to be the most important consideration in liability determinations. Autonomy is a term used to describe the extent to which a computerized program is capable of running on its own without any human assistance. The first degree of autonomy involves the use of AI purely as an automated machine that performs tasks as per instructions programmed into it without making independent decisions at all. The next stage involves using AI systems that rely on machine learning algorithms to adapt their behavior according to data input, but they continue to run under the supervision of human operators who can intervene in case something goes wrong. The third and final degree involves AI systems operating autonomously without human intervention in complex and ever-changing conditions. Some examples of fully autonomous AI systems are autonomous cars driving themselves through heavy traffic, high-frequency trading platforms conducting thousands of transactions within seconds, and autonomous lethal weapons selecting and attacking targets without human authorization. However, as autonomy levels increase, the

[2] Sayyed, Hifajatali. "Artificial intelligence and criminal liability in India: exploring legal implications and challenges." *Cogent Social Sciences* 10.1 (2024): 2343195.

[3] Dobrinoiu, Maxim. "The influence of artificial intelligence on criminal liability." *LESIJ-Lex ET Scientia International Journal* 26.1 (2019): 140-147.

connection becomes more remote, and the traditional requirement of voluntariness in the human element, actus reus, may no longer be met.

Finally, the notion of an AI agent is relevant to any discussion on this topic. Computer scientists define an agent as an entity that can perceive its surroundings and interact with it. An AI agent could be a piece of software (bots, algorithms) or embodied (robots and cars). Historically, only two types of entities have been acknowledged as able to enjoy rights and assume obligations under the law – natural person and juristic person (such as corporations and states). AI agents belong to neither category; thus, there is a lack of precedent in the application of criminal liability. The issue is further complicated by the existence of the so-called black box problem. Many advanced forms of AI, including deep learning neural networks, have a level of complexity that leaves even their makers unable to predict or understand how a specific decision was made. Thus, the actual thought processes leading to that decision become completely unknowable and unpredictable, making it impossible to prove causation or to distinguish design flaw from other factors. In criminal law where proof beyond a reasonable doubt is required, this obscurity becomes a formidable obstacle indeed. Therefore, from a legal point of view, it is necessary to realize that artificial intelligence cannot be treated in a monolithic manner. Each AI, depending on the degree of its independence and degree of transparency, requires a certain approach. Knowledge of technical realities is essential for any change in criminal law.

Foundations of Criminal Liability in Indian Law

The structure of criminal liability as provided under the Indian Penal Code, 1860 (replaced by Bharatiya Nyaya Sanhita, 2023) has two primary components which make up criminal liability: actus reus – the guilty act/omission – and mens rea – the guilty mind. In line with a basic principle in the common law system and accepted in the Indian legal system, it should be proven that the accused acted with mens rea in order to commit the crime. It should be understood as a presumption: any statutory offense will be taken to involve a mental element unless there is clear evidence in the statute that an exception should apply.[4] That means the presumption cannot be rebutted unless it is expressly excluded by the law. As a result, when one wants to prove that the accused committed a certain offense according to the Indian criminal code, it is not enough to show that there was a criminal act; additionally, the prosecution must prove that it was done with a certain state of mind – either intention, knowledge, recklessness, or negligence. The theft offense entails the dishonest intention to steal the property, whereas criminal trespass involves the intention to commit an offense or intimidate. Such a sophisticated system demonstrates that the law acknowledges moral culpability based on the mental attitude of the perpetrator.

