

# Artificial Intelligence and Criminal Liability

Rishibha Shekhawat

LLM

**Abstract**—The rapid development of Artificial Intelligence (AI) has transformed various sectors, including healthcare, finance, transportation, and law enforcement. However, its increasing autonomy raises significant legal challenges, particularly in determining criminal liability when AI systems cause harm. Traditional principles of criminal law are based on human intention (*mens rea*) and conduct (*actus reus*), which are difficult to apply to machines. This research paper examines the concept of AI, explores the challenges in attributing criminal liability, analyses existing legal frameworks in India, and suggests reforms to address emerging issues. The integration of Artificial Intelligence (AI) into critical sectors has raised complex questions regarding criminal liability, particularly when autonomous systems cause harm. A notable example is the fatal accident involving a self-driving vehicle operated by Uber in Arizona (2018), where the AI system failed to recognize a pedestrian, leading to her death. This incident highlighted the difficulty in attributing criminal responsibility among developers, operators, and corporations when an AI system acts unpredictably. This paper examines such real-world cases to analyse whether existing principles of criminal law especially *mens rea* and *actus reus* can be applied to AI-driven actions. It argues that AI systems lack intent and consciousness, thereby challenging the traditional framework of liability. Through a case-study approach, the research evaluates different models of liability, including the “perpetrator-via-another” doctrine and strict liability, to determine their applicability in AI-related offences.

**Index Terms**—Artificial Intelligence, Criminal Liability, *Mens Rea*, *Actus Reus*, Indian Penal Code, Autonomous Systems

## I. INTRODUCTION

Artificial Intelligence refers to machines or software capable of performing tasks that typically require human intelligence, such as decision-making, problem-solving, and learning. With the advent of autonomous systems like self-driving cars, predictive

policing tools, and AI-based medical diagnostics, the role of AI in society has expanded significantly.

However, as AI systems become more independent, the question arises: who is responsible when an AI system commits a harmful act? For example, if a self-driving car causes an accident or an AI-based financial system engages in fraud, determining liability becomes complex. The traditional criminal law framework struggles to accommodate such situations, making it necessary to re-evaluate existing legal principles.

### 1.1 Application of Legal Principles

To analyse this scenario, it is necessary to apply traditional and modern legal principles:

1. **Perpetrator-Via-Another Doctrine:** Under this principle, a person can be held liable for committing an offence through another agent. In the context of AI, the system may be treated as an instrument. If it can be proven that the programmer or user intentionally used the AI in a harmful manner, they can be held liable.

However, in the present case, there is no evidence of intentional misuse. The harm resulted from a system failure rather than deliberate action.

2. **Negligence-Based Liability:** Negligence arises when a person fails to exercise reasonable care. If the developers failed to adequately test the AI system or ignored known risks, they could be held liable. Similarly, if the user failed to monitor the system when required, liability may extend to them.

3. **Strict Liability:** In cases involving inherently dangerous activities, liability may be imposed without the need to prove intent or negligence. Autonomous vehicles could fall under this category, as they pose significant risks to public safety. Applying strict liability ensures that victims are compensated, regardless of fault.

## Indian Legal Context

In India, the legal framework is not yet equipped to deal with AI-related criminal liability:

- The Indian Penal Code, 1860 is based on human intent and does not recognize AI as a legal entity.
- The Information Technology Act, 2000 addresses cybercrimes but does not cover autonomous decision-making systems.

As a result, courts may rely on general principles such as negligence, vicarious liability, or strict liability to resolve such cases. However, these approaches are not sufficient to address the complexities of AI.

4. Corporate Liability The manufacturing company may also be held liable under corporate criminal liability principles. If the harm resulted from systemic failures within the organization, the company could face legal consequences.

## II. LITERATURE REVIEW

The intersection of Artificial Intelligence (AI) and criminal liability has emerged as a significant area of legal scholarship. Gabriel Hallevy is a leading authority who proposes that AI-related offences can be analysed through existing criminal law models, such as the “perpetrator-via-another” and “natural probable consequence” theories. His work highlights that although AI lacks mens rea, liability can still be attributed to human actors behind the system.

From a technological perspective, Stuart Russell and Peter Norvig, in *Artificial Intelligence*:

A Modern Approach, explain the autonomous and adaptive nature of AI systems. They emphasize the “black box” problem, where decision-making processes are not transparent, thereby complicating legal accountability.

Policy-oriented literature, particularly from the European Commission, suggests a risk-based regulatory framework focusing on accountability, transparency, and human oversight. This approach moves beyond traditional fault-based liability and promotes preventive regulation.

In the Indian context, scholarly work remains limited but highlights the inadequacy of existing laws such as the Indian Penal Code, 1860 and the Information Technology Act, 2000. These laws are primarily

designed for human conduct and do not address autonomous decision-making by machines.

Case-based analyses, including the autonomous vehicle accident involving Uber, further demonstrate the complexity of assigning liability among developers, manufacturers, and users. Overall, the literature suggests the need for a hybrid legal approach combining traditional doctrines with modern regulatory mechanisms

## III. RESEARCH METHODOLOGY

This research adopts a doctrinal and analytical methodology to examine the concept of criminal liability in the context of Artificial Intelligence (AI). The study is primarily based on secondary sources of data, including statutes, case laws, academic literature, and policy reports. The objective is to critically analyse the adequacy of existing legal frameworks and propose suitable reforms. The doctrinal approach involves a detailed examination of legal principles such as mens rea and actus reus, which form the foundation of criminal liability. Relevant provisions of the Indian Penal Code, 1860 and the Information Technology Act, 2000 are analysed to assess their applicability to AI-related offences. This method enables a structured understanding of how traditional legal doctrines interact with emerging technologies.

