
AI HALLUCINATION AS A SOURCE OF CRIMINAL HARM

Noor Fatima Khan, Government New Law College, Indore

ABSTRACT

Artificial Intelligence (AI) has become an integral part of making decision across various sectors such as finance, healthcare, law etc. However, one of the most challenging aspects is AI hallucination. AI hallucination results into the rise of criminal harm including fraud, defamation, misinformation, wrongful accusation, cyber crime etc. AI-generated content increasingly influences public opinion, because of which the legal consequences of AI hallucination have become a matter of growing concern. This article argues that AI hallucination represents a major source of criminal harm that requires a specific legal accountability mechanism. It also recommends some legal reforms aiming to ensure accountability and promote the responsible development of Artificial intelligence technology.

Keywords: Artificial Intelligence, AI hallucination, cyber crime, criminal harm, technology, accountability.

Introduction

Artificial intelligence has become a popular as well as important technology in modern society. Several industries are using artificial intelligence to simplify their tasks and seek assistance from it whenever required. Artificial intelligence has become an integral part of the legal industry as well. Advocates, students, drafters, researchers and other legal professionals are taking assistance from AI for research, translating document, drafting, contract review, and various such tasks.

AI is very useful technological tool, but it also consists of some drawbacks, of which is AI hallucination. AI hallucination refers to the instances where an artificial intelligence system generates information or responses that are incorrect or entirely fabricated but presented as factual.¹ AI hallucination is not a technical defect, it can create a significant harm in the realm of criminal law by generating fabricated and false information. Recent incident have shown that AI generated outputs containing fabricated information is mistakenly treated accurate thereby creating the potential for legal and criminal consequences.² Main legal challenge is who should be held responsible when criminal harm results from misinformation generated by AI. Traditional criminal laws are based on men's Rea and actor serious why AI system does not possess any of the above characteristics in legal sense.

Understanding AI Hallucination

In the field of AI, hallucination means instances where AI system produces false, inaccurate, fabricated and deceptive output, conveying them as reliable and true information.

Unlike ordinary errors done by software, hallucination often appear authoritative because of which users rely on false information without recognising its accuracy. AI platforms are trained on various data set to perform tasks. if these data sets are limited or incomplete, system may generate inaccurate outputs. these systems are not aware of the facts, they operate on statistical realm. since these models generate output on their training set flaws, gaps or biases in the data may propagate as hallucination.

¹ Stanford HAI

² Mata v. Avianca ,Inc.

Reasons for AI hallucination:

- Insufficient training of data.
- Prediction made by AI system on the basis of sequences.
- Ambiguous prompts provided by the user.
- Training AI on outdated or old information.
- limited understanding of context by AI during conversations.

Criminal harm arising from AI hallucination

The most pressing challenge in modern criminal law is the harm arising from AI hallucination which affect individuals, institutions and administration of justice.

Hallucination done by AI can contribute to following harms:

- **Defamation:-** AI hallucination can result into defamation when false information or allegation is generated by AI about an individual. If these information get published without review, it seriously damage the individuals reputation and give rise to legal liability for the act.

Example: If a person asks AI about a local businessman and the AI falsely says that the person is involved in fraudulent act, the information if shared online will cause harm in the reputation of businessman amounting to defamation.

- **Financial crimes:-** Fabricated financial advices and banking information generated by AI can be used to deceive consumers into fraud, resulting in financial loss. Criminals may intentionally generate false information from AI to commit fraud.

Example: A fraudster get false banking information from artificial intelligence Chatbot. He uses it to deceive people which results in financial loss.

- **Fabrication of digital evidence:-** AI may generate fake emails, documents, conversations or images that appear real. Sometimes it also generate fake case laws and judgments which in real does not exist these materials influences the legal proceedings.

Example: An investigator uses AI for summarising chat record and AI added some additional fake messages which was never sent. This mistakenly treated as evidence.

- **Wrongful accusation:** AI, by generating false evidence may identify innocent as an offender, which leads to the damage of reputation and violation of fundamental right of an innocent individual.

Example: The police asks AI to identify a suspect. The AI wrongfully took the name of an innocent person and he have to suffer from the Consequences before discovering the mistake.

- **Cyber crime:** AI hallucination may produce innocent programming code or inaccurate cyber security guidance. Users depend on these outputs without verifying it, may I have to suffer from hacking of their personal information. AI does not intentionally commit cyber crime, hence it's important to check AI generated content and does not rely on these information blindly.

Example: A criminal ask AI chatbot to write an email for a bank. The AI creates a convincing message, which the criminal uses to trick people for sharing their banking details.

Legal issues in assigning liability

- **Absence of legal personality:-** Artificial intelligence system are not considered as a legal entity because they does not possess those qualities which are normal human being have. These systems work on command of human individuals hence AI system cannot be punished under criminal law principles.
- **Lack of Mensrea:-** Men's rea is a mental element which builds the liability against wrongful act . AI system do not intentionally generate fabricated information. Therefore, it becomes a bit difficult for assigning the liability for criminal act .
- **Causation problem:-** Establishing a casual link between AI generated content and criminal harm is might be difficult.
- **Attribution problem:-** Determining the responsibility among developers deployed

users or organisations remains a significant legal challenge during assigning the liability.

Comparative legal analysis

European Union:

The European Union Has adopted a risk based Approach to AI regulations. AI system with high risk are subject to strict transparency and compliance requirements. They have introduced the most comprehensive regulatory framework which is European Union Artificial Intelligence Act.

United States:

The United States is currently relying on existing tort and sector specific regulations against AI hallucination. In cases where any criminal harm caused by AI hallucination, court decide liability on the basis of human actions, as AI lacks legal identification.

