FRAMEWORK FOR ADDRESSING LIABILITY AND ACCOUNTABILITY CHALLENGES DUE TO ARTIFICIAL INTELLIGENCE AGENTS

Ayush Gupta, Indian Law Institute

ABSTRACT

The current legal landscape worldwide, especially in India, is lagging behind the pace with which artificial intelligence capabilities are increasing in public space. The autonomous action of AI agents means that the dynamics of determination of liability or culpability are bound to come into question of law since the current laws only incriminate or place accountability on persons. Since personhood is an attribute of only a human being, this paper seeks to establish a comprehensive framework for identifying the implication of AI agents on existing laws and how liability and accountability may be placed in the new dynamics of artificial intelligence systems. The analysis considered all the latest and relevant developments in international and national legal systems till March 2025 to arrive at the principles and factors required to develop such a framework which can provide necessary implementation guidelines for a more just and fair legal system to address the impending challenges. While a single framework is important to cater to a comprehensive set of liabilities which may arise due to actions of AI agent systems being used today and in future, multiple legislative and statutory provisions are required to ensure that legal remedies are made available to all who may seek justice due to losses which may be suffered due to the actions of AI agent systems.

Introduction

Introduction: Context

The determination of liability for AI agents is an evolving area of law with a lack of clarity on how to place the responsibility for a wrong committed or recover damages caused due to the actions of an AI agent. Technological developments in recent years in the field of artificial intelligence sphere have made significant strides in the possibilities of real-life autonomous agents, in the form of software or a robot, acting based on only algorithms & programming which take external stimuli as input to generate a response. In the real-world scenario, such agents may interact with human beings or operate in ecosystems in which their actions can affect human beings. They are expected to behave in a certain manner and should exhibit a standard level of quality. Examples of AI agents commonly used in current times maybe that of self-driving cars, virtual assistants like Siri or Alexa, Open AI's ChatGPT, etc.

Introduction: Research Problem

In cases where an AI agent system makes an error or causes harm, how is the liability determined? Is it the developer of the AI agent, the user, or the AI agent itself who is liable?

Introduction: Objectives

The purpose of this research is to identify the appropriate agency on which liability and accountability can be placed so that fair and just compensation is accorded to the victim and necessary punishment is served to the wrongdoer for the prevention and control of future crimes.

Introduction: Research Questions

- What are the important qualifying conditions and contributing factors for placing the liability for harm caused by an artificial intelligence agent system?
- Who is liable for harm caused by an artificial intelligence agent system? How does it change in different scenarios?

Introduction: Significance

The absence of any human intervention or supervision needed for acts of an artificial agency which can interact and cause harm leaves a legal void in the current legal landscape which is yet to be completely figured out in the face of unanticipated challenges due to fast-evolving technology and unforeseen circumstances. The artificial intelligence agent fails to come under the ambit of criminal laws due to the absence of 'mens rea' (intent) and therefore lack of its concurrence with 'actus rea' (action).

Additionally, in case of civil liability matters, since the artificial intelligence agent is not a 'person' and therefore any/or all vicarious liability laws outlined for a principal-agent relationship in the Indian Contract Act become inapplicable. Other laws related to product liability, prevalent tort laws, etc. are also devoid of clarity as it is yet to be determined if artificial intelligence agents should be considered as a product or if the common law precedents which govern tort laws hold for more complex or unforeseen scenarios concerning artificial intelligence agents.

AI agents, lacking legal personhood, are not typically held directly liable. Instead, liability is assigned to the entities behind them—developers, manufacturers, or users—under existing frameworks like tort law (negligence, product liability) or statutory regulations. Recent cases reflect a growing willingness to explore how traditional liability principles apply to AI systems, especially in areas like copyright infringement, privacy violations, and physical harm caused by autonomous systems.

