The Indian Journal for Research in Law and Management

Open Access Law Journal – Copyright © 2026

Editor-in-Chief – Dr. Muktai Deb Chavan; Publisher – Alden Vas; ISSN: 2583-9896

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International (CC-BY-NC-SA 4.0) License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

WHEN AI GETS THE LAW WRONG: HUMAN OVERSIGHT AND ACCOUNTABILITY IN AI-ASSISTED LEGAL RESEARCH

~ Gargi Kolhe

INTRODUCTION

Generative artificial intelligence (AI) has quickly entered into legal work. It can find authorities, summarise judgements, identify arguments and can prepare drafts in no time. However, the same systems can produce fabricated cases, inaccurate citations and legal provisions that may sound binding but are wrong. In regular research, an incorrect response may waste time but in law, an incorrect authority can alter an argument, affect adjudication or the rights and liabilities of parties.

Therefore, the primary difficulty is not merely that AI can hallucinate, but rather it is what happens after the hallucination is generated. If a lawyer, judge or any other legal practitioner treats AI-generated material as a legal authority, the error shifts from machine to legal process that turns AI assistance into AI authority.

The present article argues that human oversight in AI-assisted legal research should be understood as a significant and traceable process of legal verification, proportional to the consequences of reliance. AI may assist in legal reasoning, but the responsibility of deciding what the law actually is cannot be given to the system. Legal authority infers decisions not merely by convincing language but from the source, hierarchy, jurisdiction, context and continuity. Maintaining this is important when AI becomes part of legal research.

FROM HALLUCINATION TO LEGAL ERROR

AI hallucination is not only limited to inventing a case name, it extends beyond that. The National Center for State Courts identifies the risks associated with legal hallucinations such

as generating non-existent authority, providing unsupported legal provisions, displaying false information about court procedures or blending different legal concepts from different contexts.[1]

The particular issue lies in the appearance of credibility. A fabricated citation may contain all necessary components including party names, pinpoint paragraph and formatting, resembling a valid judgement. Even if it identifies a real judgement it is not guaranteed the relevant information supports the provision.

Therefore, legal AI error should be considered as a chain of possible failures rather than a single event called hallucination. The key question is where the verification went wrong and what is the outcome. The risk is connected with fluency. Generative systems provide answers in a well-organised manner but that does not guarantee accuracy, it can conceal or omit an important source. Therefore, fluency increases the need for verification.

FROM AI ASSISTANCE TO AI AUTHORITY

The problem is not using AI as a research assistant. As described by Danaher AI research tools can be understood as "cognitive arti-facts" that assists in performing cognitive tasks. Any competent person may use it to simplify and develop further research.[2] The actual difficulty begins when the AI-generated material becomes a final submission. This process can be subtle. The person can enter a prompt in AI to act as a professional researcher. The system generates a case name, the person assumes it to be correct and adds it to the draft, it is later filed. The opposite party or court presumes that it is verified. At this point AI starts to act as if it were a reliable legal authority.

AI can help in locating, organising or finding any material, but the generated answer does not become a precedent merely because it sounds sophisticated. The system still needs an independent human decision-maker to cross-verify whether the submitted authority truly exists and is relevant.

Verification should not discard the beneficial features that make AI useful. It should rather prevent conversion of convenience into authority. AI can make the route easy by finding a source but human verification remains necessary to read, understand and check the sources

[1] TRI/NCSC AI Policy Consortium for Law & Courts, *A Legal Practitioner's Guide to AI & Hallucinations*, NCSC (Feb. 16, 2026), https://www.ncsc.org/resources-courts/legal-practitioners-guide-ai-hallucinations.
[2] Daniele Mezzadri, *The Paradox of Ethical AI-Assisted Research*, 23 J. ACAD. ETHICS 2653, 2654 (2025).

before relying on it. After all, clients hire lawyers not just for answers, but for wisdom, judgment, and advocacy, while AI can help find the information, the lawyer still provides the counsel.[3]

WHAT DOES LEGAL VERIFICATION ACTUALLY MEAN?

