

Liability for Harm Caused by Autonomous Ai Systems

Ms Purnima Gautam¹, Somya Singh²

¹ Assistant Professor, Kanoria School Of Law For Women, Jaipur

² Ballb 5th Year Law Student At Jaipur National University

ABSTRACT

“Artificial intelligence is not inherently good or bad. It is a tool. The responsibility lies with us, the humans who build and use it.”

Tim O'Reilly

AI has emerged as the most powerful and widely used technology in today's era. We can witness its presence in almost every field, reshaping industries, governance, law, and society, and transforming every sector, integrating critical domains such as healthcare, education, transportation, criminal justice, and national security etc. No doubt, it has also created numerous opportunities, innovations, progress, and development for the country's upliftment and betterment. But as we know that every coin has two sides, similarly, at the very same time, it has raised reflective legal, ethical, and policy challenges that become the pressing issue for today.

From a legal perspective, Liability in the case of damage done by AI became the foundation issue for concern and needs immediate attention on the same. When AI systems such as self-driving cars or medical diagnostic tools cause harm, it is often unclear who should be held responsible: the developer, the manufacturer, the user, or the AI system itself. Not only this, there are n' number of cases where ascertaining the liability became next to impossible. Data protection and privacy concerns also add fuel to the fire. Although policymakers are attempting to create safeguards, AI's capacity often outperforms existing regulations.

The ethical issue arises from the question of what happens if AI is biased and partial. What if their decisions are influenced by certain factors or prejudices? These biases can disproportionately disadvantage marginalized groups, undermining the principle of equality before the law. Related issues include transparency and explainability. The lack of interpretability poses serious risks in criminal sentencing, healthcare diagnostics, and financial decision-making. Self-sufficiency and human dignity are also at risk when AI begins to replace or significantly influence human policymaking and administrative functions. Additionally, it threatens to manipulate and influence public opinion, spread misinformation, and cause various other problems. Fairness, accountability, transparency, and respect for human dignity must be safeguarded.

Both at the national and international levels, countries and international organizations are engaging in creating such effective and efficient policies and frameworks that emphasize transparency, safety, and accountability. In India, the NITI Aayog's National Strategy for Artificial Intelligence (2018) and subsequent initiatives emphasize the development of AI for social good, particularly in healthcare, agriculture, and education. India's Digital Personal Data Protection Act, 2023, further aims to regulate

how data is collected and processed in AI systems. The United States has also adopted a more flexible, innovation-driven approach, exemplified by the AI Bill of Rights (2022), which outlines principles such as safe and effective systems. The European Union has taken a ground-breaking step with its AI Act (2024), which adopts a risk-based approach by categorizing AI systems into prohibited, high-risk, limited-risk, and minimal-risk categories.

Artificial Intelligence poses an unprecedented opportunity and a significant challenge to humanity. The legal system will have to adapt to redefine liability, intellectual property rights, and criminalizing behaviour in a world with smart machines. AI's ethical use demands frameworks to preserve dignity, autonomy, and equality. Additionally, policies at the national and international levels will need to find a balance between safeguarding innovation and fundamental rights.

Keywords: *Artificial Intelligence (AI), legal liability, intellectual property rights (IPR), data protection and privacy, algorithmic bias, transparency and explainability, human dignity and autonomy, predictive policing, ethical governance, AI regulation, the EU AI Act 2024, the AI Bill of Rights, India's Digital Personal Data Protection Act 2023, global AI governance, and responsible innovation.*

1. INTRODUCTION

"When an autonomous system causes harm, the legal system must decide who, if anyone, should be held responsible. But assigning liability in the absence of intent or direct control challenges traditional legal doctrines."

Ryan Calo, Legal Scholar in Robotics and Law

With the advent of advanced and sophisticated technologies, the innovation of Autonomous AI Systems marks a transformative movement in the entire history of mankind. From automatic and self-driven cars to robotics services in various sectors, such as healthcare, hospitality, and serving as a waiter in hotels and restaurants, one can see a significant and drastic revolution in the field of Artificial Intelligence and technological advancements. In simple terms, Autonomous AI Systems means such artificial intelligence that can think, perform tasks, and make decisions independently without any human interference or intervention.

The growing demand and influence of Autonomous AI in every aspect of our daily lives are proving very beneficial and convenient. Unlike the previous obsolete software, these are more capable of learning and adapting to new things according to their surroundings, and are advanced in making their own decision and performing complex tasks without any intervention or command from humans¹.