There is also a provision for strict liability under the Indian Penal Code, but it represents an exception rather than a rule. Some statutory offenses, especially those concerning public good or regulatory issues, can dispense with the mens rea element. It also applies in statutes dealing

[4] Abbas, Thamer Najm Abdullah, et al. "Artificial intelligence and criminal liability: exploring the legal implications of ai-enabled crimes." *Encuentros: Revista de Ciencias Humanas, Teoría Social y Pensamiento Crítico* 22 (2024): 140-159.

with serious social evils. In such instances, the performance of the prohibited act alone constitutes a ground for liability, irrespective of the state of mind of the offender. Nevertheless, courts have been clear that the mere fact that the statute seeks to bring about some welfare or eliminate social evil does not suffice to exempt mens rea. The exemption has to emanate clearly from the language, context, or purpose of the statute.[5]

Another significant concept in Indian criminal law is that of vicarious liability; however, it has a limited scope of applicability. While in case of civil liability, the principal could be liable for the actions performed by the agent, criminal law demands the involvement of the person himself for holding him criminally responsible. But there is one significant exception to this rule in case of common intention, where two or more persons do a thing together with common intention, then every individual will be equally responsible in law as though he were the sole doer of the deed. In order to make the other person responsible in common intention, there should be prior concert or pre-conceived plan, not necessarily formed before the commission of the offence. Under this doctrine, the conduct and intent of the senior officer as the directing mind and will of the corporation is imputed onto the corporation as a whole. Corporate criminal liability applies only to crimes that can be committed by a juristic person, which means that crimes carrying imprisonment as the sole penalty pose certain complications.[6] All of these principles presume a human individual able to form criminal intent, exert control over their behavior, and be morally culpable for their actions. They also presuppose a clear causal relationship between the conduct of the person and the prohibited result. When the actor ceases to be a human and becomes a computer program, every presumption underlying these principles is challenged. There is no mind behind the artificial intelligence that could form mens rea; all decisions made by an artificial intelligence are algorithmically programmed and data-driven. There is no voluntary act on the part of the machine but a probabilistically determined sequence of events programmed into it by humans. Principles of vicarious liability, applicable to human agents acting under human direction, fail to make a proper application to cases involving autonomous machines functioning outside the scope of human intervention. It is possible to partially overcome the issue in terms of corporate criminal liability by considering the AI's conduct the actus reus of the company utilizing the machine, but such an attribution would be far-fetched due to the element of unpredictability brought about by the machine's autonomy even in the eyes of its developers. Consequently, the fundamentals of Indian criminal law must be modified in order to cope with the new reality of AI development.[7]

AI as a Tool: Liability of Human Actors

Where artificial intelligence operates merely as a tool, carrying out programmed operations without exercising true autonomy, the current doctrine on liability could, with some adjustments, be used to hold the relevant human actor liable. When the AI acts not as a free

[5] Seifi, Sohrab, Mahdi Meyhamy, and Hossein Mehrpour. "Strict Liability in Indian Legal System: Historical Considerations, Theoretical Foundations." *Legal Research Quarterly* 27.3 (2024): 1-17.

[6] Nemeth, Charles P. *Criminal law: historical, ethical, and moral foundations*. Routledge, 2022.

[7] Sayyed, Hifajatali. "Artificial intelligence and criminal liability in India: exploring legal implications and challenges." *Cogent Social Sciences* 10.1 (2024): 2343195.

moral agent but rather as a means for implementing certain human actions, then the fundamental issue will become the matter of who performed the action and with what mens rea. There are numerous possibilities as far as human actors go, ranging from programmers writing the AI program, manufactures producing the hardware and installing the software, deployers setting up the AI in a specific location, to end users controlling the operations of the system.[8]

In the case of programmers and developers, criminal liability could be established based on negligence where the design of the AI system has been made in such a way that a reasonable person in that position would have expected potential harm to occur. Standards of care are established depending on the state of the art of the particular technology at the point when the system was developed. Negligence can therefore be charged if the developer failed to ensure the quality of coding or testing, or even foresee certain kinds of misuse and use which might lead to criminal harm. In today's complex machine learning models whose behavior is dictated by training data, it can be challenging to pinpoint a specific malfunction that could be attributed to a particular problem with designing an AI system. It is in this situation that the negligence test suggested in the knowledge base context can be useful. Under the proposed negligence test, damages arising from AI software should be attributed to negligence and not strict liability, coupled with self-regulation using damage impact assessments at all stages of development. A safe harbour regime could also apply, which would limit or insulate liability depending on whether the right steps had been taken. Such a measure would strike a balance between holding individuals accountable for their actions while also promoting innovation by making sure that developers are not strictly liable for the unintended consequences of their inventions.