The phrase "verify the AI Output" sounds like a simple instruction until one has to find out what verification actually requires. Checking whether a citation or judgement exists is necessary but it alone is not sufficient. Legal verification involves several steps and each can contain different kinds of errors. The article proposes a four-level framework.

Level 1: Existence Error

Does the cited authority actually exist?

An AI system may generate fabricated citations, cases, statutes or quotations. The first step of verification begins by finding the primary sources of these authorities. If the given authority does not exist, every idea built upon it fails. This is the most basic level, yet in some Indian cases this has also been overlooked.

Level 2: Attribution Error

Does the authority actually contain what AI has attributed to it?

The authority provided by AI can be real but still it can be represented inaccurately. The provided quotation, paragraph number or a party's submission may be incorrect. Verification not only involves confirming the original source but also identifying who said it and what was said. Here the error is not about incorrect invention of any authority but misrepresentation of the content of the authority.

Level 3: Interpretation Error

Does the authority support the idea for which it was being used?

A correct authority can still become misleading if it is taken out of context. For example, AI may quote a statement from judgement and present it as a legal rule, but in reality the statement was only a part of a party's argument or applied only to a specific proposition. Therefore, verification requires what, why and how far the Court decided something. Simply finding similar looking words is not enough. At this level verification moves out from mere fact checking to legal reasoning.

---

[3] Maryam Danishwar, *AI and the Legal Profession: Friend, Foe, or Future Colleague?*, CLARK HILL (Sept. 17, 2026), https://www.clarkhill.com/news-events/news/ai-in-legal-practice-benefits-risks-ethics/.

Level 4: Legal-Status Error

Legal authorities change over time subject to limited exceptions. A judgement or any statute can be overruled, modified or replaced by later decisions. AI can accurately provide the old framework that is no longer usable instead of providing contemporary information. The process of verification must involve looking at the continuing relevance of the authority.

The four levels are cumulative: Attribution depends on existence, interpretation depends on accurate attribution and legal-status depends on identifying the legal proposition correctly. This reveals a continuous progression from finding sources to authenticity of the content to understanding its meaning and finally to finding its continuing legal force. AI may shorten the route to find a primary source but it cannot ease the human duty to verify the authority before relying on it.

HUMAN OVERSIGHT IN PRACTICE: WHAT INDIAN CASES REVEAL

Now the problem is no longer hypothetical and can be seen in practice.

Pooja Ramesh Singh v. Jammu and Kashmir Bank Ltd. & Anr. (2026)[4]: The Supreme Court quashed the NCLT and NCLAT orders which were based on six AI-generated authorities containing non-existent citations or paragraphs. The Court adopted a "zero-tolerance" policy towards unverified AI-generated precedents and insisted on the need for "a human in the loop at every stage" of the procedure. This demonstrates how an existence or attribution error can transfer from legal research to the courtroom, compromising the sanctity of judicial decision-making.

Deepak s/o Shivkumar Bahry v. Heart & Soul Entertainment Ltd. (2026)[5]: The Bombay High Court found the written submissions made using AI tools and relied on the non-existent case "Jyoti w/o Dinesh Tulsiani v. Elegant Associates", which the Court and the law clerks could not find. The Court held that AI may assist in research but cannot replace human verification and anyone relying on AI generated material must verify its existence and genuineness. The Court also imposed ₹50,000 penalty for the judicial time spent in finding the fabricated authority. The case imposes professional responsibility fully on humans to verify the AI-generated outcomes.

[4] Pooja Ramesh Singh v. Jammu and Kashmir Bank Ltd. & Anr. 2026 INSC 668.

[5] Deepak s/o Shivkumar Bahry v. Heart & Soul Entertainment Ltd. (2026:BHC-AS:828).