In addition, the research incorporates a conceptual case study method. Real-world incidents, such as the autonomous vehicle accident involving Uber, are used illustratively to understand practical challenges in attributing liability. Hypothetical scenarios are also developed to explore potential legal outcomes and test the effectiveness of different liability models, including negligence, strict liability, and vicarious liability. A comparative approach is also employed to examine how different jurisdictions address AI-related liability. Legal developments in the European Union, the United States, and the United Kingdom are analysed to identify best practices and emerging trends. This comparative analysis helps in evaluating whether similar approaches can be adapted within the Indian legal system. Furthermore, the research uses an analytical and critical method to evaluate competing theories of liability, such as the “perpetrator-via-another” model and the concept of AI legal personality. These theories are assessed on the basis of

practicality, fairness, and consistency with established legal principles. The study is qualitative in nature, focusing on interpretation rather than empirical data collection. It does not involve fieldwork or primary surveys, as the subject is largely theoretical and evolving.

Overall, this methodology ensures a comprehensive and systematic analysis of AI and criminal liability, combining doctrinal clarity with practical insights to arrive at reasoned conclusions and recommendations.

#### IV. CRIMINAL LIABILITY

Criminal liability with Artificial Intelligence (AI) means deciding who is responsible when an AI system causes harm or commits a crime-like act. Since AI machines do not think like humans or have intentions, applying normal criminal law becomes difficult.

Why is it a problem?

In criminal law, two things are important:

- Actus Reus (guilty act)
- Mens Rea (guilty mind)

AI can perform the act (like causing an accident), but it does not have a mind or intention. So, it cannot be punished like a human.

Example

In the self-driving car accident involving Uber, the car hit a pedestrian. The AI system failed to react properly.

Now the question is:

Who is responsible?

- The company that made the AI?
- The programmer who designed it?
- The person using the car?

This confusion is the main issue in AI criminal liability.

Possible Ways to Decide Liability

- 1) Human Responsibility: AI is treated like a tool. The person using or controlling it is responsible.
- 2) Negligence: If the developer or company was careless (did not test properly), they can be held liable.
- 3) Strict Liability: In risky areas (like self-driving cars), liability can be fixed even without intention.

- 4) Corporate Liability: Companies can be punished if the problem comes from their system or management.

Position in India

India does not yet have special laws for AI crimes. Laws like the Indian Penal Code, 1860 are made for humans, not machines. So, courts may use general rules like negligence or strict liability.

#### V. CONCLUSION AND FUTURE WORK

Artificial Intelligence (AI) is rapidly changing the way society functions, but it also creates serious challenges for criminal law. Traditional concepts like mens rea (guilty mind) and actus reus (guilty act) are based on human behaviour, making them difficult to apply to AI systems that do not have intention or consciousness.

Cases such as the autonomous vehicle accident involving Uber clearly show how complex it is to decide liability when harm is caused by AI. There is often confusion between the roles of developers, manufacturers, and users. Existing laws in India, such as the Indian Penal Code, 1860 and the Information Technology Act, 2000, are not fully equipped to deal with such modern technological issues.

Therefore, it can be concluded that AI cannot be held directly criminally liable. Instead, responsibility must be placed on human actors and organizations connected to the AI system. A balanced legal approach combining negligence, strict liability, and corporate responsibility is necessary to ensure justice while still encouraging innovation.

Future Work

To effectively deal with AI-related criminal liability, several steps need to be taken in the future:

- Development of Special AI Laws: India should introduce dedicated legislation specifically dealing with AI and its legal consequences.
- Clear Allocation of Responsibility: Laws should clearly define the roles and liabilities of developers, companies, and users.
- Adoption of Strict Liability in High-Risk Areas: Sectors like autonomous vehicles, healthcare, and finance should have strict liability rules to protect victims.

- Focus on Explainable AI: AI systems should be designed in a way that their decisions can be understood and explained in court.
- Establishment of Regulatory Authorities: A specialized body should monitor AI systems, ensure safety standards, and handle disputes.
- Continuous Research and Policy Development: As AI evolves, ongoing research and international cooperation will be essential to keep laws updated.

#### REFERENCES

- [1] The Indian Penal Code, 1860.
- [2] The Information Technology Act, 2000.
- [3] G. Hallevy, Liability for Crimes Involving Artificial Intelligence Systems. Springer, 2015.
- [4] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed. Pearson, 2010.
- [5] European Commission, White Paper on Artificial Intelligence: A European Approach to Excellence and Trust, 2020.
- [6] European Parliament, Resolution on Civil Law Rules on Robotics, 2017.
- [7] UK Government, Automated and Electric Vehicles Act, 2018.
- [8] Uber, Autonomous Vehicle Incident Reports, 2018.
- [9] NITI Aayog, National Strategy for Artificial Intelligence, 2018.
- [10] Various research articles and journal publications on artificial intelligence and criminal liability.