In the area of manufacturers, manufacturers might also face product liability lawsuits because the AI system is defective by design, production, and failure to warn about defects. In criminal law, such charges will require negligence or strict liability in the case of offences in terms of certain regulations. The knowledge base area, on the other hand, presents an approach for sharing responsibility instead of joint liability. This approach would ensure that the parties liable are those whose actions lead to losses and damage and not one party alone. It would also promote an approach in which any legal proceedings against AI systems would require proof of actual damage, and not just speculation about future harms. This would prevent the situation in which lawsuits were filed based on fears of future damages, which would be an unfair practice.[9]

In addition, those deploying AI systems are also highly accountable. In situations whereby there is a human supervisor overseeing the AI system and with a capacity to intervene or overrule the system's decision, then the human supervisor can face criminal liability for lack of supervision and reckless or negligent control of the AI. For instance, a person using a self-driving vehicle who fails to oversee the road and ends up being involved in a road traffic accident will face a charge of criminal negligence. Equally, a hospital using a medical

[8] Mukherjee, Anirban, and Hannah Chang. "Operational Agency: A Framework for Tracing Intent and Liability in Multi-Agent Artificial Intelligence Systems." *Available at SSRN 5344615* (2025).
[9] Demir, Şamil. "Legal Liability Of Artificial Intelligence (AI) Operators: A Global Analysis." (2025).

diagnostic system whose deployment and testing were not conducted adequately or supervised by a human will face corporate criminal negligence in case there is harm caused by the AI. In relation to knowledge-based context, there is an insightful comparison to airline industry where any accident results into rigorous analysis of the exact failure mechanism involved in order to avoid future occurrence. It is suggested that a culture of this level of investigation in any AI-related harm should prevail.

Vicarious liability may be cited as well to make employers or principals criminally liable for the conduct of their agents in cases where the agent uses AI technologies. But as pointed out above, vicarious liability in criminal law is not well-developed and, generally, needs to prove the fault of the principal itself, like negligent hiring or supervision. Corporate criminal liability, meanwhile, allows to impute to the company, which uses the technology, the conducts and intents of its managing personnel who have knowledge of or were behind the decision to employ the AI system. Such liability may arise in situations where the AI technology is used to carry out any crimes, like algorithmic collusion, automated fraud, and data privacy infringement. The company's "directing mind" may be criminally liable for any harmful acts performed through the use of the technology if there is a possibility to establish that the deployment of the technology was deliberate or reckless enough to inflict damage. Finally, liability of the humans involved in the use of AI, especially as a tool, will depend on the level of control, foreseeability, and duty of care of the participants in the process.[10]

AI as an Autonomous Agent: The Problem of Attribution

The situation becomes particularly complicated when the system acts with a high degree of autonomy, whereby it makes decisions independent of direct human input, learns through data, and changes its conduct based on the new situation. Under such conditions, the autonomous machine acts as an actor of the crime, rather than being simply a tool used by an individual. This is where one faces what has been referred to as the "problem of attribution" – that of identifying a responsible party among humans in relation to crimes committed by an autonomous agent. It is an intricate question involving several issues such as causality, mens rea, and the identification of the perpetrator.[11]

The first and foremost challenge is the challenge of causation. It is a well-established requirement of criminal law that the prosecution must prove beyond reasonable doubt that the defendant's conduct is both the factual and legal cause of the unlawful outcome. In the case where the act that caused damage was committed by an autonomous AI system, the line of causation tends to be lengthy, fragmented, and non-linear. The code was written by a programmer, but the action performed by the computer was not instructed but learned through training datasets. The equipment was made by a company, but the decision of the machine is based on the external variables encountered during its operation. The machine was operated by the user, but he/she may have neither the skill nor knowledge to foresee the decision of the AI.

