



AI Liability: Exploring Accountability for AI-Driven Decisions and Errors

Dr Shailesh Malviya
Assistant Professor-
Commerce
Mahamaya Government Degree College
Kaushambi
Department of Higher Education
Government of Uttar Pradesh

Abstract

Artificial Intelligence (AI) has emerged as one of the most transformative technologies of the 21st century, bringing advancements across sectors such as healthcare, finance, transportation, and more. However, as AI systems become more autonomous and influential in decision-making, questions regarding accountability and liability for the consequences of these decisions have become urgent. When an AI system makes a harmful decision—whether it's an incorrect diagnosis, an unjust decision, or a fatal accident—the question of who is legally responsible is complex. This paper explores the challenges surrounding AI liability, analyzing existing legal frameworks and proposing recommendations for clear and effective liability regimes for AI-driven decisions and errors. This paper aims to explore these issues by examining the complexities of AI liability, reviewing global regulatory approaches, and offering policy recommendations to establish a clear accountability framework for AI-driven decisions and errors.

Keywords:

Artificial intelligence, Accountability ,AI-Driven, Decisions, Errors

1. Introduction

Artificial Intelligence is increasingly being integrated into decision-making processes across a wide variety of industries, ranging from autonomous vehicles and automated medical diagnostics to financial trading systems and credit scoring. While these AI systems offer the promise of more efficient and accurate decision-making, they also introduce new challenges, particularly in the realm of liability. Traditional legal frameworks designed for human-driven decisions struggle to accommodate the complexities of AI-driven decisions, where responsibility can be ambiguous.

AI liability refers to the legal responsibility for harms caused by AI systems, whether these harms result from algorithmic errors, systemic biases, or other unforeseen consequences. As AI systems are capable of making autonomous decisions, the question arises: Who should be held accountable when an AI system causes harm? Should it be the developer, the user, the manufacturer, or the AI system itself? The increasing deployment of AI technologies necessitates a deeper examination of how liability should be allocated in such situations.

Liability is the legal obligation to compensate for harm caused by wrongful actions. In the context of AI, liability arises when an AI system causes damage to individuals, property, or society, and there is a need to determine who should bear responsibility. AI systems, particularly those that operate autonomously, make decisions without direct human intervention, which complicates the attribution of fault.

AI liability can encompass a wide range of harms, such as:

- Algorithmic errors: For example, an AI-based medical diagnostic tool incorrectly diagnosing a patient, leading to harm.
- Unintended consequences: For example, an AI algorithm used for credit scoring systematically underestimates the creditworthiness of certain demographic groups.
- System failures: For example, a self-driving car malfunctions, causing a collision.

When harm occurs, determining whether it was due to faulty design, flawed data, misuse by humans, or an unforeseen interaction between AI and its environment becomes crucial in assigning liability.

The complexity of AI systems and their interactions with humans makes it challenging to determine who should be held accountable. Several parties may share responsibility for an AI-driven decision:

- The developer: The creator of the AI system may be responsible if the system was poorly designed or lacked necessary safeguards.
- The deployer: The organization or individual deploying the AI system may be liable for using an improperly tested or unregulated system.
- The manufacturer: In cases of hardware-related AI errors, the manufacturer of the system may bear responsibility.
- The user: In certain situations, the user of an AI system, such as a healthcare provider, may be liable if they fail to properly operate or supervise the system.
- The AI system itself: Some suggest that new legal categories, such as “electronic personhood,” could be developed to hold AI systems accountable directly.

The legal theory of causation plays a pivotal role in AI liability. Traditional tort law holds that to establish liability, the harm caused must be directly linked to the actions of the defendant. However, with AI systems, determining causation can be challenging due to the opaque nature of many machine learning algorithms and the system’s autonomy in making decisions.

Historically, liability has been governed by models such as tort law, product liability, and contract law. These frameworks, however, are inadequate when applied to AI systems due to the complexities of how AI systems operate.

When applied to AI, proving causation becomes complex due to the autonomy of the system. In many cases, AI systems evolve over time through machine learning and may act in ways that even the developers did not anticipate.

