A Note on Responsibility and Accountability for Artificial Intelligence in Court Proceedings: Issues and Considerations

DAKSH

May 2026

In the recent past, high courts and the Supreme Court have been repeatedly pointing out hallucinated case titles and details in arguments and pleadings arising out of lawyers using artificial intelligence (AI). The Supreme Court asked the Bar Council of India to constitute a committee of experts, including field experts, to examine issues arising from the use of artificial intelligence in court proceedings, in a case where a trial court relied on non-existent judgments cited by a litigant. The Supreme Court observed that reliance on non-existent, AI-generated judgments is not a mere error in decision-making but would amount to misconduct and that legal consequences would follow. [1]

This note seeks to describe the technical nature of AI, some ways that AI may cause issues, and how its use by advocates may be regulated.

At the outset, it is important to understand what AI is and how it works. The UNESCO Guidelines for the Use of AI Systems in Courts and Tribunals define AI systems: "Computational systems that can process data and information in a way that resembles intelligent behaviour, and typically includes aspects of reasoning, learning, perception, prediction, planning or control." [2] At the core of AI as we are familiar with today are large language models (LLMs) that are statistical systems trained to predict the most likely word fragments given everything that came before it.

LLMs operate using what is known as a transformer architecture with self-attention mechanisms that capture long-range dependencies and contextual relationships between words. When a person asks an LLM what colour an apple is, it does not only link to an authoritative document (like a Google search would). It also generates an answer choosing red as the most probable word associated with apple. LLMs are trained on vast datasets to capture patterns in text, thus enabling a range of language tasks.[3] LLMs are at their core probabilistic tools and are not capable

[1] Gummadi Usha Rani vs. Sure Mallikarjuna Rao SLP (C) No. 7575/2026 Order dated 27-02-2026.

[2] UNESCO and Javier Gutiérrez. Guidelines for the Use of AI Systems in Courts and Tribunals. Paris: UNESCO, 2025. https://unesdoc.unesco.org/ark:/48223/pf0000396582.

[3] Ferraris, Andrea Filippo, Davide Audrito, Luigi Di Caro, and Cristina Poncibò. "The Architecture of Language: Understanding the Mechanics behind LLMs." Cambridge Forum on AI: Law and Governance 1 (2025): e11, 1–19. https://doi.org/10.1017/cfl.2024.16.

of reasoning like a human mind. More importantly, in the Indian judicial context, especially in district courts and high courts, case-level data such as pleadings and order sheets are still paper-based and therefore not reliably available for training LLMs. Even for legislation, there is still no single, authoritative database that consolidates all sources of law, from primary legislation to delegated legislation and subsequent amendments and repeal, in a consistent, machine-readable format.[4] These gaps in India's legal data ecosystem mean that many AI tools operating in this context are trained largely on foreign legal data. Law, unlike the sciences, is jurisdiction-specific and hence the probability of the outputs of inadequately- trained AI being inaccurate is significant.

The other factor that may make AI outputs unreliable is hallucinations. Hallucinations are generally described as generated content that sounds plausible or well-drafted at first glance but is actually fabricated, misrepresented, or incorrectly described. Hallucinations are "a feature of LLM design, not a bug" and while technical attempts are underway to identify or reduce such hallucinations, models do not always recognise when they have made an erroneous connection.[5] Researchers at Stanford in 2024 attempted to create a typology of legal hallucinations: (i) when generated text is inconsistent with the given prompt, (ii) when it extends information it was trained on to apply it to a new set of circumstances, and (iii) when it misstates or misrepresents facts. Any of these three errors may give rise to information presented in court that an advocate may not be able to defend if questioned.[6]

Table 1: Typology of legal hallucinations

Domain	Type of hallucination	Legal example
Closed	response inconsistency with the prompt	Mischaracterization of an opinion
Open	response inconsistency with the training corpus	Creative argumentation
	response inconsistency with the facts of the world	Misstatement of the law

It is against this backdrop that the use of AI, specifically generative AI, by advocates raises important questions about legal ethics and the duties of advocates towards their clients and courts. Rules 15 and 24 of Chapter II of the Bar Council of India Rules lay down a duty to

[4] DAKSH & Digital Futures Lab. (2026). *AI for justice: Ethical, fair and robust adoption in India's courts.* https://www.dakshindia.org/ai-for-justice-ethical-fair-ofand-robust-adoption-in-indias-courts/

[5] Roepstorff, Andreas. "AI Hallucinations Are a Feature of LLM Design, Not a Bug." *Nature* 639, no. 8053 (March 4, 2025): 38. https://doi.org/10.1038/d41586-025-00662-7.