[10] Arcila, Beatriz Botero. "AI liability in Europe: How does it complement risk regulation and deal with the problem of human oversight?." *Computer Law & Security Review* 54 (2024): 106012.

[11] Franklin, Matija, et al. "Causal framework of artificial autonomous agent responsibility." *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*. 2022.

While the requirement of foreseeability might be fulfilled by several parties simultaneously, the requirement of novus actus interveniens becomes tricky when the AI itself takes a decision which was not foreseeable for any human agent.

The second issue is the difficulty with mens rea. Mens rea, or guilty mind, is an essential element of criminal liability. It means that to be found liable for an offense, one should prove that the accused had intention, knowledge, or even reckless disregard or negligence as regards the occurrence of the prohibited consequence. The autonomous AI program, however, does not have a mind; hence, it lacks intention or deliberate actions. The decisions taken by AI are computational calculations without any mental component; moreover, although one can use such terminology for AI programs, its consciousness can hardly be compared to the human's in terms of mens rea. On the contrary, the producer or the developer might be found guilty of negligence, yet it cannot be proved that he or she intended the exact consequence to occur. With the increase of the level of the machine's autonomy, there will be an increased discrepancy between the anticipated risks and the actual occurrence. In light of this fact, it has been suggested by some experts that a strict liability regime should be implemented in situations involving AI harm, whereby the simple fact that the AI caused the harm makes the user or manufacturer liable without establishing mens rea. But it should be noted that in criminal law, the doctrine of strict liability is rare; hence, for the purpose discussed here, legislation is necessary.[12]

Another problem associated with autonomous AI criminal liability is the issue of the identity of the actor. The principle of individual responsibility lies at the heart of the criminal law—it is necessary to identify a particular individual who committed the actus reus. An autonomous AI that commits an offence cannot be viewed as having done so in a way similar to how a person acts. A machine is not able to bear a criminal responsibility because machines cannot be regarded as legal persons, which means they have neither duties nor rights. Corporate criminal liability allows holding the corporation responsible for the action, although this presupposes proving that there was a "directing mind." In the case of autonomous AI, it can be difficult to prove that this mind was indeed directing a certain decision. This issue can be complicated by another problem known as the "black box" problem. Most of advanced AI technologies, especially neural network models used for deep learning, are black boxes. Even the creator cannot predict or explain all elements of a system's internal operation, such as the specific composition of weights, biases, and hidden layers that resulted in a particular decision.[13]

Three models have been suggested so far for dealing with the problem of attribution. The first model is the strict liability model in which the deployer or manufacturer would be criminally liable for the harms done by the AI system without requiring fault on his part, except for specific defenses. Such a model offers the advantage of ease in operation and guarantees compensation

[12] Carro, Maria Victoria, and David Lagnado. "Human Attribution of Causality to AI Across Agency, Misuse, and Misalignment." *arXiv preprint arXiv:2603.13236* (2026).
[13] Ziemke, Tom. "Understanding social robots: Attribution of intentional agency to artificial and biological bodies." *Artificial Life* 29.3 (2023): 351-366.

to the victim; however, it would stifle innovations and punish persons who had taken all the necessary precautions. The second model is the negligence model, under which the deployer or manufacturer is held criminally liable only when he was negligent in developing, testing, and overseeing the use of the system. Such a model maintains the moral element of the criminal law, but its practicality remains doubtful considering the complexity and opaqueness of the technology. The third model is the corporate criminal liability model under which the corporation deploying the AI will be made responsible if the AI commits an illegal act authorized or recklessly ignored by the corporation's management. The fourth, and most extreme option, is the recognition of legal personhood for AI, whereby the AI itself is a defendant who can be punished, fined, or even "sanctioned" through changes to its programming or deactivation. Such an approach presents numerous challenges both philosophically and practically, such as how to guarantee that punishment will deter a non-conscious AI from repeating its actions, or how to provide sufficient funds to compensate victims of an AI-defendant.