Greenopolis Welfare Association (GWA) v. Narendar Singh & Ors. (2025)[6]: The petition cited non-existent judgment and incorrect passages that were not present in genuine judgement, including citing paragraph 73 of *Raj Narain v. Indira Nehru Gandhi*, although the judgement contained 27 paragraphs. The petition was subsequently withdrawn. The discrepancies strongly indicated that the material was generated using AI. This illustrates that mere verification of the existence of a judgement is not enough, the specific paragraphs, numbers, relevance must also be checked.

These cases reveal different aspects of the same problem. *Pooja Ramesh Singh* case shows how unverified AI-generated authorities can enter judicial proceedings. On the other hand, the *Deepak* case places responsibility on the person relying on AI, while the *Greenopolis* case demonstrates that verification must extend beyond the existence of any authority. The common point is that AI-generated authority cannot be considered as a legal authority without human verification.

FROM ACCOUNTABILITY TO TRACEABILITY

Merely saying humans remain accountable if relying on AI is necessary but incomplete. Accountability requires traceability. Effective traceability requires records such as a timestamp, input source, AI model or version, generated output, provenance link and record of human intervention.[7]

The purpose is not to capture every interaction but to keep enough information so that, if needed, it can be later determined who used the AI, for what purpose, what material was used and whether it was verified and who relied upon it. Therefore, traceability is more than retaining just an AI prompt.

When responsibility is distributed across a working system this becomes significantly important. For example, an advocate may use AI, the junior will verify the generated authorities, a senior may approve the pleading and later the Court may rely upon it. If later an error is found and there is no record of all the processes happened at each stage, it may be difficult to find out at which stage the error occurred, whether independent verification was done and who had an opportunity to notice the error. Therefore, traceability makes

[6] Greenopolis Welfare Association (GWA) v. Narendar Singh & Ors. CM(M) 1909/2025.

[7] *AI Recommendation Traceability Systems for Contract Source Linking*, SIRION (June 29, 2026), https://www.sirion.ai/de/library/contract-insights/ai-traceability-contract-recommendations/.

accountability practical and reviewable. Instead of blaming someone for the mistake, the record can help to identify what happened, at which stage and by whom.

Traceability can also act as a corrective function. If later an error is found, the documented process can help to trace how the error occurred, whether from a fabricated authority, inefficient verification, misinterpretation or due to outdated law. This allows the error to be corrected or to improve the process, rules and checks so that similar errors are less likely to happen in future.

This principle is reflected in various AI governance. UNESCO's Recommendation on the Ethics of Artificial Intelligence recognises accountability, auditability, traceability and human oversight as governing principles.[8] The Supreme Court's Draft Regulations for Use of Artificial Intelligence in Courts, 2026[9] also highlights human primacy, accountability and transparency. As a draft, they represent a developing approach.

INTERNATIONAL RESPONSE: FROM AI USE TO PROFFESSIONAL ACCOUNTABILITY

International approaches also increasingly address traceability, institutional safeguards and the risks of unverified AI outcomes entering into legal procedures.

UNESCO's Recommendation on the Ethics of Artificial Intelligence[10]: This is significantly relevant to traceability. It connects accountability with auditability and traceability and holds human or legal actors responsible for the actions involving AI.

UNESCO's Guidelines for the use of AI systems in courts and tribunals[11]: These guidelines specifically focus on recognising AI-related risks arising in courts and tribunals. In research or preparation of any legal material, if AI is used in this kind of work, the error from its output can enter into the reasoning process at an earlier stage. The guidelines therefore recognise human review as a part of judicial process, rather than considering it as a safeguard applied only at the point of final decision-making.

---

[8] *Recommendation on the Ethics of Artificial Intelligence*, UNESCO (last visited Sept. 17, 2026), https://www.unesco.org/en/artificial-intelligence/recommendation-ethics.

