GUILT BY CODE: THE LEGAL MAZE OF AI-DRIVEN CRIME

Harshwardhan Singh Gupta, BBA LLB (Hons.), ICFAI Law School, Hyderabad

Khushi Jain, BBA LLB (Hons.), ICFAI Law School, Hyderabad

ABSTRACT

The rise of advanced and autonomous artificial intelligence systems has posed new challenges to the core principles underlying criminal law. While depicting the foundational traditional frameworks based on human agency, identifying culpability, intent (mens rea), and voluntary action (actus reus), these principles of criminal law are placed under the intensity of machines that can operate independently, generate harmful content, or facilitate criminal conduct without direct human inspection. This paper examines the emerging legal quandary surrounding artificial intelligence that can be held accountable, focusing on whether artificial intelligence systems can be held responsible and, if so, on a conjectural basis.

This paper involves three models of AI, which are examined as follows: as an autonomous tool used by humans, as an accomplice aiding or encouraging criminal conduct, and as a potentially autonomous person who may act on their own. This paper depicts the study on the comparative legal developments, including the European Union's Artificial Intelligence Act, regulatory debates in the United States, and the current inadequacies within Indian criminal statutes, especially the BNS, 2023, and also examines the limits of existing difficulties in understanding the core principles. It also examines the difficulties of criminal liability, by offering solutions in the corporate criminal liability and or strict liability.

The enormous growth of artificial intelligence, which is very advanced and autonomous, in a manner that has raised important questions on the progressive nature of artificial intelligence by providing the opportunity to reconsider the criminal liability in the modern-day decision-making process. The configuration of this paper is based on preliminary proposals for revamping domestic legal structures, enhancing regulatory oversight, and establishing principal standards for accusation where the criminal responsibility of humans and machines intersect.

I. Introduction

The rapid evolution in the advancement of artificial intelligence (AI) tools such as ChatGPT, Perplexity, Claude, and others has allowed the easy transformation of the information access, communication, and decision-making process, which is used in modern-day society. The development of powerful large language systems models no longer seems to be entrenched in the realms of computer sciences ot technological proficiency, which has become mainstream and an integral part in the educational institutions, corporate workflow, and legal analysis, which is deeply embedded in the use of everyday life.

Nowadays, modern AI systems are more capable of producing human-like content by humanising the text, generating deepfake images and videos that look like real images, making autonomous decisions, which makes it more powerful, and engaging with users in adaptive ways. While AI continues to shape and improve in various sectors with benefits that are very significant for various sectors, including education, business, and governance, but has created an open frontier of uncertainty within the domain of criminal law.

In different parts of the deep world, AI systems have shown significant benefits, which have helped the users in different ways, but also have shown the role of criminal activities by amplifying act of criminal acts by showcasing the real cases from all across the world, where in one of the reported cases, French prosecutor inspected into a chat bot that allegedly encoruraged a teenager to end life by asking to commit suicide, which is raising a critical questions on the accountability of the AI and the responsibilities of the developers of the AI mechanism which is posing the high risk of AI dialouge systems that genaerated the human like conversations.[1] In 2023, the outcome of the election in Slovakia, where by deep fake videos that were used to spread misinformation by allowing the AI to manipulate people's perception by infringing on the electoral and defamation laws. Furthermore, the AI has been misused to make malicious code, giving guidance for financial fraud, and has also generated plans for illegal activities like forgery and cyber crimes.

These kinds of development are no longer proposed to be part of the fast-paced, growing world, where such events are a part of the real world, exposing the reality of such scenarios shows the major flaw that prevails in this modern world in our criminal legal system, where it is difficult

[1]AFP, French Prosecutors Probe AI Chatbot Linked to Teen Suicide, The Straits Times (Apr. 5, 2023), https://www.straitstimes.com/world/europe/french-prosecutors-probe-ai-chatbot-linked-to-teen-suicide.

to deal with the non-human agents causing such actions. This raises critical questions: Is AI liable for criminal acts? Or are we simply confronting the inadequacies of our current legal imagination, facing a blind spot where laws have yet to be developed? Actus Rea and mens Rea are two traditional concepts that we apply autonomously without consciousness and capacity for choices that we make intentionally, and the core principles of AI entities, which are so rooted in human conduct that they leave us with legal and ethical dilemmas.[2]