Legal Personality for AI: A Comparative Perspective

The attribution problem, as discussed in the previous section, has made lawyers and policymakers think about adopting an extreme solution, where Artificial Intelligence could be accorded with the legal personality. Legal personality is not an alien concept to the law. Traditionally, the law has known not just natural persons (natural beings), but also juristic persons (Artificial beings), which include corporations, trusteeship, and even rivers in some countries. The common factor between these juristic persons is that while they do not have the power of consciousness or morality, they still get the recognition as a person who can enjoy legal rights and responsibilities. In essence, the question remains – is it possible to accord AI, particularly autonomous AI systems, with legal personality for attributing responsibility and liability? This section discusses how various countries are handling this matter, and what India can learn from them.[14]

The leading example of a legal proposal regarding the legal personality of AI was brought up in the European Union. As early as in 2017, the European Parliament passed a resolution recommending that the European Commission consider the possibility of granting "electronic personhood" to advanced autonomous robots capable of making independent decisions. According to the proposal, such a legal person should be capable of being held accountable for his actions, entering into legal agreements, and paying taxes. This AI should be registered and insured, and a compensation fund must be established to help those who have been harmed by the AI. Nevertheless, the proposal has received criticism from the legal, technological, and ethical community. They argued that the creation of an "electronic person" was simply a facade that would protect the real persons responsible for the actions of AI from any liability. Secondly, they queried whether there was a case for punishment where the object punished cannot feel pain or cannot even be deterred from committing an offense. The proposal by the European Commission was rejected, and later EU attempts at regulation in AI have concentrated on regulating AI using the risk-based approach rather than through legal

[14] Chesterman, Simon. "Artificial intelligence and the limits of legal personality." *International & Comparative Law Quarterly* 69.4 (2020): 819-844.

personality. The European Artificial Intelligence Act of 2024 has made it mandatory for producers and users of AI systems to meet some standards and faces different penalties for breaching these standards without recognizing AI as a legal person.[15]

However, the position in the UK has been more conservative, and the reform efforts have been industry-specific. For instance, there is no legislation establishing a general provision of AI legal personality in the UK. Instead, the government has attempted to regulate liability regimes within certain industries. For example, under the Automated and Electric Vehicles Act 2018, insurers of autonomous vehicles are strictly liable for accidents caused by these vehicles when they operate without human assistance. They can recover any losses incurred from the manufacturer of the vehicle, provided that the accident occurred due to a defect in the machine. The UK Law Commission has looked into the issue of legal personality for AI while undertaking a broader study on smart contracts and autonomous systems. It did not recommend the establishment of a distinct class of legal persons based on AI. In the US, the approach has been fragmented. There is no federal law regarding AI legal personality in the US. The product liability laws, tort laws, and criminal laws still apply to human beings involved in the operation of AI devices. Certain states have passed autonomous car laws imposing liability on manufacturers or operators of these cars, but not one has made provisions for AI having legal personality. In America, the concept of responsibility is still tied strictly to human agency and control.[16]

The matter of legal personality in India with regard to artificial intelligence is only just starting to be discussed. It has to be noted that Indian law recognizes a variety of juristic persons including corporations, societies, trade unions, and even religious idols. Moreover, in the context of corporate criminal liability, it has been recognized that a corporation could be found guilty of certain offenses requiring mens rea if such offenses were committed with the "directing mind and will" of a corporation. Such liability has been found analogous to AI. Thus, if an AI system were to represent the directing mind in some domain, the corporation would be responsible for AI actions in this domain. At the same time, one has to bear in mind that such corporate criminal liability cannot lead to imprisonment of a corporation, nor can it fully take into account AI independence. Whereas the strategy outlined by the Indian government, as previously referred to in this paper, does not provide any insights regarding legal personality, it highlights the importance of having a comprehensive legal system, which would include negligence-based liability, division of damage between parties, and proof of actual harm caused. It states that India ought to adopt relevant international models, adjusted to the Indian legal system.

