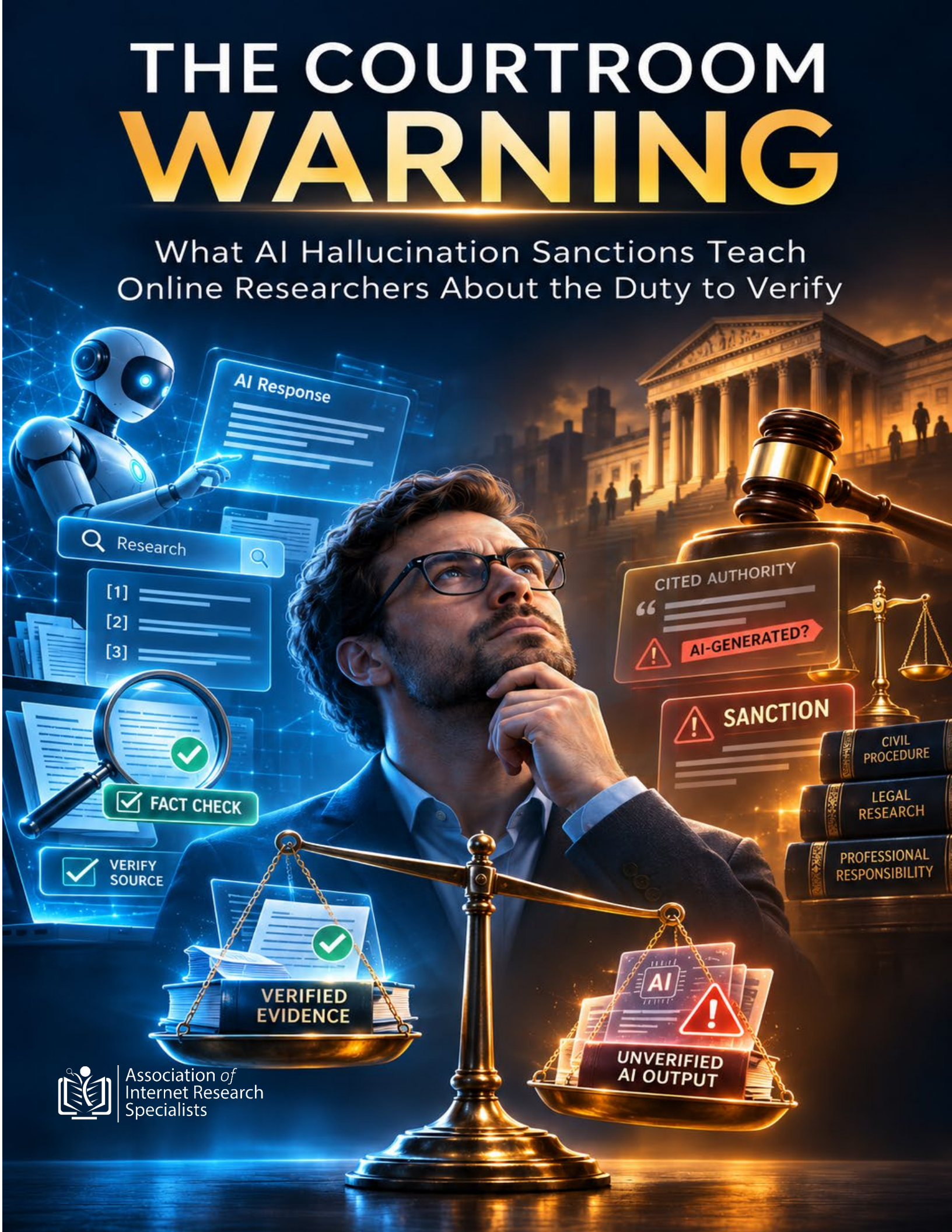
THE COURTROOM WARNING

What AI Hallucination Sanctions Teach Online Researchers About the Duty to Verify



Association *of* Internet Research Specialists

AOFIRS RESEARCH REPORT · NOVEMBER 2026

The Courtroom Warning: What AI Hallucination Sanctions Teach Online Researchers About the Duty to Verify

Judicial sanctions for AI hallucinations, the non-delegable duty to verify, and a verification protocol for online research professionals

Naveed Manzoor, MS
Founder & CEO, Association of Internet Research Specialists (AOFIRS)
Creator of the Certified Internet Research Specialist (CIRS™) Program

Table of Contents

Abstract iv

1. Introduction 1
- 1.1 Why the Courtroom Matters to Online Researchers 1
- 1.2 Research Questions and Approach 1

2. Understanding AI Hallucination 3
- 2.1 Why Models Make Things Up 3
- 2.2 How Often: The Peer-Reviewed Evidence 3
- 2.3 Agentic Research Tools: More Citations, Same Problem 5