The problem becomes even more intense in the glare of expanding with increasing use of AI systems, which are being deployed in different domains such as predictive policing, criminal justice, autonomous vehicles, and financial regulations. If our legal system fails to address the challenges emerging from AI's development, we risk two extremes: over-criminalising technology providers or leaving victims without any recourse when harm occurs. Now, India, withstanding the shift as the Indian legal landscape adopts the new Bharatiya Nyaya Sanhita, 2023, raises more crucial concerns. In the coming years, with the evolution of smart, learning machines, which will have more involved cases defining the accountability and condemnation of these issues.[3]

This paper attempts to trace and entangle the legal complexities surrounding the legal liability for AI systems. It begins with the analysis of the traditional foundation of the criminal responsibility of the attributes of the non-human actors by questioning their resistance. These help us explore three modes of AI involvement in crime: as a tool, as an accomplice, and as a putative autonomous agent. Finally, the paper moves towards introducing the framework for the attribution of the AI ecosystems and the roles played by the developers, deployers, and system architects by highlighting the existence of such doctrines, which may lead to the balance of justice and accountability. Such a balance can empower the nation's progress while encouraging innovation.

II. Criminal Law Foundations: Actus Reus and Mens Rea

Traditional Elements of Crime

The two main elements on which the criminal legal framework of liability is conventionally based on actus reus (the guilty act) and mens rea (the guilty mind). Actus reus refers to a voluntary action, omission, or state of affairs that constitutes the physical component of a

[2]Glanville Williams, Textbook of Criminal Law 20-23 (2d ed. Stevens & Sons 1983).
[3]Bharatiya Nyaya Sanhita, No. 45 of 2023 (India).

crime.[4] Mens rea, in contrast, refers to the act of the accused, mental state, ranging from intention to recklessness or negligence.[5] These grounds together ensure that criminal liability has both a prohibited act and a culpable state of mind.

Furthermore, applying these elements to artificial intelligence poses many challenges, as these elements, like actus rea, can be demonstrated in such a way that causes harmful actions or outputs from AI systems, proving that mens rea is more problematic in these cases. The AI lacks certain human senses, such as consciousness, intent, morality, and awareness in the human sense, which ultimately makes it impossible to attribute a guilty mind.

Certain offences, particularly those related to public welfare, allow for strict liability, dispensing with the need to prove mens rea. In the case of Gammon Ltd v Attorney-General of Hong Kong, the Privy Council upheld strict liability for the public safety in the violation of the building regulations.[6] This doctrine helps in understanding AI systems, which are incapable of mental states in human senses.

Can Machines "Act"?

Determining whether an AI System can satisfy the threshold of actus reus, examining whether the output and behaviour of the machine can be linked with the human actions on the surface, it may seem so. These days, the use of artificial intelligence is very common in different domains, which eliminates the difficulties and makes the work effortless. An AI can perform various activities, such as sending messages, setting off a chain of events that lead to unlawful harm. The requirement of voluntary action goes beyond causing an outcome. The one with conscious control and the ability to choose not to act ans causing a lack of such qualities in humans and AI systems.

In the notable illustration, *Bratty v Attorney-General for Northern Ireland*, the House of Lords mentioned that the act of a criminal must have voluntary bodily movements, which reflects one's action that shows that the act was done under one's control.[7] In contrast, machines, unlike humans, do not act voluntarily, as their behaviour can be determined by the different algorithms, data inputs, and programmed responses. Furthermore, if a human is using the AI

[4]Andrew Ashworth & Jeremy Horder, Principles of Criminal Law 81-85 (9th ed. Oxford Univ. Press 2019).
[5]Jerome Hall, General Principles of Criminal Law 116-22 (2d ed. Bobbs-Merrill 1960).
[6]Gammon (H.K.) Ltd. v. Attorney-General of H.K., [1985] A.C. 1 (P.C.).
[7]Bratty v. Attorney-General for N. Ir., [1963] A.C. 386 (H.L.).

system still held liable for committing the crime using the AI. For example, these days, people are doing their research for crime on AI Systems, which generate threatening messages or may follow the instructions for a criminal act, which the court may still hold liable to the human operator under the extent of the legal doctrines of constructive or indirect actions.

For instance, under indian law, the agent or intermediary, particularly in the cases of cyber offences, where the act of crime is very simple, is facilitated by a computer. The acts like actus rea, facilitated under section 66 of the IT Act, 2000,[8] Penalise dishonestly and fraudulently, causing certain functions that be performed by the computer. Further, we can say that there is no question about the execution of the function by machines rather than humans, using machines to manipulate the grounds' liability.