[9] SUPREME COURT OF INDIA, *Notice: Regulations for Use of Artificial Intelligence (AI) in Courts, 2026 (Draft)* (June 3, 2026), https://cdnbbsr.s3waas.gov.in/s3ec0490f1f4972d133619a60c30f3559e/uploads/2026/06/2026060342.pdf.

[10] *Recommendation on the Ethics of Artificial Intelligence, supra note 8.*

[11] Juan David Gutiérrez, *Guidelines for the Use of AI Systems in Courts and Tribunals*, UNESCO (Dec. 3, 2025), https://www.unesco.org/en/articles/guidelines-use-ai-systems-courts-and-tribunals.

The EU Artificial Intelligence Act[12]: This act adopts a more specific risk-based approach. It identifies certain AI systems that judicial authorities use in research and interpreting facts and law as high-risk. It recognises the consequences of error and hindrance in judicial processing. It also requires human oversight to be the final decision-making body.

Murray v. Victoria (2025)[13]: This concern is also reflected in practice. A junior solicitor used AI-assisted research and produced documents containing false citations in Court. The Federal Court stated that unverified use of AI caused costs, inconvenience and delay, emphasising that legal practitioners must verify the AI-generated outcomes. The law firm was ordered to pay costs to respondents on an indemnity basis.

Taken together, the above approaches show that now the concern is not only whether can be AI used, but what conditions are to be followed to use the outputs generated by AI.

FROM PRINCIPLES TO PRACTICE: A RISK-BASED VERIFICATION FRAMEWORK

The earlier analysis already built the basis of what must be verified and why traceability is important. But now the question arises about these checks can be applied without subjecting every AI use to the same level of intensity as not every outcome has the same consequences.

The article therefore proposes different-level verification framework. The more serious consequence of any outcome the more serious its verification should be.

Low Risk: It includes tasks where AI is used in basic organisation or preliminary research. It includes organising documents, finding keywords and generating preliminary search terms. The material provided by AI still needs to be reviewed but there is no need for comprehensive source verifying.

Moderate Risk: This occurs when AI starts working more than basic assistance and begins doing actual legal work such as finding authorities, summarising judgements and interpreting statutes. As AI is now dealing with legal authorities and propositions of law, the primary sources should be verified independently and examined for proper context.

High Risk: In this the output has a direct effect on legal proceedings. Tasks such as court filing, legal submissions, reasoning, affecting rights and liablities of any person. Verification at this

[12] FUTURE OF LIFE INSTITUTE, *EU Artificial Intelligence Act*, EU ARTIFICIAL INTELLIGENCE ACT (last visited Sept. 18, 2026), https://artificialintelligenceact.eu/.
[13] Murray v. Victoria (2025) FCA 731 (Austl.).

stage should be extensive and must check existence of the authority, attribution, interpretation, current legal status and relevance. Additional review must be done if the consequences of the error justify it.

The level of verification, must be aligned with the consequences of relying on AI-generated materials. The framework can be applied through six stages:

Generate →Verify →Evaluate →Decide →Record →Correct.

First, AI generates the information, secondly, it is verified through original source. The information is then examined for its meaning, relevance and validity. A human considers and decides whether the information can be reliably used. Wherever necessary, the use of AI can be recorded for accountability and review. Finally, if any error occurs after relying on the information, it can be corrected. This framework turns human verification merely from a theoretical requirement to a practical process.

CONCLUSION

The key problem with the use of AI in research is not the fabricated information it provides but the acceptance of information without being correctly checked because it sounds too sophisticated. Convincingly incorrect material is hard to detect. Therefore, the role of verification is not only in preventing AI errors but also in maintaining reliability of the legal reasoning. The system does not have to choose either to use AI or to maintain research standards for accurate material. AI can be used, provided that the output is being checked, questioned and corrected. So, the responsible use of AI is not in how quickly AI enters the working system but how difficult it is for an unchecked error to enter and remain there.