AI AND LEGAL LIABILITY: WHO IS RESPONSIBLE WHEN MACHINES GO WRONG?

Shreya Goyal, NIMS University Jaipur Rajasthan

ABSTRACT

Artificial Intelligence (AI) has experienced swift growth across various sectors, such as healthcare, transportation, finance, and legal practices, introducing both distinct opportunities and challenges. A key legal concern that emerges is the issue of accountability when AI systems fail or inflict harm. Conventional legal frameworks, including tort, product, and contract law, often struggle to address the complexities introduced by autonomous decision-making, algorithmic opacity, and the evolution of machine learning technologies. This research paper conducts a comprehensive analysis of the legal principles that currently dictate liability for damages caused by AI, assesses significant case studies, and explores comparative approaches across multiple jurisdictions, including the US, India, and the EU. It also examines the viability of ideas such as digital personhood and frameworks of collective responsibility. The piece proposes the development of a tailored legal framework that integrates human responsibility with algorithmic accountability, ensuring both progress in technology and comprehension of the law. It recommends a united strategy of responsibility and regulatory oversight that can more effectively meet the demands of a future shaped by AI.

Keywords: Artificial Intelligence (AI), Legal Liability, Accountability, Autonomous Decision-Making, Algorithmic Opacity, Machine Learning, Tort Law, Product Liability, Contract Law, Comparative Jurisprudence, Digital Personhood, Collective Responsibility, Human Responsibility, Regulatory Oversight, Tailored Legal Framework, Technological Innovation, AI Governance, EU AI Act, India AI Policy**,** US AI Regulation

CHAPTER 1: INTRODUCTION

Artificial Intelligence (AI) is no longer just a concept from the future confined to science fiction; it has become an integral part of everyday life, influencing sectors such as healthcare, transportation, finance, education, and the legal field.[1] AI technologies are currently responsible for operating vehicles[2], assisting medical assessments, managing investment strategies[3], and delivering predictions in the criminal justice realm[4]. As this rapid development continues, a complex and pressing legal question arises: Who is accountable when AI systems malfunction, yield erroneous results, or cause harm?

Unlike traditional machines or tools which are entirely controlled by human operators, many AI systems function autonomously and make decisions based on algorithms that evolve over time.[5] This autonomy creates challenges for the application of existing legal principles, that typically rely on human intent, control, and predictability.[6] For example, if a self-driving vehicle is involved in a collision or an AI-based diagnostic tool advises incorrect treatment leading to harm, who should be held accountable—the developer, the manufacturer, user, or AI system itself?[7]

Existing legal frameworks like tort law, product liability, and contractual obligations provide certain mechanisms for evaluating responsibility, yet they often overlook the distinct features of AI, including machine learning, reliance on data, and unclear decision-making processes. Furthermore, the absence of consistent global standards contributes to the ambiguity and variation in determining legal accountability.

This paper seeks to explore the deficiencies in the existing legal structures regarding AI liability, examine various comparative methods from different legal systems, and determine whether current doctrines can be modified to tackle these emerging issues. Furthermore, it considers proposals such as creating frameworks for shared accountability, requiring transparency in algorithms, and acknowledging electronic personhood. The primary aim is to find a balanced solution that fosters technological advancement while ensuring responsibility,

[1] Ryan Calo, *Artificial Intelligence Policy*, 51 U.C. Davis L. Rev. 399, 401 (2017).
[2] Karen Hao, *MIT Tech. Rev.* (Mar. 25, 2020).
[3] U.S. Sec. & Exch. Comm'n, *AI in Investment Management* (2021).
[4] *State v. Loomis*, 881 N.W.2d 749, 761 (Wis. 2016).
[5] Thomas Burri, *Machines and Law*, 7 Eur. J. Risk Regul. 439, 443 (2016).
[6] Ugo Pagallo, *Legal Challenges of AI*, 33 Comput. L. & Sec. Rev. 603, 606 (2017).
[7] Shivangi Gangwar, *AI in Medical Negligence*, 48 Indian B. Rev. 32, 35 (2021).

fairness, and the safeguarding of rights in an era increasingly dominated by artificial intelligence.