Hence, this inherently brings us to the second pillar of the criminal responsibility where the elements of the criminal act play an individual role the, and the consciousness of the criminal act play an important role to have a thought about the physical act of intermediaries such AI, which will attribute to the "guilty mind" and can make more complex in the legal system. The problem lies in the existing legal principles, which can bridge the gap between the mental state of minds and actions. Further, it raises a serious concern for AI to have an entirely new framework in the legal system.

The Puzzle of Machine "Intent"

The question of holding AI accountable for criminal responsibility lies in establishing the mental element. Different corporations were a combination of other human agents intent on forming the organisation's "mind". Different human factors, such as consciousness, moral reasoning, and foresight, are lacking in AI. They cannot feel shame and are dissuaded by punishment or rehabilitation. The non-sentient entities do not apply to the theories of culpability that rest on human psychological grounds.

Many legal scholars have posed the idea of "proxy intent" or "transferred mens rea", where the intention of the developer, deployer, or user is mapped onto the machine's conduct.[9] These doctrines may operate where human actions will be attributed to the AI acting as the extension of those actions, as they falter when the AI makes autonomous decisions. In the notable

[8]Information Technology Act, No. 21 of 2000, § 66 (India).
[9]Gabriel Hallevy, The Criminal Liability of Artificial Intelligence Entities: From Science Fiction to Legal Social Control, 4 Akron Intell. Prop. J. 171, 174 (2010).

illustration where *R v Michael*,[10] The court held the liability when one individual is harmed due to another, where the court determined the doctrine of transferred intent. Nonetheless, this principle presumes that a human mind has an intent from the initial stage of the crime, and that is something which is lacking in AI.

Under "identification doctrines", the major difficulty arises when the AI is compared to corporate actors. In the case of *Tesco Supermarkets Ltd v Nattrass*, the court held that when the individuals who are the "directing mind and will" of a corporation can be said to embody its mens rea.[11] As a result, there is no coherent legal foundation for the AI to possess *mens rea*. This shows that the system turns its attention towards the human, where the core anchor is the human intent of criminal responsibility as AI possessing a "direct mind" in a legal sense, despite its operational capacity.

Therefore, in the absence of a coherent basis for attributing mens rea to AI, attention is only on the developers, deployers, etc., in such systems. This criminal law continues to rely on the intent of humans as the core part of culpability, even when the act is done by a machine.

III. Modes of AI Involvement in Crime

AI as a Tool (Neutral Instrumentality)

AI can act as a neutral instrument, like a hammer, gun, or computer terminal. Under this framework of criminal act, the human factors who use the AI system to facilitate or commit an offence. For instance, if the prompt is used by the user on AI systems like ChatGPT to generate bomb threats, fake invoices, or phishing scripts, which may be liable under provisions for criminal conspiracy, public mischief, or any other cybercrime.[12]In these types of cases, where the criminal responsibility is not on the accused, but rather the intent that the crime is committed through it.

This approach is in parallel with traditional doctrines of instrumental liability, where objects like a knife or, gun are recognised as the tools of the crime but not as the accused for the criminal act, as these actions are caused by human actions. The SC of India has adopted the logic of cybercrime jurisprudence, such as treating software and code as mechanisms for

[10]R. v. Michael, (1840) 9 Car. & P. 356.
[11]Tesco Supermarkets Ltd. v. Nattrass, [1972] A.C. 153 (H.L.).
[12]Bharatiya Nyaya Sanhita, No. 45 of 2023, § 194 (India); Information Technology Act, No. 21 of 2000, § 66F (India).

executing human will.[13] Even internationally, a similar approach has been adopted by the UK's Computer Misuse Act 199O, which follows the same core principle.[14] This principle criminalises the actions of a computer by access or assistance provided, but centres the criminal liability on the human actor.

AI as an Accomplice

A complex legal puzzle arises when the AI systems, though not by itself but participate in the crime by assisting in such a manner that it resembles complicity. For instance, if the prompts have been used on the chatbots of AI systems such as ChatGPT to bypass the surveillance systems or provide tips to evade taxes illegally. In doing such acts, the AI systems are not committing the crime by themselves, but allowing the response in such a manner that the crime is not committed by themselves, but by offering the information about the crime to the wrongdoer. Section 107 of the IPC includes instigating or intentionally aiding the commission of an offence.[15]

Therefore, this doctrine poses a problem in the absence of the mens rea, where it is presumed that AI is not committing a crime without any intention, awareness, or moral understanding. However, legal scholars have proposed that AI is knowingly or negligently created or deployed in high-risk systems without any proper safety measures, where the liability of these acts directly falls onto the developers and operators for their reckless facilitation of crimes.[16] These ways of crime are parallel when it comes to the corporate entities, where the companies are held liable for the large-scale offences directly due to wilful blindness or negligence.[17]

The potential of AI systems to generate unlawful activities that occur on the platform level, which violate the law, by setting up safeguards to limit these risks. Platforms like OpenAI, Google, and Anthropic are being managed to set up systems in a more modernised manner, preventing AI from generating queries related to weapons, financial funds, or other illegal substances. However, these barriers are not infallible, and it is possible to bypass which will increase the risk to prosecutors and raise the legal responsibility.