3. The Scale of the Courtroom Problem 7
- 3.1 Growth of the Documented Case Record 7
- 3.2 Who Submits Hallucinated Material 7
- 3.3 What Kinds of Errors Courts Find 8
- 3.4 Where Courts Confront the Problem 9

4. Landmark Cases: From Warning to Sanction 11
- 4.1 Mata v. Avianca (2023): The Case That Verified Itself 11
- 4.2 Park v. Kim (2024): The Appellate Floor 11
- 4.3 Wadsworth v. Walmart (2025): The Signature Is the Verification 11
- 4.4 Lacey v. State Farm (2025): The Multi-Tool, Multi-Firm Pipeline 11
- 4.5 Johnson v. Dunn (2025): When Fines Stopped Being Enough 12
- 4.6 Ayinde and Al-Haroun (2025): England's Profession-Wide Warning 12
- 4.7 Kohls v. Ellison (2025): The Expert Who Trusted the Machine 12
- 4.8 Couvrette v. Wisnovsky (2025–2026): The Six-Figure Outlier 12
- 4.9 Fletcher and Whiting (2026): Appellate Courts Close the Door 13
- 4.10 From Fines to Licenses 13
- 4.11 The Bench Is Not Exempt 13
- 4.12 Many Smaller Cases, Same Direction 13

5. The Sanctions Architecture 16
- 5.1 Legal Bases for Sanctions 16
- 5.2 The Escalation Ladder 16
- 5.3 Why the Response Matters More Than the Mistake 17
- 5.4 Consequences Beyond Money 18

6. The Non-Delegable Duty to Verify 19
- 6.1 Doctrinal Roots 19
- 6.2 Delegation to People Was Never a Defense 19
- 6.3 Delegation to Machines Is No Different 20
- 6.4 The Chain of Accountability 20

7. When Verification Itself Depends on AI 22
- 7.1 Circular Verification: Asking the Model to Vouch for Itself 22
- 7.2 Summaries and "Research-Grade" Tools as Stand-Ins for Sources 22
- 7.3 Prompting Is Not Verification 23
- 7.4 Why Capable People Stop Checking 23
- 7.5 Consequences of AI-Dependent Verification 23

8. Lessons for Online Researchers 25
- 8.1 Why the Courtroom Standard Applies to Research 25
- 8.2 Seven Transferable Principles 25
- 8.3 A Six-Gate Verification Protocol 26

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

8.4 Governance for Research Teams 28
8.5 Implications for Professional Training and Certification 28
9. The Regulatory Horizon **29**
10. Conclusion **30**
References **31**
Appendix A. Methodological Note on the Case Data **34**

List of Figures

Figure 1. Hallucination and Fabrication Rates Reported in Peer-Reviewed Studies 4
Figure 2. Citation Failure in Deep-Research Agents Versus Search-Augmented Models 5
Figure 3. Growth of Documented AI-Hallucination Decisions 7
Figure 4. Parties Responsible for Hallucinated Submissions 8
Figure 5. Nature of Hallucinations Identified by Courts 9
Figure 6. Decisions Addressing Hallucinated Material by Jurisdiction 10
Figure 7. Escalation of Monetary Consequences for AI-Hallucinated Filings, 2023–2026 15
Figure 8. The Sanctions Escalation Ladder for Unverified AI Output 17
Figure 9. The Non-Delegable Duty: Work Can Be Delegated, Accountability Cannot 20
Figure 10. Circular AI-Dependent “Verification” Versus Independent Verification 22
Figure 11. The Six-Gate Verification Protocol for AI-Assisted Research 26

List of Tables

Table 1. A Typology of AI Hallucination Across Legal and Online Research Contexts 4
Table 2. Landmark AI-Hallucination Decisions and the Verification Lessons They Establish 14
Table 3. Principal Legal Bases for Sanctioning Reliance on Hallucinated AI Output 16
Table 4. Aggravating and Mitigating Conduct After Hallucinated Material Is Discovered 18
Table 5. Delegation Did Not Transfer Responsibility: Who Relied on Whom 19
Table 6. AI-Dependent Verification Compared With Independent Verification 24
Table 7. Translating Courtroom Lessons Into Online Research Practice 25
Table 8. Pre-Delivery Verification Checklist for AI-Assisted Research Reports 27

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

Abstract

Generative AI is now a routine part of professional research, and courtrooms have become the place where its failures are most visibly caught and punished. This report looks at how judges in the United States, England and Wales, Canada, Australia, and elsewhere have dealt with filings that contained AI hallucinations, such as invented authorities, false quotations, and misstated holdings, and asks what those decisions mean for people who do research online.