[6] Dahl, Matthew, Varun Magesh, Mirac Suzgun, and Daniel E. Ho. "Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models." *Journal of Legal Analysis* 16, no. 1 (2024): 64–93. https://doi.org/10.1093/jla/laae003.

fearlessly uphold the interests of his client by all fair and honourable means without regard to any unpleasant consequences to himself or any other, and a bar on doing anything whereby he abuses or takes advantage of the confidence reposed in him by his client. These duties are particularly significant in our age, where the risks of over-reliance of AI are well-documented. Given these professional duties, it is important for advocates to use AI tools responsibly and cautiously and to regard AI outputs as a starting point for further inquiry and not a substitute for their own legal research and reasoning.

What could go wrong?

An AI system may:

- cite cases that do not exist
- fabricate excerpts from real cases
- fabricate statutory provisions
- fabricate the ratio decidendi of a case
- reinforce the advocate's initial position or ideas without presenting contradictory views
- make mistakes with calculations and other mathematical functions
- not be internally consistent in a draft and may contradict text generated in prior sentences or paragraphs
- overgeneralise, e.g. making broad assumptions when simplifying or summarising text
- get access to confidential and privileged information

Duties and responsibilities of various stakeholders

Stakeholder	Duties & Responsibilities
Drafting advocate	Primary duty to verify everything cited in the pleadings Do not over-rely on AI tools
Senior	Review AI-assisted work of juniors Establish verification protocols in the office or for any material presented for their signature
Opposing Counsel	Do not assume opposing citations are correct Flag suspected hallucinations promptly

	Independent cross-check cases relied on by the other side
Judge	Verify suspect citations during hearings Impose sanctions on errant lawyers Refer errant lawyers to the Bar Association
AI Product Developers	Disclose where/how AI is used Build in technical guardrails for verification
Parties in person	Should be given a checklist of verification steps May be required to file an undertaking regarding verification

What Can Accountability Look Like?

1. **Professional associations**

The Bar Council and Bar Associations should treat the submission of AI-fabricated citations as professional misconduct. A graded approach to sanctions for such misconduct would be advisable:

- First instance: formal censure, mandatory remedial training and costs
- Repeat offenders: escalating sanctions, including suspension from practice

2. **Court-imposed sanctions**

Courts may take action against advocates submitting inaccurate information generated through AI in the following ways:

- Censure
- Imposition of costs on the advocate (non-transferable to the client)
- Striking out of pleadings or cases with inaccurate AI-generated information

3. **Mandatory training**

The Bar Council should mandate AI literacy training as part of the curriculum in law schools. Such professional bodies should also arrange for online modules and videos on AI literacy for practising advocates.

4. **Disclosure norms**

Mandatory disclosure regarding the usage of AI creates a paper trail for accountability and alerts opposing counsel and courts to material that requires heightened scrutiny. Advocates must disclose to their clients and the court when they use AI to draft pleadings and written submissions.

5. **Chamber-level protocols**

The advocate who signs the pleadings must take full responsibility for their contents. It is thus necessary for advocates' chambers and law firms to establish protocols at the level of their chamber or firm that may include:

- Clear policies on permitted and prohibited uses of AI tools
- A list of approved AI tools selected by evaluating tools for security, appropriateness of training, accuracy and suitability for the Indian legal context.
- Guidelines on not sharing confidential and privileged information with AI tools
- Mandatory human verification of all AI-generated content in pleadings and written submissions
- Specific procedures for reviewing pleadings received from opposing parties
- Appropriate disclosures to clients

6. **Designation of authoritative legal sources**

Courts should designate authentic, official sources for case law against which all citations must be verified. These may include the official website of the High Courts and Supreme Court websites and any other designated commercial or third-party databases.

Parties-in-person

Parties who represent themselves in court, rather than being represented by advocates, should be treated differently from advocates. They should be granted leniency and not be held to the same standards as trained advocates. However, a calibrated approach including some of the following steps may be advisable to tackle their use of AI:

- Provision of a checklist in simple language or verification steps to be completed before filing
- An affidavit acknowledging that citations have been independently verified

AI Product Developers

Developers of products using AI should be obliged to disclose the datasets on which their models were trained so that advocates can decide whether their tools are reliable. Where an AI tool is specifically designed for use by advocates, developers should build in guardrails such as providing authoritative sources for all outputs.

Examples of Advisory Notes or Guidelines developed for advocates in other jurisdictions:

1. American Bar Association, Standing Committee of Ethics and Professional Responsibility. (2024). Formal Opinion 512: Generative Intelligence Tools
2. Bar Council of England and Wales, The Information Technology Panel. Considerations when using ChatGPT and generative artificial intelligence software based on large language models. (2024), updated in 2025)
3. Canadian Bar Association, Ethics of Artificial Intelligence for the Legal Practitioner. (2024) https://cba.org/resources/practice-tools/ethics-of-artificial-intelligence-for-the-legal-practitioner/
4. Ministry of Law, Singapore. (2026) Guide for Using Generative AI in the Legal Sector https://www.mlaw.gov.sg/files/Guide_for_using_Generative_AI_in_the_Legal_Sector_Published_on_6_Mar_2026_.pdf