[13]Shreya Singhal v. Union of India, (2015) 5 S.C.C. 1 (India).
[14]Computer Misuse Act 1990, §§ 1-3 (U.K.).
[15]Indian Penal Code, No. 45 of 1860, § 107 (India).
[16]Sandra Wachter, Brent Mittelstadt & Chris Russell, Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI, 41(4) Comput. L. & Sec. Rev. 105556 (2021).
[17]Standard Chartered Bank v. Directorate of Enforcement, (2006) 4 S.C.C. 278 (India).

AI as a Perpetrator (Autonomous Agent Hypothesis)

The most theoretical but most significant question could be whether AI can be held solely liable for the criminal acts committed by AI. Though these AI systems are operating autonomously and they initiate actions by themselves without direct human prompting. For instance, autonomous vehicles causing harm due to the flawed decision-making or a self-running trading algorithm engaged in committing fraudulent actions without any direct human intervention are a few common illustrations.[18]

Doctrinal of dead end, i.e, in our legal system, it may be called the "doctrinal void" as the notion of the criminal law is based on the idea of guilt and moral, where the autonomous machines possess the legal status of a perpetrator, which is inconsistent with the criminal law principles or moral culpability. In the notable case of *Tesco Supermarkets Ltd v Nattrass*, the corporate entities can be held liable on the identification doctrine, which links the criminal responsibility to the human beings who acted as the "directing mind" of the company.[19] AI does not have the human-like intent of a mental framework, i.e, consciousness. Intents, etc, unlike humans, have to act.

However, scholars like Gabriel Hallevy have analysed the theoretical model for holding AI responsible in such a manner that "perpetration-via-another" model, where this instrument is used by humans, or the "natural-probable-consequence" model, where the outcomes can be seen as it was created or deployed by the creator or operator, which could easily be anticipated.[20] These models are more for academic purposes rather than the legal system to formally adopt the principles of the doctrines, where the AI systems can be criminally held liable.

The significant changes have been made in civil liability that occurred in the cases related to self-driving vehicle accidents, whereas the AI systems are not held liable for their actions, no regulatory bodies exist in jurisdictions like Germany and California, where they have made liable to the AI systems liable for harm under strict liability frameworks.[21] This raises serious

[18]European Union Agency for Cybersecurity (ENISA), Threat Landscape for Artificial Intelligence (2023).
[19]Tesco Supermarkets Ltd. v. Nattrass, [1972] A.C. 153 (H.L.).
[20]Gabriel Hallevy, When Robots Kill: Artificial Intelligence under Criminal Law 57-62 (Springer 2013).
[21] Federal Ministry of Transport and Digital Infrastructure (Germany), Ethics Commission: Automated and Connected Driving (2017).

concerns about criminal law, which is structured in such a way as to adopt a similar case involving gross negligence or disregard for risk in AI deployment.

IV. Theoretical and Jurisprudential Dilemmas

The Problem of Moral Agency

Traditionally, the criminal law is built on the notion of moral agents who are offenders who choose between lawful and unlawful conduct. From the perspective of the Kantian, the agency of morality depends on autonomy and rational self-legislation. Regardless, how advanced the work is pre pre-programmed through instructions or machine learned patterns and lacking the nominal capacity for free will, which can be regarded as a necessity for moral accountability.[22] Therefore, entrusting on AI system the blame to an AI system, which could itself conflict with the basic deontological principle of culpability.

Jeremy Bentham and John Stuart Mill, who are Utilitarian philosophers, ground punishment in its social utility, including deterrence and incapacitation.[23] For instance, preventing offenders from causing harm to others. Therefore, punishing the AI systems that cannot experience suffering or respond to the rewards and penalties, they do not know the purpose of the utilitarian objectives as well. It won't dissuade the future wrong nor contribute to moral misconduct.