The comparative study proves that there are no jurisdictions that grant legal personality to AI at the moment. Firstly, such an issue poses a number of difficulties related to registration, representation, and punishment of AI entities. Moreover, from the philosophical perspective,

[15] Filipova, Irina A., and Vadim D. Koroteev. "Future of the artificial intelligence: object of law or legal personality?." *Journal of Digital Technologies and Law* 1.2 (2023).
[16] Novelli, Claudio, Giorgio Bongiovanni, and Giovanni Sartor. "A conceptual framework for legal personality and its application to AI." *Jurisprudence* 13.2 (2022): 194-219.

the definition of legal personhood implies that a particular entity can have certain rights and duties, which means that there must be consciousness, interests, and will. AI entities lack all of the aforementioned traits, since it requires human intervention. Moreover, the concept of corporation can also be regarded as a legal personhood; however, it is nothing more than a fiction, aggregating intentions of humans. On the contrary, the AI may undertake actions that were not planned or instructed by any human, thereby eliminating the applicability of the fiction of personality altogether. As evidenced by the comparison, the best approach would be to improve upon the existing legal framework of liability in favor of humans, via strict liability, mandatory insurance, transparency regulations, and supervision, while at the same time leaving open the option of reassessing the issue in case of development of artificial general intelligence (AGI). For the present time, liability of the human agent as well as regulation of the AI itself appear to be more pertinent than the idea of granting legal personhood to the latter.

Recommendations for Reform

Based on the in-depth analysis conducted in the paper, it becomes clear that the current framework of Indian criminal law lacks the necessary structural foundation needed to deal with the challenges posed by artificial intelligence. Therefore, effective changes should not just be confined to the substance of the law but must take procedural, institutional, and regulatory dimensions into account as well. First and foremost, it is imperative that the lawmaking process should be initiated. The Bharatiya Nyaya Sanhita, 2023, and the Information Technology Act, 2000 do not have specific laws pertaining to harm caused by AI. As such, an entirely new chapter or even a separate piece of legislation should be created to lay down a liability regime that takes into consideration the peculiarities of artificial intelligence. The negligence criterion would be ideal for doing so, considering the knowledge base context in which the issue of harm is analyzed. The negligence standard would oblige the courts to evaluate whether the deploying party (the developer or the manufacturer) has done enough to develop, test, and improve the AI program. Nevertheless, the negligence standard should also be supported by safe harbor provisions, as advised by the circumstances, whereby the person is shielded from liability in case the person took all necessary precautions, including conducting damage impact assessments for each development phase and maintaining thorough audit trails. The balance is struck between responsibility and the drive to foster innovation because the creators cannot be made strictly liable for anything that may arise because of their creation.[17]

The second significant recommendation in addressing AI and criminal culpability is the adoption of a mandatory transparency and documentation framework for AI systems deemed high-risk. The black box issue is the one that stands out as the biggest hindrance to proving causation and culpability.[18] The law requires that all AI systems utilized in sectors such as autonomous transport, healthcare, finance, security, and police operations maintain logs for every input, output, and internal process in the system. In the event that an AI system operates as a black box, then the burden of proving that it adheres to all relevant standards will rest on

[17] Sayyed, Hifajatali. "Artificial intelligence and criminal liability in India: exploring legal implications and challenges." *Cogent Social Sciences* 10.1 (2024): 2343195.

[18] Abbas, Thamer Najm Abdullah, et al. "Artificial intelligence and criminal liability: exploring the legal implications of ai-enabled crimes." *Encuentros: Revista de Ciencias Humanas, Teoría Social y Pensamiento Crítico* 22 (2024): 140-159.