As every coin has two sides, similarly, these Autonomous AI Systems pose difficulties and put forward the important yet deliberative questions about accountability and liability. What happens when a self-driving car causes an accident? Who is liable if a medical AI system makes an incorrect diagnosis? These questions point out the lacunae and the loopholes present in the most advanced technology, which was overshadowed by its few capabilities. Unlike humans, who are natural persons, having their morality, consciousness, intent, or legal personality, can sue and can be sued, AI, on the other hand, have nothing in common and making it difficult to assign blame or liability directly to the machine².

¹ Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 Ga. L. Rev. 109 (2017), <https://digitalcommons.law.uw.edu/faculty-articles/572/>.

² Bryant Walker Smith, *Automated Driving and Product Liability*, 2017 Mich. St. L. Rev. 1, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2923240.

These systems expose users to harm and injury, including physical, mental, privacy infringement, damage to property, and loss of confidentiality and authenticity. It creates a legal vacuum when it comes to assigning liability for the actions taken by the machine. As a result, ascertaining the fault became a huge and complex task, especially when multiple parties are involved in the use of the AI system, such as software developers, hardware manufacturers, data providers, and end-users. It became necessary to strive for a perfect balance between technological innovations and ensuring proper compliance and prevention of any grievances to the users and the public.

The current legal frameworks were seen as inefficient, inadequate, and ineffective in resolving the disputes revolving around the autonomous AI systems. Law of Negligence, Strict Liability, Vicarious Liability, and Product Liability; all these existing laws are not responsive and are not able to retaliate against the harmful effects of the autonomous systems³.

This article explores the existing legal doctrines, legal challenges, legal framework, global responses, and proposed models for assigning liability in cases involving harm caused by autonomous AI systems.

THE NEED FOR LEGAL ACCOUNTABILITY

Imagine a place where technology controls everything, where drones deliver groceries, self-driving taxis take passengers to their destinations, and AI doctors diagnose patients. Everyone's life becomes seamless and convenient until one day, a self-driving taxi breaks down and crashes into a pedestrian, a drone delivers the goods to the wrong address, and an AI doctor kills someone by giving the wrong medicines instead of the right ones. Then the question arises: Who is responsible? There are certain plausible reasons why we need legal accountability⁴:

The very first reason is the protection of the Public Interest. An individual should be protected from harm and injury, whether physical, mental, financial, or any emotional trauma accrued by the AI, and it can be possible only when there is some kind of Legal Accountability. Without clear laws, victims of AI errors may be left without justice.

We all refer to the traditional doctrines like negligence or intentions whenever we need to ascertain the liability or the fault of the person, but as AI lacks morality, consciousness, intent, or legal personality, it becomes difficult to apply those doctrines. This creates a legal vacuum that can only be filled with accountability laws.

In addition to this, when there is the existence of accountability laws, the developers and companies also have a mindset to adopt safety-by-design, conduct risk assessments, and ensure more carefulness in designing the system, as they will be held accountable for the casualties that happen to the users. This will work as a deterrence theory, which will create a fear of punishment in the minds of those developers and companies, and they will not put innocent lives at stake.

Technology without responsibility, and Laws without accountability are just like a compass without a needle. As the latter will not be able to guide the direction, the former cannot guide justice. The reparation to the victims for their losses is a necessary part, and accountability ensures that victims will receive that, and there's a legal path to that.

³ *Restatement (Third) of Agency* § 2.04 (Am. L. Inst. 2006), <https://www.ali.org/publications/show/agency/>.

⁴ Karen Yeung, *Algorithmic Regulation: A Critical Interrogation*, 12 *Regulation & Governance* 505 (2018), <https://doi.org/10.1111/rego.12160>.

Lastly, as AI operates across borders, unambiguous liability frameworks help to establish global standards, ensuring AI doesn't become a legal loophole. Accountability makes international collaboration more practical and ensures global justice, interaction, cooperation, and facilitates more friendly and harmonious approaches⁵.

In the game of chess, every piece has rules. But what if a rook started flying? Or did the queen make her own rules? That's what AI without accountability looks like: a board where strategy is meaningless because nothing is held accountable. Legal accountability places rules back on the board, so the game remains fair and unbiased.

Challenges in Attributing Liability

Attributing liability in cases involving autonomous AI systems presents challenges those traditional legal doctrines struggle to address. They are as follows;

The existing doctrines of negligence, mens rea, cannot be applied in those conditions where AI lacks consciousness, intent, and legal personality. Absence of Human command made it difficult to ascertain and establish the liability based on intention and negligence.

Working on its own intelligence, AI's decisions become complex and uncertain as its work behaviour is based on learning algorithms, making it difficult to establish the link between the working and the harmful outcome, as AI decisions are not always predictable⁶.