CHAPTER 2: LEGAL STRUCTURE: RESPONSIBILITY UNDER EXISTING LEGISLATION

The legal framework, which has typically relied on human participants and established technologies, is encountering serious difficulties in determining who is at fault when autonomous artificial intelligence (AI) systems malfunction or inflict harm. The distinctive features of AI, including its autonomy, the lack of transparency in algorithms, and its ever-changing behavior, fall outside the purview of traditional concepts such as tort law, product liability, vicarious liability, and contractual obligations, even though these concepts offer some means for recourse. This section examines the applicability of these legal principles to artificial intelligence and highlights the areas where they fail to assign responsibility.

2.1 Tort Law: Negligence and Strict Liability

2.1.1 Negligence

Negligence, which makes an individual responsible for damages that arise from not exercising proper care, is a core principle of tort law. The key components include a duty of care, a violation of that duty, causation, and the resulting harm. In the context of artificial intelligence, negligence can be attributed to developers, manufacturers, or users who do not anticipate or address the risks associated with using intelligent systems.

Nonetheless, assigning negligence to AI poses significant challenges. Because AI systems frequently evolve and adapt through self-directed processes, it becomes difficult to entirely anticipate their actions. This unpredictability is further exacerbated in situations involving machine learning algorithms that change their conduct over time without human oversight. Additionally, from both legal and technological standpoints, establishing a direct connection between human negligence and a detrimental action taken by AI may prove to be complex.

2.1.2 Strict Liability

Strict liability is imposed on all individuals participating in inherently hazardous activities or in situations involving defective products, irrespective of their intent or fault. Certain uses of

AI, including robotic surgery or self-driving vehicles, could be considered highly perilous.

However, imposing a stringent accountability framework on AI comes with its drawbacks. Identifying the origin or nature of errors in AI systems can be challenging, as they are continually evolving due to algorithmic updates and data processing. Additionally, assessing what constitutes a "fault" in a decision-making AI can be highly subjective and context dependent.

2.2 Product Liability

Retailers and producers are responsible for any harm resulting from faulty products according to product liability laws. AI systems could potentially fall under the same three types of product defects: mistakes in manufacturing, issues with design, and lack of proper warnings.

Design flaws can occur when the structure of an AI system allows for hazardous decisions in certain circumstances. However, as AI technology progresses, it may start functioning in ways that its creators did not intend. In such instances, pinpointing a design defect can become challenging. Neglecting to inform users about the limitations or potential dangers associated with an AI system could also lead to liability concerns. The "black-box" characteristic of numerous AI systems, where the decision-making process is not transparent, further complicates these issues and makes it harder for those harmed to prove liability based on this concept.

2.3 Vicarious Liability

The idea of vicarious responsibility, which holds employers liable for wrongdoing by their employees during work hours, raises theological questions regarding artificial intelligence. Can an AI system be regarded as an agent of the individual or entity that operates or owns it?

Due to the absence of legal personality and intent in AI systems, courts have been reluctant to regard them as conventional agents. This situation complicates the assignment of responsibility since AI does not hold a recognized legal status. Unless the behavior or negligence of the operator can be directly linked to the actions of the AI, vicarious liability is unlikely to provide an adequate legal remedy.

2.4 Contractual Liability

Contract law can offer solutions in situations where AI systems bound by particular agreements cause damage. Responsibility may occur from violations of contracts, warranties, or service-level agreements. For example, a service provider could be liable if an AI diagnostic tool does not fulfil established performance criteria.

Most commercial AI solutions, however, come with licenses that include disclaimers and limit responsibilities. Additionally, agreements often overlook the independent evolution of AI behavior as time goes on. Contractual terms may not adequately address the fair distribution of accountability in cases involving unexpected AI actions.

2.5 Statutory and Sector-Specific Regulations

In various regions, legal structures for data protection, consumer welfare, and accountable algorithm usage have been established. Regulations in particular industries, such as finance, medical care, or transportation, might impose obligations that indirectly govern the application and deployment of AI.

Despite their importance, these regulations tend to be reactive and often fall short of incorporating thorough responsibility clauses that align with the risks and capabilities of autonomous systems. For example, although EU data protection legislation such as the General Data Protection Regulation (GDPR) mandates transparency and accountability, it does not explicitly tackle liability for harm resulting from decisions made by autonomous AI.

2.6 Limitations of Existing Frameworks

While traditional legal systems can be beneficial, they often fail to effectively address key issues related to harm caused by AI:

- Autonomy: AI systems typically function without direct human control or specific programming.