the shoulders of the deploying agency or organization. Another critical proposition that should be considered within the context of the knowledge base context is the need for an actual harm requirement. Criminal proceedings cannot be initiated if it is not clear whether there is any actual harm or only a potential one, with the exception of inchoate offenses. The proposed rule would help make sure that there would be no attempt at punishing a person for creating a risk but not acting upon it. Thirdly, there needs to be proportionate apportionment of damages and liability, rather than joint and several liability. Indeed, in case of a typical damage done by an AI system, the programmer, the manufacturer, the deployer, and even the user might all be to blame to some extent. However, under joint and several liability, the least guilty actor could end up paying for the whole amount of the claim. Proportionate apportionment means that each party pays for what it caused, and this is especially true when the AI was used beyond the scope of permission. However, criminal liability entails a greater burden because it ought only to be meted out on people whose actions were well below the threshold of reasonable care and who caused harm substantially through their own fault. As an analogy from the context of knowledge-based systems is quite illustrative in this instance, every single air crash undergoes a very detailed investigation aimed at discovering which element failed in order to ensure that it does not happen again in the future and not necessarily to assign any culpability. Therefore, criminal liability ought only to be applicable when clear acts of negligence and gross negligence can be established, while fines and other forms of liabilities will be used in dealing with ordinary negligence cases.[19]

The final reform would entail having a specialized tribunal or court for handling AI disputes. Due to the nature of AI systems and the evidence that would be presented in such disputes, a normal criminal court might lack the expertise to handle them effectively. As in the case of the tribunal to be set up for resolving anti-defection law cases, a special tribunal to resolve issues pertaining to AI-based crimes must be formed. The tribunal will have the mandate to settle cases related to AI systems within the country. Judges on the tribunal will be specialists in AI technology, with the ability to make judgments regarding causation, standards of care, and liability distribution. Such judgments include ordering the alteration of the source code of AI systems, suspending or decommissioning certain systems altogether if there is no other means of preventing harm from occurring in the future. [20]

Lastly, the law ought to include a requirement for the establishment of a mandatory insurance and compensation regime for risky AI. Even with the toughest of liability regimes, there will always be instances where some form of loss occurs due to the actions of an AI and where no person can possibly be criminally liable. Such situations can arise either due to the inability to satisfy the criteria necessary to prove fault or because the person responsible cannot be traced. In such situations, the victim of the wrongdoing must not be left to fend for himself or herself, especially since there would be someone capable of paying for the damages, even if he or she may not have committed a crime. Therefore, it is important to establish a compensation fund

[19] Osman, Yasein Hassan M., et al. "Criminal Responsibility in the Age of AI: Rethinking Legal Accountability." *African Journal of Law and Justice System* 4.3 (2025): 179.
[20] Raza, Ahmed, et al. "Artificial Intelligence and Criminal Liability; Rethinking Criminal Liability in the era of Automated Decision Making." *Rethinking Criminal Liability in the era of Automated Decision Making (July 31, 2023)* (2023).

for such victims based on a strict liability criterion, whereby all that is required is proof that the AI was responsible for causing the damage and not necessarily any proof of culpable behavior. However, since artificial intelligence is a rapidly developing subject, a rigid approach to formulating laws in relation to it cannot be followed since it will make such laws redundant very soon. It is recommended that a permanent committee made up of representatives from the Ministry of Law & Justice, Ministry of Electronics & IT, Law Commission, and technology experts monitor developments in this regard and suggest amendments to the law accordingly. The collective efforts of these proposed amendments would result in a logical and futuristic approach to criminal liability in the era of artificial intelligence in the country.