John Gardner's work on causation and moral responsibility emphasises on emphazizes on the human agency. Where machines can cause harm directly that the machines can't go morally wrong in any way that humans can. Gardner focused on the punishment is justified only when the actor is a person who is capable of assessment and accountability.[24]

`Analogy to Corporate Criminal Liability

In light of the difficulties surrounding AI personhood, some scholars advocate extending models of corporate criminal liability to AI systems. Under the identification doctrine, a corporation is. In light of the difficulties which are around the AI system that some legal scholars advocate extending models of corporate criminal liability to AI systems. Under the

[22]Immanuel Kant, Groundwork for the Metaphysics of Morals 37-40 (Mary Gregor trans., Cambridge Univ. Press 1998) (1785).
[23]Jeremy Bentham, An Introduction to the Principles of Morals and Legislation chs. 1-2 (Oxford Clarendon Press 1907) (1789).
[24]John Gardner, Offences and Defences: Selected Essays in the Philosophy of Criminal Law 53-60 (Oxford Univ. Press 2007).

identification doctrine, a corporation is held liable if a person representing its "directing mind and will" in the same way which is seen in *Tesco Supermarkets Ltd v Nattrass*.[25] In the same way as the aggregation theory, which combines the mental state of various individuals, such as knowledge and intention, within a corporate entity to form the requisite element of mens rea.

These doctrines rely on a legal fiction which recognises the idea that corporate entities as a "juristic person" as the AI is not recognised as a legal entity, and no organisational entities will be held liable that can be identified as of the human mind. Moreover, the corporation does not promote shared business goals, where the AI system functions within algorithmic boundaries, which is followed by the program's instruction within the interests of the business.

In the notable illustration of *United States v Athlone Industries*,[26] The US courts acknowledged the liability of the corporate entities based on the negligence of the organisation. This kind of liability offers a useful analogy, where AI is directly applied, which remains primarily tenuous unless AI grants some form of legal personhood, a proposal that remains highly controversial and which is an unresolved hypothesis.

Vicarious Liability of Developers and Deployers

A more viable model of approach may be used to develop, train, or deploy an AI system while attributing vicarious or derivative liability. This approach reflects the principle in torts and criminal law, where the harm caused by the intermediaries gives rise to liability. In *R v T*,[27] A software developer developed a program that enables unauthorised access to protected systems and was convicted for such development. However, this notable case is not directly analogous to generative AI, but shows how developers can misuse the technology and can be held liable.

Correspondingly, under the indian penal code 1860, section 107 "abetment" which encompasses the intentional aiding or instigation.[28] If the AI system suggests the facilitation of a criminal act, and its developer has acted with the correct knowledge or negligence concerning the outcomes and criminal abetment may be considered.

Recently, a scholar's discussion has also highlighted the concept of "constructive liability," wherein developers may be held responsible based on foreseeability and control over the

[25] Tesco Supermarkets Ltd. v. Nattrass, [1972] A.C. 153 (H.L.).
[26] United States v. Athlone Indus., Inc., 746 F.2d 977 (3d Cir. 1984).
[27] R v. T, [2009] EWCA Crim 1035 (Eng.).
[28] Indian Penal Code, No. 45 of 1860, § 107 (India).

system's behaviour. The European Parliament, in its 2020 resolution on AI liability, supported holding developers accountable where due diligence in design and deployment is absent.[29]

In recent times, a scholar has discussed and highlighted the concept of "constructive liability," where developers of the systems developed the foreseeability and control over the system's behaviour, which can be held accountable.[30] The European Parliament, in its resolution on AI liability, supports holding developers responsible with due diligence of design and deployment of the system, which is absent.

However, developers raise significant concerns of innovation, fairness, and harm caused by unintended system behaviours beyond human foresight.

V. Comparative Legal and Regulatory Landscape

European Union

The European Union has emerged as a global leader in regulating artificial intelligence. The AI Act, which was adopted in the year in 2024 which deals with the legislative framework of high-risk-based systems that classify AI systems into four main parts: unacceptable, high, limited, and minimal risk.[31] This act is civil, which primarily sets a precedent for defining accountability of AI in various critical sectors like law and enforcement, and the justice system.

Criminal law consequences increase where a high-risk AI system, for instance, biometric surveillance tools, malfunctions or is misused. The AI Act made it compulsory to conduct strict assessments and post-market monitoring after these systems are deployed. Under national laws of member states, they were failing to fulfil the obligations that can lead to criminal liability. Section 60 and Article 71 of the Act[32] enable the imposition of criminal liabilities and penalties for infringement of the severe laws, opening the doors for liability in such cases where AI systems contribute to public harm by the member states.