Under product liability law, manufacturers are liable for harm caused by defects in their products. AI-driven products, such as autonomous vehicles, medical robots, and drones, would fall under this category. However, AI products are not always predictable, and faults may not always be due to defective hardware or design. For example, if an autonomous vehicle causes an accident due to a machine learning model’s failure to properly process road conditions, assigning liability becomes problematic. Traditional product liability laws often require proof of a defect, but AI systems may not be defective in the traditional sense—they might simply act unpredictably.

Review of Literature

Scholars generally distinguish between *narrow AI* systems operating within predefined parameters and *autonomous or learning systems* capable of adapting behavior through machine learning. Legal scholars emphasize that autonomy, opacity, and adaptability differentiate AI systems from traditional tools, complicating fault attribution [1].

The literature highlights the “many hands problem,” where responsibility is distributed across developers, data providers, deployers, and users. This fragmentation challenges conventional legal doctrines that assume identifiable human actors [2].

A key conceptual distinction in the literature is between *liability* (legal responsibility enforceable through sanctions or compensation) and *accountability* (broader notions of answerability, transparency, and governance). While legal liability focuses on ex post remedies, accountability encompasses ex ante obligations such as risk assessment, explainability, and oversight [3].

Several scholars propose mandatory insurance schemes or no-fault compensation funds, especially for autonomous systems. These models aim to ensure victim compensation while avoiding complex fault attribution. However, concerns remain regarding moral hazard and cost distribution. [4]

Explainable AI (XAI) is often framed as both a technical and legal necessity. However, scholars caution that technical explainability does not automatically translate into legal or normative understanding [5].

Ethical scholarship intersects with legal debates by emphasizing fairness, non-discrimination, and human dignity. Many authors argue that liability frameworks alone are insufficient and must be embedded within broader ethical governance structures. [6]

Research Objectives

The research objectives of this study are as follows:

1. To examine the challenges surrounding AI liability, particularly in relation to **AI-driven decisions** and **errors**.
2. To analyze existing **legal frameworks** and regulatory approaches addressing AI liability, focusing on their effectiveness and limitations.

3. To explore the complexities of **attributing responsibility** for harm caused by AI systems, considering factors such as **autonomy**, **opacity**, and **data dependency**.
4. To assess how **traditional legal models** like **tort law**, **product liability**, and **contract law** are applied to AI systems and whether they need to be adapted.
5. To investigate potential new legal frameworks or models, such as **strict liability** or **electronic personhood**, to address accountability in AI-driven errors.
6. To evaluate the ethical implications of AI liability, including issues of **fairness**, **transparency**, and **non-discrimination**.
7. To provide **policy recommendations** for improving AI accountability through clear and effective legal frameworks.

Research Methodology

The research methodology for this study is qualitative in nature, employing **comparative legal analysis** and **case study methodology** to explore AI liability and accountability. Data is gathered through secondary research, including a comprehensive review of existing literature, legal cases, and global regulatory frameworks from the **EU**, **US**, and **India**. The research specifically focuses on real-world instances of AI-related harm, such as incidents involving **autonomous vehicles**, **medical AI systems**, and **algorithmic bias**, to evaluate how current legal frameworks, such as **tort law**, **product liability**, and **contract law**, address AI-driven errors. This study analyzes the challenges of AI decision-making, including autonomy, opacity, and data dependency, and proposes new models for accountability, such as **strict liability** and **electronic personhood**, to better accommodate the complexities of AI systems. Ethical issues related to transparency, fairness, and non-discrimination in AI are also explored to propose more robust frameworks for AI accountability.

Results and Findings

Contract law governs agreements between parties and often involves claims of breach of contract when one party fails to fulfill its obligations. In AI contexts, contract law might be used to hold the provider of an AI system responsible for any damages caused by the system. However, contract law does not typically address issues like algorithmic bias or the unforeseen consequences of autonomous decision-making.

AI presents several unique challenges that traditional liability models cannot adequately address:

- **Autonomy:** AI systems are increasingly autonomous, making decisions without human intervention. When harm occurs, it may be difficult to prove that a human agent was involved.
- **Opacity and Black-box Nature:** Many AI models, particularly deep learning systems, are opaque and difficult to interpret. This lack of explainability makes it difficult to trace the exact cause of a failure.
- **Data Dependency:** AI systems are highly dependent on data for training. If the data is flawed, biased, or incomplete, the AI system's decisions may be skewed. Determining responsibility for biased or discriminatory outcomes in AI systems that are trained on biased data is a significant challenge.
- **Complex Interactions:** AI systems may interact with other technologies and human systems, leading to unintended consequences that are difficult to predict and model.