Literature Review

Literature Review: Theoretical Framework

To understand comprehensively the scope of liabilities, a mutually exclusive and collective exhaustive list of types of Legal Liabilities[1] May be considered for evaluating the artificial intelligence agent's role in the case of each scenario:

- Contractual Liabilities

[1]AI, G. (2025). *can you give me mutually exclusive collectively exhaustive list of legal liabilities?* Retrieved from Grok: grok.com

Obligations arise from agreements between parties, where one party fails to fulfil a promise or breaches a contract (e.g., non-payment, non-delivery of goods or services).

• Tort Liabilities

Obligations arise from civil wrongs (not based on contracts), where one party causes harm or loss to another. These include:

Intentional Torts: Liabilities from deliberate acts (e.g., assault, defamation, fraud).

Negligence: Liabilities from failure to exercise reasonable care (e.g., car accidents, medical malpractice).

Strict Liability: Liabilities imposed without fault, typically for inherently dangerous activities or defective products (e.g., manufacturing flaws, hazardous materials).

• Statutory Liabilities

Obligations imposed by specific laws or regulations, independent of contracts or torts (e.g., tax liabilities, fines for regulatory violations, environmental cleanup costs).

• Equitable Liabilities

Obligations arise from principles of fairness enforced by courts, such as restitution or unjust enrichment (e.g., returning profits gained unfairly at another's expense).

• Criminal Liabilities

Obligations to the state or society for violating criminal laws result in penalties like fines or imprisonment (note: while primarily punitive, these can include restitution to victims).

Literature Review: **Critical Synthesis**

Relevant decided case laws.[2] Which are related to the research problem:

1. Thomson Reuters Enterprise Centre GmbH v. ROSS Intelligence Inc. (USA, filed 2020,

[2]AI, G. (2025). *Can you share latest relevant decided case laws which have placed liability on artificial intelligence agents?* Retrieved from Grok: grok.com

Decided in Part 2023-2024)

- Court: U.S. District Court for the District of Delaware
- Date: Key rulings in 2022 and 2023; summary judgment on certain claims in April 2022, with a significant opinion in 2023.
- Context: Thomson Reuters sued ROSS Intelligence, alleging that ROSS used its Westlaw database to train an AI-powered legal research tool, infringing copyright.
- Ruling: In 2023, the court ruled that ROSS's use of Westlaw data did not constitute "fair use" for training its AI. The court found that copying protected material for commercial purposes outweighed any transformative argument. While the case did not hold the AI itself liable, it imposed liability on ROSS Intelligence (the company) for how its AI was developed.
- Implication: This case highlights that companies can be held liable for the actions of their AI systems, particularly when training data involves copyrighted material. It is one of the earliest substantive decisions in AI litigation and sets a precedent for scrutinizing AI development processes.

2. The Elaine Herzberg Case (USA), 2020

Elaine Herzberg's death is a landmark case in autonomous vehicle liability.

- Incident: On March 18, 2018, Elaine Herzberg, a 49-year-old pedestrian, was struck and killed by an Uber self-driving Volvo XC90 in Tempe, Arizona, USA while pushing a bicycle across a four-lane road. The vehicle was in autonomous mode, with a human safety driver, Rafaela Vasquez, behind the wheel.
- Details: The National Transportation Safety Board (NTSB) found that the Uber system detected Herzberg 5.6 seconds before impact but failed to classify her correctly as a pedestrian due to its design not accounting for jaywalking scenarios. The safety driver was distracted, streaming a TV show, and looked up only 0.5 seconds before the crash, too late to intervene.

- Legal Outcome:
    - Uber's Liability: On March 4, 2019, the Yavapai County Attorney determined there was "no basis for criminal liability" against Uber as a corporation, shifting focus to the safety driver. This decision reflected Arizona's industry-friendly stance on autonomous vehicles.
    - Safety Driver's Liability: In September 2020, a Maricopa County grand jury indicted Rafaela Vasquez on one count of negligent homicide. On July 28, 2023, Vasquez pleaded guilty to a reduced charge of endangerment and was sentenced to three years of supervised probation in Maricopa County Superior Court. No jail time was imposed.