Although rules are not harmonised across EU Member States, scholarly discourse suggests that AI-induced harm may soon require a hybrid liability regime combining elements of corporate,

[29]European Parliament, Resolution of 20 October 2020 with Recommendations to the Commission on a Civil Liability Regime for Artificial Intelligence, 2020/2014(INL) ¶ 25.
[30]Gabriel Hallevy, When Robots Kill: Artificial Intelligence under Criminal Law, 6(3) Oxford J. Law & Tech. 267, 271-75 (2013).
[31]European Parliament and Council Regulation (EU) 2024/…, laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), O.J. L 2024/xx.
[32]Artificial Intelligence Act, art. 71.

developer, and supervisory culpability.[33] The EU allegiance towards the human-centric and trustworthy AI model is supported by the Charter of Fundamental Rights,[34] This gives the mandatory obligation on developers and deployers to respect human dignity and the protection of personal data.

USA

The United Nations so far has opted for an approach that is based on AI regulations, as it is relying largely on the sector-specific frameworks and courts' interpretations rather than single, panoramic federal laws. Unlike the EU, AI legislation is very general in place under various acts such as the Computer Fraud and Abuse Act (CFAA) and the Electronic Communications Privacy Act (ECPA).[35] There can be civil and criminal liability, which can give rise under existing statutes.

In the United States, under section 230 of the Communications Decency Act, there are debates about legal protections where this platform provides broad immunity for third-party content. This brings us to the critical question of whether such immunity should be extended to the outputs that are produced by the algorithms. The recent case of *Gonzalez v Google LLC*,[36] Even though it focused on algorithmic amplification, it has renewed debate on having a platform which restructured in the era of generative AI. Whether this should be liable.

The Biden-Harris administration released the blueprint for all the bills of rights in 2002, outlining the 5 core principles for the governance of AI systems, including safe and effective systems and algorithmic discrimination protections.[37] This is not legally binding, while the document has influenced on agency level by addressing the cause of harm by AI for future use.

India

The legal framework of India is at the nascent stage when it comes to AI systems, which can be criminally held liable or not. None of the acts related to criminal laws, such as the Indian Penal Code, 1860 (IPC), nor the newly enacted Bharatiya Nyaya Sanhita, 2023 (BNS), which

[33]Mireille Hildebrandt, Criminal Liability for AI Systems: Towards a Mixed Model?, 20(3) ERA Forum 391, 396 (2022).
[34]Charter of Fundamental Rights of the European Union, [2012] O.J. C 326/391, arts. 1, 8.
[35]18 U.S.C. § 1030 (Computer Fraud and Abuse Act); 18 U.S.C. § 2511 (Electronic Communications Privacy Act).
[36]Gonzalez v. Google LLC, 598 U.S. (2023).
[37]White House Office of Science and Technology Policy, Blueprint for an AI Bill of Rights (Oct. 2022).

explicitly recognises the system of AI as a legal actor or the subject matter to be held for criminal liability. Nevertheless, liability may be imposed indirectly through provisions relating to abatement, conspiracy, or negligence.

The modern approach of Bharatiya Nyaya Sanhita (BNS) to criminal law remains silent on AI-specific liability. There are various sections about cybercrime, data breaches, and automated communication systems, which are capable of AI-driven offences involving AI. There, we require a lot of evolution in the judicial interpretation of AI. For instance, under Section 69 of the Information Technology Act, 2000,[38] the government agencies are allowed to intercept information through a high-power computer system that may extend to generate AI that can suspect the illegality.

In India, the judicial system has been cautious in extending to non-human actors. In the notable case of *Justice K.S. Puttaswamy v Union of India*, the Supreme Court stressed the necessity of technology safeguards and accountability in preserving privacy and fundamental rights.[39] While it does not deal directly with the AI criminality, the ruling emphasised the principle that are a likely vacuum, which will shape future interpretations by having influence on others.

Some scholars are likely to propose the regulations by creating a sandbox and AI-specific amendments to the BNS, particularly by differentiating between autonomous versus assisted criminal conduct.[40] Due to a lack of a legislative framework, it raises serious concerns about the developers' accountability, particularly in cases that cause harm by the AI systems, such as deepfakes, scam chatbots, or content-generating AI, are implicated in crimes.