There are multiple actors involved in the line of bearing the responsibility, and ascertaining who should bear responsibility is a bit challenging. They can be:

- **Developers** (who write the code)
- **Manufacturers** (who create the hardware)
- **Data providers** (who supply training data)
- **End-users/operators** (who deploy the AI)

AI's decision-making processes are not easily understood or traceable because they use a complex model and technology to function, making it hard to identify errors or assign responsibility for harm.

AI systems operate across borders. Determining which legal system applies or whose country's court will have the power of jurisdiction can complicate liability attribution.

Biasness and incomplete data may pose more problems as it becomes difficult to ascertain the correct information. They can cause financial as well as economic losses, and neither liability lies with the data provider, the developer who failed to correct it, or the AI that acted on it.

AI systems can learn and evolve post-deployment. Any new changes and developments may affect past learnings, making it difficult to handle problems effectively and efficiently.

Multiple actors (developers, integrators, third-party providers) might all contribute indirectly to harm. Dividing liability fairly among parties in proportion is very difficult, as nothing is certain and confirmed⁷.

⁵ Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 Harv. J.L. & Tech. 353 (2016), <https://jolt.law.harvard.edu/assets/articlePDFs/v29/29HarvJLTech353.pdf>.

⁶ Luciano Floridi & Josh Cowls, *A Unified Framework of Five Principles for AI in Society*, 5 Harv. Data Sci. Rev. (2021), <https://doi.org/10.1162/99608f92.8cd550d1>.

⁷ Luciano Floridi & Josh Cowls, *A Unified Framework of Five Principles for AI in Society*, 5 Harv. Data Sci. Rev. (2021), <https://doi.org/10.1162/99608f92.8cd550d1>.

Legal systems rely on precedents to interpret liability. Autonomous AI is a new field with few decided cases, leaving courts without any proper guidance.

If no one is held liable for AI-caused harm, there may be reduced incentives for safe design. On the other hand, imposing liability without clear fault may lead to unfair burdens on AI developers or users⁸.

RELEVANT ILLUSTRATIONS

Here are a few real-life illustrations of how the Autonomous AI can be proven fatal and dangerous for the users as well as for the public at large.

➤ SELF-DRIVEN CARS

Although it seems to be very fascinating that the cars have the auto-pilot mode, making it very convenient for the other person to reach their destination without driving themselves, who will be held liable if there is any accident caused by that self-driven car because it fails to detect the pedestrian on the road, whether it was the car owner, car manufacturer, the coder of the algorithm, or the supplier of the faulty sensor?

➤ MEDICAL MISDIAGNOSIS

AI or Robotic technology, which is used for the diagnosis of the patient's condition, and what will happen if it misdiagnoses the health condition and medical requirement of the patient, and exposes him to risky surgery, which can be detrimental to him who will be responsible if the patient dies or sustain any grievous injury, whether it was the hospital, the developer of the tool, or the doctor responsible?

➤ WRONG SURVEILLANCE

If there is a crime has taken place and with the help of an AI surveillance system, the criminal was detected and punished accordingly, but after some time, it was realised that the actual criminal was still out and moving freely while the innocent one was punished and it is all because of the glitch happened in an AI surveillance system, who will be held responsible, whether the law enforcement agency, the AI software provider, or the data trainer?

➤ AI SYSTEMS IN BANKING AND FINANCIAL SECTOR

What if an AI system faces a certain setback and goof up with the credit and debit of money in the account of a person other than the one in whose account it should be deposited actually, or what will happen if the application of a serious candidate for the job or the qualified applicant who is seeking loan was rejected due to the mismanagement of AI System, who will be responsible, the bank, the customer, third-party AI developer, or the creator of the learning algorithm?

It is not the end of the list of the blunders of an Autonomous AI System, but if we do not take any requisite measures and essential steps, the day is not too far away when a lot of destruction will happen, and the list will be increasing day by day with such incidents.

INTERNATIONAL POSITION ON LIABILITY FOR AUTONOMOUS AI SYSTEMS

European Union (EU) – A Global Leader in AI Regulation:

The European Union is leading the process of creating a functional regulatory framework for artificial intelligence, particularly regarding safety, ethics, and liability.

A. EU Artificial Intelligence Act (AI Act, 2024)

⁸ Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 Ga. L. Rev. 109 (2017), <https://digitalcommons.law.uw.edu/faculty-articles/572/>.

- The first comprehensive law on AI in the world, adopted in 2024⁹.
- AI systems are classified into four categories:
 - Unacceptable Risk (banned)
 - High Risk (subject to strict compliance)
 - Limited Risk (transparency obligations)
 - Minimal Risk (free use)

B. Product Liability Directive & AI Liability Directive (2023 Proposal)

PLD is being revised to explicitly consider AI systems as “products”. Proposes strict liability for any damages caused consequence AI systems, physical injuries, property damage, etc., harms that are expected with easy-to-predict circumstances in complex machinery built by humans. AI Liability Directive seeks to shift the burden of proof: victims of harm resulting from AI technology may only need to establish causation without proving fault.