Current Laws in India Governing AI Liability[3]:

India relies on existing legal frameworks to address AI-related liability indirectly, as no dedicated AI liability statute exists. These laws focus on the actions of humans or organizations rather than AI itself:

Tort Law (Civil Liability):

Negligence: Under common law principles, liability arises if a duty of care is breached, causing harm. For AI, this could apply to developers or deployers who fail to ensure the system functions safely (e.g., a poorly designed autonomous vehicle causing an accident). Compensation depends on proven damages, with no fixed penalty.

Strict Liability: Limited to specific scenarios (e.g., hazardous activities under the Rylands v. Fletcher rule), but not broadly applicable to AI without legislative expansion. The Supreme Court in Meenu Mehta v. State of Rajasthan (1977) ruled strict liability requires statutory backing, which is absent for AI.

Consumer Protection Act, 2019 (CPA):

Product Liability: Sections 82-87 allow claims against manufacturers, sellers, or service

[3]AI, G. (2025). *What are the current laws governing liability of AI agents?* Retrieved from Grok: grok.com

providers for defective goods or services causing harm. AI systems (e.g., software or hardware like self-driving cars) could be deemed "products," making developers liable for design flaws without proving fault. Remedies include compensation or product replacement, with fines of up to ₹10 lakh for non-compliance (Section 90).

Application: If an AI-powered medical device misdiagnoses, the manufacturer could be liable under CPA, not the AI itself.

Information Technology Act, 2000 (IT Act):

Section 66 (Computer-Related Offenses): Covers unauthorized acts involving computers (e.g., hacking via AI). Punishable by up to 3 years imprisonment or ₹5 lakh fine. Liability falls on the human operator or programmer, not the AI.

Section 79 (Intermediary Liability): Exempts platforms from liability for third-party content if they act as mere conduits, but this does not directly apply to standalone AI agents. Amendments proposed in 2023 to regulate AI-generated content (e.g., deepfakes) remain pending.

Relevance: If an AI chatbot spreads misinformation, the deploying company could face penalties under IT Act provisions.

Indian Penal Code, 1860 (IPC):

Criminal Liability: Sections like 304A (causing death by negligence, up to 2 years imprisonment or fine) or 420 (cheating, up to 7 years imprisonment and fine) could apply to individuals or entities misusing AI. For example, a developer knowingly deploying a faulty AI system causing harm might be prosecuted.

Limitation: Requires mens rea (intent), which AI cannot possess, so liability shifts to humans.

Motor Vehicles Act, 1988 (Amended 2019):

Section 184: Addresses dangerous driving, with fines up to ₹5,000 or 6-months imprisonment. For autonomous vehicles, liability typically falls on the owner or operator, not the AI.

Relevance: In a hypothetical self-driving car accident, the manufacturer or programmer could be liable if the system's failure is proven, but no case has tested this yet.

Contract Law (Indian Contract Act, 1872):

If AI is part of a service agreement (e.g., AI software sold to a client), breaches could lead to liability for damages under Section 73. This targets the contracting party, not the AI.

International Developments Influencing India

While India lacks AI-specific liability laws, global trends provide context and potential models:

European Union – AI Act (Effective August 2026)[4]:

Classifies AI systems by risk (e.g., high-risk systems like autonomous vehicles). Providers of high-risk AI face strict obligations (e.g., transparency, safety)

Proposed AI Liability Directive (2022): Still under review, it aims to ease claims for AI-related harm by presuming causality if providers fail compliance. This could inspire India to adopt similar rules.

United States: No federal AI liability law exists, mirroring India's fragmented approach. Relies on tort law (negligence, product liability) and state regulations (e.g., Colorado AI Act, effective 2026, mandates risk assessments). Cases like Lee v. Tesla, Inc. (2023) emphasize human oversight, a principle Indian courts might adopt.

Singapore – Model AI Governance Framework (2019, Updated 2024)[5]:

Non-binding guidelines suggest human accountability for AI decisions. India's NITI Aayog has referenced this model, hinting at future alignment.