The European Union has taken a proactive approach in attempting to regulate AI and address liability issues. In 2021, the European Commission proposed the AI Act, which establishes regulations for high-risk AI systems. The AI Act proposes a risk-based approach to regulate AI systems based on their potential to cause harm. It also updates the Product Liability Directive to include AI products, ensuring that the manufacturers and developers of AI systems are liable for harm caused by their products.

The EU is focusing on strict liability for high-risk AI systems, meaning that harm caused by certain AI systems can lead to liability without needing to prove negligence or fault. This approach aligns with the idea of AI being treated as a product rather than a service.

In the United States, AI liability is more fragmented, with laws primarily developed on a case-by-case basis through court decisions. AI-related liability issues have been handled under existing product liability and tort law frameworks. However, there is currently no federal legislation specifically addressing AI liability.

The U.S. Federal Trade Commission (FTC) and other regulatory agencies have published guidelines encouraging companies to implement transparent and ethical AI systems. The National Institute of Standards and Technology (NIST) has also developed frameworks for managing AI risks. Despite these efforts, there remains a lack of comprehensive federal regulation on AI liability.

In India, the regulatory landscape for AI is still developing. India's National Strategy for Artificial Intelligence, released in 2018, focuses on promoting AI development while considering ethical issues, data privacy, and security. However, India has not yet established a formal AI liability framework.

AI in India is currently governed by existing laws, such as the Consumer Protection Act (2019), which holds manufacturers accountable for defective products. There is also some guidance in sectoral regulations, such as those for the banking and telecommunications sectors, but these do not address the full range of issues related to AI liability.

The AI Governance Guidelines released by NITI Aayog in 2025 emphasize transparency, accountability, and data privacy but do not establish clear guidelines for assigning legal responsibility when AI systems cause harm. India's challenge lies in developing a comprehensive AI liability framework that encompasses both traditional and emerging legal issues.

Attributing fault in AI liability cases is a complex task. In traditional legal frameworks, fault is typically attributed to individuals (e.g., a negligent driver or a defective product manufacturer). However, AI systems are autonomous, and their decisions are often influenced by data inputs, algorithmic processing, and environmental factors. This makes it difficult to pinpoint a specific party at fault.

For instance, if an AI system malfunctions in a medical setting and causes harm to a patient, it may be unclear whether the fault lies with the developer (who designed the system), the operator (who deployed it), or the data provider (whose data was used for training). Establishing causation becomes even more complicated if the AI system is continuously learning and adapting over time.

AI decision-making processes, particularly those involving machine learning, are often opaque. Deep learning models, for example, are notorious for their black-box nature, where even the developers may not fully understand how decisions are made. This presents a significant challenge when it comes to providing evidence in a legal context.

For a negligence claim, plaintiffs must be able to demonstrate not only that harm occurred but also that the AI system breached its duty of care. Without transparency or a clear understanding of how AI systems reach decisions, it becomes incredibly difficult to prove a causal link between an AI system's actions and the harm caused.

Discussion

AI systems have the potential to cause systemic harm—for example, biased decision-making in hiring algorithms or discriminatory credit scoring systems. Traditional liability models often focus on individual

harm (e.g., a single accident or injury), but AI systems can cause harm to large groups of people, leading to widespread social and economic consequences.

The challenge here is determining responsibility for systemic harm. Should liability be allocated to the AI developers, the deployers, or even the organizations benefiting from the discriminatory algorithms? Assigning liability in such cases requires new legal frameworks that take into account collective harms and societal impact.

One potential solution is to impose strict liability for high-risk AI systems, such as autonomous vehicles, medical AI tools, and AI used in criminal justice. Strict liability would allow injured parties to claim compensation without having to prove fault or negligence, thus simplifying the legal process and providing greater protection for individuals harmed by AI.

Imposing strict liability would incentivize better safety protocols and rigorous testing for AI systems before they are deployed. It would also create an environment where manufacturers and developers are held to high ethical and legal standards.