OECD (Organisation for Economic Co-operation and Development):

OECD Principles on Artificial Intelligence (2019)¹⁰

- Endorsed by over 40 countries, including the U.S., EU states, Japan, India, etc.
- Five core principles:
 1. Inclusive growth and sustainable development
 2. Human-centred values and fairness
 3. Transparency and explainability
 4. Robustness and security
 5. Accountability for AI actors

UNESCO – Global Ethical Framework¹¹

◆ UNESCO Recommendation on the Ethics of Artificial Intelligence (2021)

- Adopted unanimously by 193 countries.
- Emphasizes ethical AI use, protection of human rights, and establishing liability mechanisms for harm caused.
- Calls for:
 - Risk assessment and human oversight
 - Legal remedies for victims
 - Clear assignment of responsibility across the AI value chain

⁹ **European Commission**, *Proposal for a Directive of the European Parliament and of the Council on Adapting Non-Contractual Civil Liability Rules to Artificial Intelligence (AI Liability Directive)*, COM (2022) 496 final (Sept. 28, 2022), available at <https://eur-lex.europa.eu/LEGAL-CONTENT/EN/TXT/?uri=CELEX%3A52022PC0496>

¹⁰ Org. for Econ. Co-operation & Dev. [OECD], *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449 (May 22, 2019), <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

¹¹ UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, SHS/IGM-AIETHICS/2021/1 (Nov. 25, 2021), <https://unesdoc.unesco.org/ark:/48223/pf0000381137>.

Council of Europe (CAHAI)◆ Committee on Artificial Intelligence (CAHAI) Report (2022)¹²

- Recommends a legally binding convention on AI and human rights.
- Focuses on:
 - Transparency of AI decision-making
 - Redress mechanisms for individuals
 - Civil and criminal liability models

United Nations (UN) – Soft Law, but Growing Involvement¹³

- The UN Secretary-General's Roadmap for Digital Cooperation (2020) stressed the need for global AI governance frameworks.
- The UN Human Rights Council has cautioned against AI systems that interfere with privacy, freedom of expression, and due process.
- The ITU (International Telecommunication Union) and WIPO (World Intellectual Property Organization) are working on AI in telecom standards and IP law, respectively.

United States

- No single AI law, but sectoral regulations and court rulings apply.
- NHTSA (transportation), FDA (medical devices), and FTC (consumer protection) regulate AI systems.
- Courts use common law doctrines (negligence, strict liability, product liability) for AI harm.
- Blueprint for an AI Bill of Rights (2022) recommends:
 - Right to safe & effective systems
 - Right to protection from algorithmic discrimination
 - Right to explanation and recourse

China

- Released AI regulation draft (2021) emphasizing¹⁴:
 - No algorithmic discrimination
 - Mandatory accountability for deepfakes and recommendation systems
 - High penalties for violating safety and content obligations

CONCLUSION

There is no doubt that an Autonomous AI system has marked significant growth in the technology, offering a wide range of benefits, including efficiency, convenience, and innovation. However, it also poses a challenge of determining the harm caused by the non-human element capable of independent decision-making, posing a threat to traditional legal frameworks.

The lack of well-defined liability structures to govern AI technologies opens up possibilities of victims being left without compensation. Developers may also lack the motivation to create safe systems. It also

¹² Council of Eur., Ad hoc Comm. on Artificial Intelligence [CAHAI], *Feasibility Study on a Legal Framework on Artificial Intelligence Design, Development and Application Based on the Council of Europe's Standards* (Dec. 2022), <https://www.coe.int/en/web/artificial-intelligence/cahai>.

¹³ U.N. Secretary-General, *Roadmap for Digital Cooperation*, U.N. Doc. A/74/821 (June 2020), <https://www.un.org/en/content/digital-cooperation-roadmap/>.

¹⁴ Cyberspace Admin. of China, *Regulations on the Administration of Algorithmic Recommendations in Internet Information Services (Draft)* (Aug. 2021), unofficial English translation available at <https://digichina.stanford.edu/>.



poses serious ethical questions, especially in instances with harm results from decisions made by an automated system without direct human control. The situation becomes more complex when there is an array of different participants, such as the developers, manufacturers, data providers, and users, all of whom add to the complexity of determining blame.

To ensure AI systems help out in ways that are safe, respectful, and uphold human dignity, nations need to enact strong and flexible laws that balance individual rights with the need for responsible advancements. As such, AI will have to be governed with the same level of sophistication in laws that are meant to ensure blame is assigned properly, justice is maintained, and the public retains faith in technology as it grows more intelligent and autonomous.