- NITI Aayog's Principles for Responsible AI (Part 1)[6] Lists seven principles: safety and reliability, inclusivity and non-discrimination, equality, privacy and security, transparency, accountability, and reinforcement of human values. These closely mirror

[4](FLI), F. o. (2025). *The EU Artificial Intelligence Act*. Retrieved from The EU Artificial Intelligence Act: https://artificialintelligenceact.eu/

[5] Authority, I. M. (2024, 10 9). *Singapore proposes framework to foster trusted Generative AI development*. Retrieved from Singapore proposes framework to foster trusted Generative AI development: https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2024/public-consult-model-ai-governance-framework-genai

[6]Aayog, N. (2021, 8). *RESPONSIBLE AI #AIForAll Approach Document for India: Part 2 - Operationalizing Principles for Responsible AI.* Retrieved from RESPONSIBLE AI #AIForAll

Singapore's focus on explainability, transparency, fairness, and human-centricity from the 2019 framework.

- Singapore's influence is implicit in NITI Aayog's adoption[7] Of a balanced, innovation-friendly yet ethical approach, a hallmark of Singapore's model. Both frameworks prioritize trust, accountability, and innovation without heavy-handed regulation—Singapore through its nine dimensions (2024) and NITI Aayog via its seven principles (2021).

Literature Review: **Identification of Gaps**

No AI Personhood: Indian law does not recognize AI as a legal entity, so liability always traces back to a human or corporate actor.

Legislative Gap: Proposals like NITI Aayog's "Responsible AI" framework (2021) suggest ethical guidelines, but these lack legal enforceability. A draft Digital India Act (DIA) floated in 2023 aims to regulate AI, including liability for high-risk systems, but it is not yet passed as a law.

Comprehensiveness Gap: There is no framework or legislation yet which proposes an exhaustive guideline or straightforward civil laws on applicability in different scenarios in which legal liabilities may occur making the question of AI agent's liability subject to the interpretation of presiding judge or authority on case-to-case basis.

Methodology

Methodology: Research Design

Doctrinal approach (non-empirical, focused on legal texts)

Methodology: Sources

- Primary: statutes, regulations, and case laws

[7] Aayog, N. (2022, 11). *RESPONSIBLE AI #AIForAll Adopting the Framework: A Use Case Approach on Facial Recognition Technology.* Retrieved from RESPONSIBLE AI #AIForAll Adopting the Framework: A Use Case Approach on Facial Recognition Technology: https://www.niti.gov.in/

Methodology: Analytical Approach

Analysis of legal rules and doctrines using statutory interpretations and case law synthesis to interpret the law to apply to different scenarios of liabilities.

Methodology: Limitations

- Jurisdictional focus may be required to contextualise the findings of this paper
- Temporal scope of the considerations due to rapidly evolving artificial intelligence capabilities

Analysis

Analysis: Historical Development

Trace the doctrine's evolution through key legislative/judicial milestones.

- Legal Maxim, maybe traced back to the common law case of Ashby v. White (1703), "where there is a wrong, there is a remedy", ubi jus ibi remedium, meaning that for every legal wrong, the law provides a remedy or means of redress.
- Doctrine of Negligence (Nonfeasance, Misfeasance, Malfeasance), In Blyth v. Birmingham Waterworks Co., was defined as Negligence is the failure to do something a person of ordinary prudence would do or the taking of an action that a person of ordinary prudence would not take. Establishing a consideration of necessary 'standard of care' subject to reasonability.
- Doctrine of "Res Ipsa loquitur," meaning "the thing speaks for itself" allows plaintiffs to infer negligence from the nature of an accident or injury when direct evidence of negligence is lacking if certain conditions are met.
- "No fault" liability: both strict and absolute liability operate under the "no-fault" principle, meaning the defendant is liable regardless of their specific actions or negligence.
- Elements of Strict Liability as laid down in Rylands vs Fletcher (1868): For strict

liability to apply, certain elements must be present:

- A person must have brought something dangerous onto their land.
- That dangerous thing must escape and cause damage.
- The use of land must be non-natural

- Absolute Liability: As per Mehta v. Union of India, 1987 SCR (1) 819; AIR 1987 965, Absolute Liability shall apply to only those enterprises shall be held liable which are involved in 'hazardous or inherently dangerous activities', this implies that other industries not falling in this ambit shall be covered under the rule of Strict liability.