- Opacity: Numerous AI models, particularly those utilizing deep learning techniques, behave as "black boxes" that cannot provide explanations for their decisions.

- Actor Multiplicity: Given that AI development involves a variety of stakeholders, such as manufacturers, users, developers, and data providers, pinpointing accountability becomes challenging.

These shortcomings highlight the necessity for a revised legal framework that considers the realities of contemporary AI systems.

CHAPTER 3: CASE STUDIES AND COMPARATIVE APPROACHES TO AI LIABILITY

To gain insights into how various legal systems are addressing the challenge of AI accountability, this section analyses several case studies that illustrate the actual repercussions of AI malfunctions. It also examines comparable legal measures implemented in the US, India, and the EU to evaluate the suitability and flexibility of different frameworks for governing and determining responsibility for damage inflicted by AI systems.

3.1 Case Studies

3.1.1 Uber Autonomous Vehicle Fatality (Arizona, 2018)

In March 2018, a pedestrian was struck and killed by an Uber-operated self-driving vehicle in Tempe, Arizona. The human safety driver failed to respond in time, while the AI technology incorrectly recognized the pedestrian. Investigations revealed that Uber had disabled the car's emergency braking feature during its autonomous operation.

Legal Result: Uber escaped criminal charges, but the safety driver faced charges of negligent homicide.

The case emphasized the uneven distribution of responsibility among the operator, manufacturer, software developer, and human supervisor. This event demonstrates the accountability void that arises when several entities oversee and manage autonomous artificial intelligence systems.

3.1.2 State v. Loomis (Wisconsin, USA, 2016)

This matter brought before the Wisconsin Supreme Court involved the application of the AIdriven risk assessment tool known as COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) during the sentencing phase. The defendant contended that using the proprietary algorithm infringed upon his due process rights, as its reasoning was unclear and could not be critically examined.

Legal Outcome:

- The court upheld the sentence, stating that the algorithm was only a supporting tool.
- However, it acknowledged the concerns over **algorithmic bias, lack of transparency**, and **explain ability**.

The Loomis case highlights the significance of accountability in algorithms along with the conflict between corporate motivations and basic rights.

3.2 Comparative Legal Approaches

3.2.1 European Union: Risk-Based Regulation and the AI Act

The European Union is adopting a proactive strategy for AI regulation with the proposed Artificial Intelligence Act (AI Act), which received approval from the European Parliament in 2024. High-risk AI systems must adhere to more rigorous compliance requirements according to the AI Act, which categorizes them as unacceptable, high, limited, or minimal risk.

Key Features:

- High-risk AI systems must maintain audit trails and undergo conformance testing.
- Developers and deployers share joint responsibilities.
- The concept of granting AI "electronic personhood" was discussed but ultimately rejected in favor of holding human actors accountable.

Although the AI Act does not create a new system for liability, it improves preventive accountability by introducing requirements for transparency, documentation, and monitoring.

3.2.2 United States: Common Law and Sector-Specific Responses

The United States does not have an all-encompassing federal law governing AI. Rather, issues of liability related to AI are managed through:

- **Laws concerning torts, liability for products, and negligent behavior.**
- **The Federal Trade Commission's guidelines regarding fairness and transparency in algorithms.**
- **Regulations specific to industries, including those for finance (SEC), healthcare (FDA), and transportation (NHTSA).**

The legal system remains primarily reactive. Courts consistently apply established principles to novel situations with varying success. The absence of federal regulation results in state-level discrepancies, creating legal uncertainty for both developers and users.

3.2.3 India: Nascent Framework with Policy Aspirations

The regulations surrounding AI in India are still quite recent. The legal framework mainly utilizes basic tort law, information technology law, and constitutional rights to resolve conflicts related to AI. While the Information Technology Act of 2000 does not address autonomous decision-making, it does deal with matters concerning data.

Policy Initiatives: The "Responsible AI for All" initiative by NITI Aayog emphasizes ethical AI principles such as accountability, transparency, and inclusivity. Currently, there is no specific legal framework or regulations addressing AI responsibility. Indian courts have not made any significant rulings regarding AI liability, even though such cases are likely to increase as the use of AI expands.

India confronts the simultaneous challenge of fostering innovation within its expanding technology sector while also establishing a regulatory and liability framework that guarantees equity, safety, and justice.