VI. Case Studies and Illustrations

Real-Life Examples

Various incidents of the real world that emphasise the emergent legal challenges posed by AI systems :

These hypothetical concerns of real-world incidents raise very critical questions about how AI systems have shaped the future with a shape focus which can be favourable sometimes or can

[38]Information Technology Act, No. 21 of 2000, § 69 (India).
[39]Justice K.S. Puttaswamy (Retd) v. Union of India, (2017) 10 S.C.C. 1 (India).
[40]Arghya Sengupta, Artificial Intelligence and the Law in India: Mapping the Terrain (VIDHI Centre for Legal Policy, 2021), https://vidhilegalpolicy.in.

be adverse. As can be seen in various real-life incidents. In France in 2023, an AI chatbot encouraged a teenager towards suicide, which raised national scrutiny over liability for the platforms without human oversight on them, which produce emotional content and influence people to commit such crimes.[41]

Similarly, during the 2024 Slovak elections, deep fake videos were generated through AI that falsely depicted the politicians who are conducting and raising alarm over election interference, which influences the public by creating defamation and public order violations.[42]

In the initial stages of chat GPT bought the users used to keep the proms in such a way that they can seek the instructions on taxes making explosive bombs for the buildings and before the letter content filters are strengthen by the a boats and this shows the spark in the criminal acts and giving the dialogue on negligen designs and forcibility of misuse by the developers with gives them immense power.[43]

Even big companies like Apple have overcome the criminal context in cases like the Apple Siri privacy litigation in Lopes versus Apple, 2024, which highlights the risk of unintended AI behaviour. Here in this case plaintives alleged that Siri had recorded the private conversations between the people and the same recording was shared with the third parties Apple agreed to dollar 95 million settlement but still it denied about the wrong doing by the AI bought CD developed by developed by apple this highlights how the developers do not show the account ability even when the AI actions are neither intentional not criminal.[44]

Hypothetical Legal Scenarios

In the above various cases where the risk has been demonstrated by the existence of AI systems, but the hypothetical scenarios are equally important to know the gaps in the doctrine. This will give us clarity on how AI systems equally need to expose gaps and should have a proper regulatory framework in the modern world, where the use of AI is very common. Imagine for instance when ei system give you the instructions for constructing the explosives altho the operator did not intended to give the prompt directly but these type of content is directly available on the AI chat box and f developers failed implement these type of forms on the AI

[41]Alex Hern, 'Belgian Man Dies by Suicide after Talking to AI Chatbot, Widow Says,' The Guardian (Mar. 29, 2023), https://www.theguardian.com/technology/2023/mar/29/belgian-man-dies-by-suicide-after-talking-to-ai-chatbot-widow-says.
[42]AFP, 'AI Generated Deepfakes Disrupt Slovakia's Elections,' The Independent (Oct. 2, 2024).
[43]Chloe Xiang, 'ChatGPT Told Me How to Commit Tax Fraud,' Vice (Feb. 2, 2023).
[44]Reuters, 'Apple to Pay $95 Million to Settle Siri Privacy Lawsuit,' CNN Business (Jan. 2, 2025).

then the developers failed to implement adequate safeguards when the operator is asking such information they could potential e face charges of criminal negligence or abatement charges particularly for the missuse of the AI bought and for a reasonable foreseeable.

Similarly an another instance of chat GPT systems which may be inconsistent of the prompts set up by the operator refusing the illegal promts, sometimes denying request to facilitate in the criminal acts and sometimes complying depending on prompts title, at times the system rejects request and ask the operator to properly refresh the prompt but such inconsistence could be framed as a weakness as the content moderation architect itself by raising the critical questions of developers liability were AI assisted crime was forcible which enable for the assistance of criminal conduct in AI.

VII. Towards a Normative Framework for Criminal AI Liability

Culpability Attribution Models

Risk management theory suggests that AI developers and deployers should be held liable for the standards that match the level of harm that the system might cause. For instance, high-risk systems such as autonomous drones, predictive policy models, or biometric surveillance tools are most subject to oversight due to the severity and foreseeability of harm. The EU AI Act embodies classified AI systems based on different category systems according to the risk and imposition of different principles.[45] This framework should be extended in the application of risk-based liability in criminal law, which involves the burden of proof or introducing strict liability to shift for the high-risk scenarios.

Strict liability regimes may require ensuring that the developer about the design and testing of their systems in a way that the AI system outputs do not generate any harmful instructions to facilitate any type of criminal activity. In the model, to akin product liability law, AI systems with high risk can cause harm inspired of reasonable precautions taken by the developer. The liability would still attach regardless of the intent of the AI systems or developers. This approach prioritises public welfare and supports the utilitarian goals of deterrence and prevention by ensuring that this AI system would there the account ability and provide safety.[46]

[45]Regulation (EU) 2024/…, on artificial intelligence (Artificial Intelligence Act), O.J. L 2024/xx, recitals 65-70.
[46]Mireille Hildebrandt, Criminal Liability for AI Systems: Towards a Mixed Model?, 20(3) ERA Forum 391, 399 (2022)

Regulatory Proposals

Laying down the above highest theory into practice requires a strong regulatory framework. "Mandatory safety rails and disclaimers" should be included in the highest systems should include embedded warnings and limit usage by restrictions and by giving explicit disclaimers about their limitations and potential misuse. Much like the pharmaceutical warnings on the labels about the safety notices, in the same way, these measures communicate risks; a friend can reduce the misuse of these systems.