- Exceptions to No Fault Liability:
  - Volenti non-fit injuria (where the person affected consented to the risk).
  - Act of God
  - Act of the third party
  - Contributory negligence or fault of the plaintiff
  - Statutory Authority

Analysis: Statutory Interpretation

Examining related statutory provisions,

- The Indian Contract Act, of 1872, defines the agent-principal relationship in Sections 182-238, with Section 182 defining "agent" and "principal" as a person employed to act for another or represent them, and the person for whom the act is done, respectively, therefore it absolves the consideration of AI Agents as there is no established personhood.

- The Information Technology Act, of 2000 has provisions which concern certain aspects of Artificial Intelligence but do not specifically talk about Agents' liability:

  - Data Privacy and Protection: Sections 43A, 72 and 72A
  - Cybercrimes and Offenses: Section 43, Section 65, Section 66D, Section 67, Section 66F
  - Intermediary Liability read with The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021: Section 79

- Consumer Protection Act, 2019: Applies to AI embedded in goods or services (e.g., defective AI-powered devices like smart appliances).
  - Liability: Manufacturers, sellers, or service providers can be held liable for defective products or services causing harm to consumers. This includes:
    - Product Liability (Section 82-87): Strict liability applies if an AI product is defective, with compensation for harm caused (e.g., injury, property damage)

- Although there is no legal backing yet through legislation, as per NITI Aayog's AI Framework, Part 2 - Operationalizing Principles for Responsible AI (August 2021), the following mechanisms are prescribed for setting accountability:
  - Risk-Based Approach to Accountability
  - Stakeholder Roles and Responsibilities
  - Audit and Reporting Mechanisms
  - Grievance Redressal and Liability (Strict liability is not necessarily prescribed)
  - Regulatory and Policy Interventions
  - Public Accountability and Transparency

- Digital Personal Data Protection Act, 2023 (DPDP Act): this law governs data processing and could impose obligations on AI agents handling personal data, though it does not specifically address AI autonomy or liability.

Analysis: Case Law Analysis

There are almost no decided precedents as of 24 March 2025 in India related to AI Agent Liability. There is one major ongoing case:

- ANI v. OpenAI (Ongoing as of March 2025)
    - Context: This case, currently before the Delhi High Court, marks India's first significant AI-related copyright dispute. The news agency ANI has accused OpenAI, the creator of ChatGPT, of using its copyrighted content to train AI models without permission, seeking INR 2 crore in damages. The hearing is ongoing.
    - Relevance to AI: This case raises critical questions about the liability of AI agents (or their developers) for copyright infringement when training data includes protected material. It could set a precedent for how AI-generated content and the use of third-party data are regulated in India.

In a closely related case decided outside India therefore not a legal precedent,

- In Thomson Reuters Enterprise Centre GmbH v. ROSS Intelligence Inc. (United States District Court for the District of Delaware, filed 2020, Decided in Part 2023-2024),
    - The court ruled that ROSS's use of Westlaw data did not constitute "fair use" for training its AI. The court found that copying protected material for commercial purposes outweighed any transformative argument. While the case did not hold the AI itself liable, it imposed liability on ROSS Intelligence (the company) for how its AI was developed

In another decided case in the USA in a case of criminal liability of an AI agent,

- The Elaine Herzberg Case (USA, 2020) was the first criminal case against a human operator of a fully autonomous vehicle in a fatal incident. It set a precedent that safety drivers bear significant responsibility, even in highly automated systems, rather than the AI or company being directly liable.