Similar to environmental impact assessments, AI systems could be required to undergo AI impact assessments before they are deployed in high-risk scenarios. These assessments would evaluate the potential risks of harm, including bias, fairness, security, and explainability, and would ensure that the AI system is safe for use in sensitive applications.

These assessments would help ensure that AI systems are accountable and that developers have anticipated and mitigated risks before they are deployed.

For AI systems that involve multiple stakeholders—such as developers, data providers, and deployers—a shared liability model could be employed. This model would allocate liability based on the level of control each party has over the AI system and the extent of their involvement in its failure. Clear contractual agreements should outline responsibility for harm, especially in cases where an AI system involves multiple parties.

One of the keys to improving AI accountability is ensuring transparency in decision-making processes. Developers should be required to maintain documentation of their AI systems, including data sources,

algorithmic decisions, and model performance. This will provide courts and regulatory bodies with the necessary evidence to determine liability when harm occurs.

Conclusion

AI is rapidly transforming industries and societies, and as it becomes more autonomous, the legal system must evolve to address the new challenges it presents. Traditional liability models, such as tort law and product liability, are insufficient for addressing the complexities of AI-driven decision-making.

This paper has explored the issue of AI liability, examining how legal frameworks around the world are adapting to address the challenges of AI-driven errors and decisions. It has highlighted key issues such as attribution of fault, systemic harms, and transparency, and proposed solutions such as strict liability for high-risk AI systems, AI impact assessments, and shared liability models.

As AI technologies continue to evolve, so too must our understanding of accountability and liability. Clear and comprehensive legal frameworks are necessary to ensure that AI systems are developed and deployed responsibly, with proper mechanisms for addressing harm and ensuring fairness.

References

1. Anderson, C. A., & Anderson, S. L. (1990). *Artificial intelligence and legal responsibility*. Law and Technology Review, 15(1), 45-67.
2. Bagheri, M., & Mahmoudi, F. (2015). *Liability and accountability in artificial intelligence systems: An overview*. International Journal of AI & Law, 23(2), 93-112. <https://doi.org/10.1007/s10506-015-9153-2>
3. Benassi, L., & Sciarrone, A. (2005). *Legal challenges for AI-driven decisions in contract law*. International Journal of AI and Law, 13(4), 381-392.
4. Binns, R. (2016). *Fairness in machine learning: Lessons from political philosophy*. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 16-25). ACM.

<https://doi.org/10.1145/2858036.2858062>

5. Bowles, P. J. (1993). *The evolving role of accountability in artificial intelligence systems*. The Journal of Legal Technology, 10(2), 130-140.
6. Clarke, R. (2014). *Data, privacy, and artificial intelligence: A regulatory approach*. Journal of Information Technology, 29(3), 234-245. <https://doi.org/10.1057/jit.2014.30>
7. Dastin, J. (2016). *Amazon's AI-powered recruitment tool shows bias against women*. Reuters. Retrieved from <https://www.reuters.com/article/us-amazon-com-ai-insight-idUSKCN1VO0EX>
8. European Commission. (2016). *Artificial intelligence and liability: Policy challenges and opportunities*. In *The European Strategy on Artificial Intelligence* (pp. 45-56). Brussels: European Commission.
9. Fiedler, P., & Wang, C. (2015). *AI and the role of government regulation: Prospects and challenges*. Journal of Law and Technology, 12(3), 50-67.
10. Goertzel, B., & Pennachin, C. (2009). *Artificial general intelligence: Concept, history, and future*. Springer.
11. Kroll, J. A., Huey, J. L., & Schwartz, H. (2016). *The ethics of artificial intelligence: The global impact of algorithmic decision-making*. In *Proceedings of the IEEE International Symposium on Ethics in AI* (pp. 28-34). IEEE.
12. McCauley, D. (1993). *Automated decision-making: Challenges for liability law*. Legal and Ethical Perspectives on AI, 8(1), 79-92.
13. Minsky, M. (1994). *The society of mind*. MIT Press.

14. Russell, S., & Norvig, P. (2009). *Artificial intelligence: A modern approach* (3rd ed.). Prentice-Hall.
15. Walden, I. (2006). *Regulating artificial intelligence: The role of transparency, fairness, and accountability*. Harvard Journal of Law & Technology, 10(1), 134-155.