3.3 Observations and Analysis

The comparative analysis indicates that there are no consistent global standards regarding AI liability. While the European Union emphasizes preventive governance and classification-driven regulation, the United States depends on private litigation and a patchwork of regulations. On the other hand, India is in the process of drafting its policy and could benefit from both approaches.

Legal systems face similar challenges when trying to apply conventional legal concepts to AI systems that are autonomous, constantly changing, and not transparent.

- One of these issues involves determining the authority within a multi-stakeholder environment.
- Tackling issues of bias and discrimination present in algorithms.
- Guaranteeing that due process is maintained when AI is utilized in governance or the criminal justice system.

CHAPTER 4: EMERGING DOCTRINAL SOLUTIONS AND PROPOSED LIABILITY MODELS

Due to the shortcomings of current legal systems in addressing the harms associated with artificial intelligence, scholars, legislators, and regulators are exploring new legal approaches that reflect the dynamic nature of AI. This section reviews some of the most important emerging concepts, such as the provocative notion of providing AI with legal personhood, the creation of shared liability frameworks, and the necessity for enhanced algorithmic transparency. These developments strive to strike a balance between justice and innovation by closing the divide between technological autonomy and legal responsibility.

4.1 Electronic Personhood: A Contested Proposal

One of the most debated concepts in AI legal theory is the acknowledgment of electronic personality, which advocates for giving autonomous AI systems—especially those capable of making decisions independently—restricted legal recognition.

Arguments in Favour:

- Reflects the idea of corporate legal identity, where entities that are not human possess rights and responsibilities.

- Might enable the allocation of liability when there isn't a distinct human responsible.
- Could promote self-regulation and insurance solutions within sophisticated AI systems.

Criticisms:

- AI does not possess the moral agency, awareness, or intention required to assign legal culpability.
- It might provide a legal safeguard for creators and users by shifting responsibility onto machines.
- There are dangers of eroding essential principles of tort and criminal law by assigning blame to non-sentient beings.

Although it sounds good in theory, the European Parliament officially turned down the proposal in 2017, highlighting that accountability should lie with the human individuals who create, implement, or manage AI systems.

4.2 Shared Liability Model

Given the intricate design of AI and the various human participants involved, numerous researchers and legal frameworks support a collaborative or tiered liability approach that allocates accountability among developers, users, manufacturers, data suppliers, and platform operators.

Key Elements:

- **Developers** may be liable for faulty algorithms, poor training data, or lack of safeguards.
- **Users or operators** may be responsible for misuse, neglect, or failure to intervene when required.

- **Third-party data providers** could bear liability for corrupted or biased datasets that cause discriminatory outcomes.
- **Insurers** can help compensate victims through compulsory AI liability insurance, especially for high-risk applications like autonomous vehicles or medical diagnostics.

This model promotes **risk-sharing** and encourages all parties to adopt best practices, while also ensuring that victims are not left uncompensated due to legal ambiguities.

4.3 Algorithmic Transparency and Explain ability

One of the most pressing concerns in AI liability is the **opacity of algorithmic decision-making**, often referred to as the "black box" problem. Victims harmed by AI decisions may not understand how or why the system reached a particular outcome, making it difficult to challenge or seek redress.

To address this, legal scholars propose mandating:

- **Explainable AI (XAI)**: Systems designed to provide human-understandable reasoning for their outputs.
- **Auditability**: Third-party mechanisms to inspect, monitor, and test AI systems for compliance and fairness.
- **Impact Assessments**: Requiring developers to assess and disclose potential risks prior to deployment (as seen in the EU's AI Act).

Algorithmic transparency not only improves accountability but also ensures compliance with constitutional principles such as **due process, non-discrimination,** and **fair trial rights** when AI is used in governance or criminal justice.

4.4 AI Insurance and Compensation Funds

An increasingly popular proposal is to establish **compulsory insurance schemes** or **AI compensation funds** to provide swift remedies in cases of AI-related harm, particularly where fault is difficult to prove. These mechanisms could:

- Shift the burden from the victim to the system.
- Promote innovation by providing developers with predictable risk coverage.
- Be financed through **risk-weighted contributions** from AI producers and users.

This approach mirrors successful models like **no-fault compensation in motor accident tribunals** and could serve as a transitional solution until doctrinal clarity is achieved.