Conclusion

The convergence of artificial intelligence with criminal responsibility constitutes perhaps one of the most formidable challenges to modern legal theory. As the foregoing discussion has made clear, the fundamental tenets of Indian criminal law, namely actus reus, mens rea, causation, and personal accountability, are inherently incapable of resolving the problem of harm caused by autonomous machines without the element of human intent, human consciousness, and human predictability. The inquiry has ranged from a situation where artificial intelligence is simply treated as a means, where established legal theories of negligence, vicarious responsibility, and corporate criminality can, with some modification, be used to hold humans accountable, to a scenario where artificial intelligence is regarded as an autonomous agent, where the attribution issue becomes particularly vexing. The study of AI legal personhood in various legal systems shows that no legal regime recognizes AI as a legal person, and there appears to be little prospect of doing so in the near term. The concrete problems arising in each of these areas—autonomous vehicles, AI in healthcare, algorithmic trading, AI-created content, and lethal autonomous weapons—demonstrate that the theoretical difficulties are not only relevant in an abstract manner but also have real-world implications calling for legislative intervention. Together, the suggested reforms outlined in this paper, namely, the implementation of negligence liability with safe harbor clauses, mandating transparency and audibility, apportioning liability on a proportional basis, imposing a requirement of actual harm, adapting corporate criminal liability, establishing an adjudicatory tribunal, mandatory compensation mechanisms, and periodical reviews constitute a roadmap that will allow striking the delicate balance between the necessity of holding responsible parties liable while fostering technological innovation. India, being a nation that is witnessing rapid digitalization and has set up a strategy for the implementation of AI technology in key industries, needs to take a proactive stance regarding this issue. The law should not lag behind until a disastrous incident occurs, which then shows the inefficiency of current regulatory schemes. In fact, it is imperative to reform the existing laws with due foresight and principles before it is too late.

References

Statutes and Constitutional Provisions

The Constitution of India, 1950.

Bharatiya Nyaya Sanhita, 2023 (No. 45 of 2023).

Indian Penal Code, 1860 (No. 45 of 1860) [repealed].

Information Technology Act, 2000 (No. 21 of 2000).

Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.

The Consumer Protection Act, 2019 (No. 35 of 2019).

The Motor Vehicles Act, 1988 (No. 59 of 1988).

European Union, 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).

United Kingdom, Automated and Electric Vehicles Act 2018 (c. 18).

United Kingdom, House of Lords Select Committee on Artificial Intelligence, *AI in the UK: Ready, Willing and Able?* (HL Paper 100, 2018).

Books

S.N. Dhyani, *Fundamentals of Jurisprudence: The Indian Approach* (Central Law Agency, 2019).

Justice M. Rama Jois, *Legal and Constitutional History of India: Ancient, Judicial and Constitutional System* (LexisNexis, 1st edn., 2022).

Seema Singh & Raman Mittal, *Law and Spirituality* (Kitabwale, New Delhi, 2024).

Ronald Dworkin, *Taking Rights Seriously* (Harvard University Press, 1977).

H.L.A. Hart, *The Concept of Law* (2nd edn., Clarendon Press, 1994).

John Austin, *The Province of Jurisprudence Determined* (Cambridge University Press, 1995).

M.P. Singh, *V.N. Shukla's Constitution of India* (13th edn., Eastern Book Company, 2017).

Durga Das Basu, *Commentary on the Constitution of India* (8th edn., LexisNexis, 2011).

K.D. Gaur, *Textbook on Indian Penal Code* (6th edn., Universal Law Publishing, 2020).

Journal Articles

S. Ganesh, "Vedic Concept of Dharma" (2021) 12 *Purvamimamsa* 45.

Allen Buchanan, "What is so Special about Rights?" (1984) 2 *Social Policy & Philosophy* 61.

M.P. Singh, "The Anti-Defection Law: A Critique" (1993) 35(2) *Journal of the Indian Law Institute* 211.

Vikram Raghavan, "The Tenth Schedule: The Anti-Defection Law in India" (2002) 44(1) *Journal of the Indian Law Institute* 1.

A.N. Veera Raghavan, "Legal Profession and the Advocates Act, 1961" (1972) 14 *Journal of the Indian Law Institute* 228.

Lon L. Fuller, "The Case of the Speluncean Explorers" (1949) 62 *Harvard Law Review* 616.

Lawrence B. Solum, "Legal Personhood for Artificial Intelligences" (1992) 70 *North Carolina Law Review* 1231.

Matthew U. Scherer, "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies" (2016) 29(2) *Harvard Journal of Law & Technology* 353.

Ryan Abbott, "The Reasonable Computer: The Impact of Artificial Intelligence on Negligence Law" (2018) 68 *University of Toronto Law Journal* 130.