United Kingdom:

The United Kingdom favours the framework which focuses on safety, accountability and fairness . They have not adopted any AI specific statutes , instead they adopted some principles which are compulsory to be followed by organization and institutions.

India:

At this point of time India doesn't have any AI specific legislation. For most of the cases related to AI hallucination, Information technology act, 2000 ; Bhartiya Nyay Sanhita, 2023 and Digital personal data protection Act, 2023 is referred during legal proceedings for assigning liability, but still gap exists.

China:

China has adopted preventive and government led approach. They did not waited for any harm to be occurred by AI hallucination, instead they choose preventive measures for their country by introducing *Interim measures for the management of generative AI service*, which came into effect in 2023. Developers may have to face regulatory penalties if any harm is caused to public

from AI hallucination.

Need for legal accountability:

At present, India does not witness a reported criminal conviction where AI hallucination itself has been recognized as the direct reason for a particular criminal act, but some cases like *Gummadi Usha Rani V. Sure Mallikarjuna Rao*³, where trial court relied upon AI generated non-existent case citations, and *Deepak Rakheja V. Omkara Assets Reconstruction Pvt. Ltd*⁴, where supreme court was informed that pleadings contained more than 100 AI generated, non-existent case citations, has affected the legal proceedings.

This highlights the need for a legal accountability framework before a hallucination results in more serious consequences. It is the responsibility of AI developers to test their system, reduce hallucination and clearly warn users about the inaccurate output. Hence, developer is to be liable for foreseeable risk.

Users should also verify the data before publishing or relying on it and if users knowingly do the act then they will be responsible for intentional misuse. Furthermore, mandatory transparency obligations must also be adopted by the government for security purposes.

Therefore, a framework with effective legal accountability should combine responsibility of developers, user liability, independent auditing and transparency. This kind of framework will protect individuals from harm caused by AI hallucination and ensure the continuation of innovation in a responsible and legally accountable manner.

Recommendations:

- **Enactment Of a Dedicated AI Accountability Law:** India should have considered to bring AI specific laws, which lay down the idea for governing the criminal or civil offences committed by the AI generated content. These laws must define basics of AI and its careful usage. It should also prescribe the liability where AI generated content brings substantial harm.

³ *Gummadi Usha Rani V. Sure Mallikarjuna Rao* (Andhra Pradesh high court Jan 21, 2026)

⁴ *Deepak Rakheja V. Omkara Assets Reconstruction Pvt. Ltd*

- **Clearly Define Criminal Liability:** The law should clearly define the liability among developers, service providers, organization deploying AI system and users for international misuse of AI generated content.
- **Mandatory Human Oversight:** It is very important to review the data generated by artificial intelligence system. Human oversight is necessary safeguard which can help to prevent wrongful decision making ,Based only on hallucinated AI output.
- **Transparency:** AI developers are required to disclose important information regarding the limitation and functioning of their system. Greater transparency enables users to verify information and reduces blind dependency on responses generated by AI.
- **Establishment of an independent AI regulatory authority:** India should establish a specialised AI regulatory authority. This authority will look after the licensing of high risk AI system, monitoring compliance, and imposing the penalties when required.
- **Provide effective remedies for victims:** Individuals who suffered harm from AI hallucination should also have some effective legal remedies. Victims should be entitled to get compensation for the loss of their reputation, money or having any kind of legal injury.
- **Promote AI literacy and public awareness:** Government organisations, educational institutes and technological companies should conduct awareness programs Where both the benefits and limitations of artificial intelligence should be explained by the experts. Users should also understand that the information generated by AI should not treated accurate and be independently verified before using it for important matters.

Hence, a solid framework should be introduced which not only focuses to punish the wrongdoer but also lay down preventive measures like continues monitoring of AI system.

Conclusion:

Artificial Intelligence has become an important part of our daily lives. From answering simple questions to assisting professionals in complex tasks, its role continues to grow every day. While these developments have made many processes faster and more efficient, they have also

created new legal challenges that cannot be ignored. One such challenge is AI hallucination, where an AI system produces false or inaccurate information with confidence. When such information influences criminal investigations, legal advice, court proceedings, or public decisions, the consequences can be serious. A person may face wrongful allegations, financial loss, damage to reputation, or even legal action because of information that was never true in the first place.

This is why there is a need for a legal framework that clearly defines the duties of AI developers, companies that use these systems, and individuals who rely on AI-generated outputs. The law should encourage innovation, but it should also ensure that technological progress does not come at the cost of justice and public safety. Human supervision, transparency in AI systems, regular testing, and proper accountability mechanisms can reduce the risk of criminal harm and improve public trust in AI.

In the end, the issue is not whether AI should be used, but how it should be used responsibly. Technology will continue to evolve, and the law cannot remain silent while new forms of harm emerge. A balanced legal approach can help society enjoy the benefits of AI without compromising the rights and safety of individuals. If AI is expected to play a greater role in the future, then legal accountability must grow alongside it. Only then can innovation and justice move forward together.

References:

1. Criminal liability of artificial intelligence entities
<https://docs.manupatra.in/newsline/articles/Upload/4e5c9c80-320b-4433-9f87-f56059a5345c.pdf>
2. AI Hallucinations in Indian Courts | AI Hallucination Legal Case
<https://advocateprakhhar.com/ai-hallucinations-indian-courts-complete-guide-2024-2026/>
3. AI hallucination as a form of harm
https://msarr.magnascientiapub.com/sites/default/files/fulltext_pdf/MSARR-2026-0051.pdf