4.5 Legislative Reform and Regulatory Oversight

Numerous regions are advancing towards regulations tailored to specific sectors and suggesting all-encompassing laws for AI. Proposals for these legislative changes generally consist of:

- Risk classification of AI systems (e.g., low-risk, high-risk).
- Mandatory **safety certification** and **compliance protocols**.
- Establishment of independent **AI regulatory authorities** or tribunals for adjudicating liability claims.

The continuous evolution of the EU AI Act could provide a framework for countries like India to create an effective regulatory system. India has highlighted the importance of these governance frameworks via NITI Aayog and the Digital India initiative, yet has not implemented any legally enforceable laws.

Conclusion of the Section

The changing dynamics of artificial intelligence require a legal approach that is equally flexible and responsive. Although established liability principles are still crucial, they need to be enhanced by creative strategies like shared liability models, requirements for algorithmic transparency, and compensation systems based on insurance. Instead of ascribing personhood to machines, the focus should stay on accountability centered around humans, guided by the ideas of equity, proportionality, and awareness of technology.

CHAPTER 5: CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

The incorporation of artificial intelligence in essential areas of human endeavor has undeniably changed contemporary living, providing enhanced efficiency, precision, and scalability. However, the independence and intricacy of AI systems put strain on the basic principles of legal responsibility, which have historically relied on human intent, foreseeability, and control. As shown, conventional legal doctrines—such as negligence, strict liability, and product liability—although somewhat flexible, are inadequate when it comes to addressing harm caused by self-learning and autonomous AI.

Incident analyses like the fatal accident involving an Uber autonomous vehicle and the application of risk-assessment algorithms in criminal sentencing have underscored the legal, ethical, and procedural challenges stemming from AI. Additionally, a comparative examination shows that while regions like the European Union are implementing forward-thinking, risk-oriented regulatory frameworks, others—such as the United States and India—are facing a disjointed response or remain in the early stages of policy development.

The discussion surrounding electronic personhood for AI highlights the unease in assigning legal accountability to entities that are not human, emphasizing the necessity of preserving a liability framework focused on humans. Considering this, creating shared liability models, enforcing algorithmic disclosure requirements, implementing insurance systems, and introducing legislative changes seem to be the most viable options for achieving both responsibility and progress.

In the end, the legal framework needs to find a delicate equilibrium between promoting the advancement of AI technologies and protecting individuals from potential risks while guaranteeing equitable access to justice. This requires not only an evolution in legal doctrine but also ethical awareness, preparedness within institutions, and collaboration across various disciplines.

5.2 Recommendations

Based on the findings of this research, the following recommendations are proposed to address the legal vacuum surrounding AI liability:

1. **Enact a Comprehensive AI Liability Framework**

   Legislatures should introduce a dedicated law that clearly outlines liability standards for AI developers, users, manufacturers, and operators. This law must recognize the layered nature of AI development and deployment.

2. **Adopt a Shared Responsibility Model**

   Liability should be distributed across the AI lifecycle—ranging from dataset providers to end-users. This promotes accountability at each stage and discourages negligence in both design and application.

3. **Mandate Algorithmic Transparency**

   All high-risk AI systems must be subject to explainability requirements and periodic audits. Regulatory bodies should be empowered to test, review, and assess AI outcomes for bias, errors, and fairness.

4. **Establish an AI Tribunal or Regulatory Authority**

   A specialized body, comprising legal and technical experts, should be created to adjudicate AI-related disputes, determine fault, and recommend sanctions or compensation.

5. **Introduce AI-Specific Insurance Mechanisms**

   High-risk AI applications should be subject to compulsory insurance, ensuring that victims are compensated regardless of litigation outcomes. Funds may be industry-financed, modelled on no-fault compensation schemes.

6. **Encourage Cross-Border Harmonization of AI Laws**

   Given the global nature of AI deployment, international cooperation, and model law development (e.g., via UNCITRAL or OECD) are essential to prevent regulatory arbitrage and ensure consistent standards.

7. **Allocate resources for enhancing public understanding of legal matters and improving capacity**

   Legal practitioners, members of the judiciary, and enforcement officials need to receive training on AI technologies, their legal ramifications, and the concepts of algorithmic fairness.

By implementing these strategies, the legal system can progress alongside AI advancements, guaranteeing that technological growth remains aligned with human rights, justice, and legal responsibility. The future of AI should not only demonstrate intelligence but also embody a sense of responsibility.