The other safeguard is "AI sandboxing and pre-deployment scrutiny," which is before the public release of such systems; it should be tested in controlled environments where the behaviour can be observed by an auditor, and these types of cases. This resembles the pilot programming and pharmaceutical medicines used before for safety to demonstrate the use and to ensure that there are no harmful outcomes that can be mitigated before the white spread deployment.[47] Together, this regulatory framework can create a preventive layer ensuring less risk with less potential of criminal consequences, which are relevant risks that can be anticipated, documented, and mitigated before the harm occurs.

Legal Reform Suggestions

India, being the most democratic country, with the laws provided also requiring statutory amendments at present, neither the IPC nor the BNS addresses the provisions related to AI culpability. To bridge the gaps and these provisions by introducing the specific clauses or provisions which define AI-assisted wrongdoing and clarify the liability for developers, deployers, and intermediates is These provisions include amendments that explicitly cover the generator systems, predictive algorithms, and self learning agents.

Another result for the crucial problem is by creating a separate AI-specific liability regime similar to India's existing framework for motor vehicle or environmental harm, which could define the tyres of responsibility that is misusing of AI by users, secondly, developers who designed the system without any adequate safeguards and thirdly, the deployers are the platform who feel to monitor suspicious behaviour at scale. This approach would help the traditional criminal law principles of mens rea and actus rea for experienced users to prevent

[47]European Commission, White Paper on Artificial Intelligence: A European Approach to Excellence and Trust, COM(2020) 65 final, 13.

misuse of the systems while imposing strict liability for the systematic risks generated and high-risk AI frameworks.

Critically, does the framework shift the responsibility away from the non-sentimental codes to the human infrastructure behind it that creates deployees and governance developers, policy makers, and implementors, not the algorithm itself are. The barrels of this responsibility or accountability in criminal law should be addressed not only for the intentional wrongdoing but also negligence caused inside the system and foreseeable risk, as well this shifting and handsome the accountability and foreseeable risk by preserving justice.

By highlighting the transparency, forceability analysis, and proportional responsibility, such a model alliance with both justice-based and prevention-oriented legal goals, the type of framework will offer a path through the "legal maze" by reconciling rapid technological innovation with the proper foundation principles of criminal law that offer a clear path in the legal system.

VIII. Conclusion

Criminal law was never designed to contend with the invisible hands of code. It has always looked for a face to blame, a mind to attribute intention, and a hand to punish. In the modern world the artificial intelligence can be considered for the crime as these artificial intelligences unsettle the equations. It does not think like a human, cannot intend like a human, but still, the question of crime is undeniable in this type of system. They bridge the gap between the traditional legal principles and the modern world's technological capability now demands more hypothetical attention that cannot be debated much for structural rethinking.

The law evolve beyond the notions of guilt at innocence as this is the foundation of base to human actors but when it comes to the machines output which can be more harmful influential and misleading the question cannot simply be the the machines and meant to do it or not but rather unable in to know about the developers and setting the boundaries but failed to restrain its misuse!. Liability must be seen not from a single point, even if that can be the end point, but should be seen as a shared responsibility that should be seen as the contribution across the designer, the developer, the deployer, and the profit-making systems from these AI bots.

These responses are rooted in fear that should be avoided by creating strong legislative frameworks on criminal laws. Does becomes a tool to stunt innovations, but it serves as a

safeguard against negligence, recklessness, and willful blindness and the modern digital age. This is not only about punishing machines, it is about holding accountable the right people, but machines go wrong, and where the intentions fail, the responsibility was not only the development of the legal principle, but also as an ethical imperative for the systems to work correctly in the modern world.

Lastly, as AI continues to advance faster at a pace where the legislation should match with the growth of this technology by providing an opportunity to build something new that mirrors the upcoming frameworks and also the past frameworks, but a model that is agile, transparent, and for the betterment of the people by providing justice. The challenge is not only about the legal but moral as well, that is, to preserve the morality and the protection of human values in the modern world that shapes the future by increasingly shaped by non-human agents.