Analysis: Findings and Conclusion

Critically evaluating the findings about the theoretical framework of a mutually exclusive collectively exhaustive list of liabilities to answer the research questions:

Kind of Liability	Who is Liable?	Qualifying Conditions & Contributing Factors	Rationale
Contractual	Defaulting party to contract	Privity to contract	Necessary human accountability, Contractual obligations
Tort: Intentional	Human actor/s with intent (Can be user, intermediary, developer, all involved)	Intent of the human actor/s, Foreseeability of harm within the duty of care	Necessary human accountability, If it is determined that a duty of care exists and if an AI agent's actions caused foreseeable harm to another, then.
Tort: Negligence	Negligent human actor/s (Can be user, intermediary, developer, all involved)	Reasonable standard of care not observed, proximity to cause as a determining factor, Subject to contractual waive off which legally prescribes an alternate human actor as liable	Necessary human accountability, Breach of Duty to reasonable care, Law of Proximity
Tort: Strict	AI Agent system user	Subject to Statutory recognition of High-Risk/Strict/Absolute Liability AI agent systems, proximity to cause as a determining factor, Subject to contractual waive off which legally prescribes a human actor as liable	Necessary human accountability, Law of Proximity
Statutory	State	Statutory action necessary as per statute, Damages to be compensated to the injured as prescribed in the statute	Necessary human accountability, Exception of statutory action to no fault liability
Equitable	Subject to the nature of the Tort (as determined above)	Conditions met as per Specific Relief Act 1963, Indian Trusts Act 1882	Necessary human accountability, Principles of equity and natural justice
Criminal	Human actor/s with Mens Rea (intent) (Can be user, intermediary,	Concurrence of Mens Rea (intent) with ultimate Actus Reus (action) or omission by AI agent, Crime & punishment proportionate to identification as	Necessary human accountability, Constructive Liability through common intention,

	developer, all involved)	per Bhartiya Nyaya Sanhita, 2023 to all human actors found guilty	Legality and Certainty of Offense

Recommendations

- There cannot be a 'one size fits all' solution to placing liability for AI agent actions but clarity may be established using jurisprudential methods to determine accountability and liability for an exhaustive list of scenarios for not only fair and swift disposable of conflicts but to provide a preventive framework to the society to avoid unanticipated losses due to fast-evolving AI capabilities which may be overridden by judicial precedents and not vice-versa

- For Contractual Liabilities, the Indian Contract Act, of 1872 may be reviewed for any amendments required

- Necessary statutory regulations need to be defined to fix strict liability as necessary

- Exceptions to liabilities arising out of statutory actions need to have provisions for a remedy to be provided by the state as per the statute to the victims of AI agent actions

- For Equitable Liability, Specific Relief Act, of 1963 and Indian Trusts Act, of 1882 to be reviewed for any amendments required

- For Criminal liability, Bhartiya Nyaya Sanhita, of 2023 may be reviewed for any amendments required

References

(FLI), F. o. (2025). The EU Artificial Intelligence Act. Retrieved from The EU Artificial Intelligence Act: https://artificialintelligenceact.eu/

Aayog, N. (2021, 8). RESPONSIBLE AI #AIForAll Approach Document for India: Part 2 - Operationalizing Principles for Responsible AI. Retrieved from RESPONSIBLE AI #AIForAll Approach Document for India: Part 2 - Operationalizing Principles for Responsible AI: niti.gov.in

Aayog, N. (2022, 11). RESPONSIBLE AI #AIForAll Adopting the Framework: A Use Case Approach on Facial Recognition Technology. Retrieved from RESPONSIBLE AI #AIForAll Adopting the Framework: A Use Case Approach on Facial Recognition Technology: https://www.niti.gov.in/

AI, G. (2025). Can you give me a mutually exclusive collectively exhaustive list of legal liabilities? Retrieved from Grok: grok.com

AI, G. (2025). Can you share the latest relevant decided case laws which have placed liability on artificial intelligence agents? Retrieved from Grok: grok.com

AI, G. (2025). What are the current laws governing the liability of AI agents? Retrieved from Grok: grok.com

Authority, I. M. (2024, 10 9). Singapore proposes a framework to foster trusted Generative AI development. Retrieved from Singapore proposes a framework to foster trusted Generative AI development: https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2024/public-consult-model-ai-governance-framework-genai