The analysis draws on a September 2026 snapshot of the Charlotin AI Hallucination Cases Database (2,095 decisions in more than 40 jurisdictions), peer-reviewed studies of hallucination rates, and the decisions themselves. Together they show steady escalation, from a $5,000 penalty in *Mata v. Avianca* (2023) to six-figure financial sanctions, disqualification, dismissal with prejudice, and suspension from practice in 2025 and 2026.

Three findings come through clearly. Courts treat the duty to verify as personal and non-delegable, so it cannot be passed to a paralegal, co-counsel, client, intern, vendor, or AI agent. Courts also treat verification that itself relies on AI as no verification at all, whether that means asking a model to confirm its own output, relying on an AI-generated summary, or trusting a research-grade tool without reading the primary source. Finally, sanctions depend less on the original error than on how the professional responds once it comes to light. Candor, quick correction, and remediation help; denial and concealment make things worse.

The report turns these findings into a six-gate verification protocol (Exist, Retrieve, Match, Support, Current, Record), a pre-delivery checklist, and governance recommendations for OSINT analysts, investigators, due-diligence specialists, journalists, and academic researchers, whose work, like a court filing, carries an implied promise of accuracy.

Keywords: *AI hallucination; sanctions; duty to verify; Rule 11; non-delegable duty; automation bias; agentic AI; citation verification; online research; OSINT*

> **About this report**
>
> I prepared this report for the professional development community of the Association of Internet Research Specialists (AOFIRS). It is meant to teach verification practice and is not legal advice. Case descriptions summarize published decisions and reliable reporting, and the database figures reflect a snapshot taken on 28 September 2026. Please read the underlying decisions before relying on any detail.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

1. Introduction

1.1 Why the Courtroom Matters to Online Researchers

Anyone who works with sources eventually runs into an uncomfortable question: who is responsible when a cited source turns out not to exist? For most researchers the answer comes quietly, in a correction notice, a retraction, or a client who stops calling. In a courtroom it comes on the public record, with the judge's name on it and, more and more often, a sanction order attached. That openness is what makes court decisions such useful evidence. They show in unusual detail what happens when professionals let generative AI do part of their work and then skip the checking.

There is no shortage of examples. As of 28 September 2026, the AI Hallucination Cases Database kept by legal researcher Damien Charlotin listed 2,095 decisions in which a court or tribunal found, or strongly implied, that a party had relied on hallucinated AI content (Charlotin, 2026). In July 2025 the count stood at 160 (Manach, 2025). Nor is this mainly a problem of people representing themselves without legal training. Lawyers appear in 829 of the logged decisions, judges or tribunal members in 33, and expert witnesses in 15.

I wrote this report for people who do research for a living outside the courtroom: open-source intelligence (OSINT) analysts, investigators, due-diligence and background-check specialists, fact-checkers, librarians, and academic researchers. Their reports work much like a legal brief. Whoever reads one assumes the author has looked at the sources and that the sources say what the report says they say. Over the past three years, judges have spelled out, one case at a time, what that assumption requires once AI is part of the workflow. The sections that follow pull those rulings together and turn them into standards a research team can actually use.

1.2 Research Questions and Approach

Four questions guide the analysis:

1. **RQ1.** How have courts described and sanctioned reliance on hallucinated AI output, and how has the severity of sanctions changed between 2023 and 2026?
2. **RQ2.** On what legal and ethical grounds do courts treat the duty to verify as personal and non-delegable, including when the work is delegated to software tools and AI agents?
3. **RQ3.** What happens when verification itself depends on AI?
4. **RQ4.** Which of these standards carry over to professional online research?

The evidence comes from four places. The first is the decisions themselves, along with published orders, from the United States, England and Wales, Canada, and Australia. The second is a 28 September 2026 snapshot of the Charlotin database, which is published under a Creative Commons Attribution 4.0 license. The third is the peer-reviewed and preprint research that has measured how often AI systems hallucinate in legal and scholarly work. The fourth is professional-ethics guidance, chiefly American Bar Association (ABA) Formal Opinion 512 and the guidance reviewed by the Divisional Court of England and Wales. Where I could not access a decision directly, I confirmed the facts through law-library analyses, legal journalism, and public broadcasters.

Readers should keep three limits in mind. The database captures only decisions that were reported and that dealt with hallucinated material, so its counts are a floor rather than a census. Reported dollar amounts often lump fines together with fee awards, which makes them a measure of total financial

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

exposure rather than of punishment alone. And coverage is thinner outside the United States, which inflates the American share of the record.

Key terms used in this report

AI hallucination: AI output presented as fact that is actually fabricated, distorted, or unsupported. Typical examples are a case that does not exist, a quotation a real source never contained, or a real source cited for a point it does not make.

Agentic AI: systems that carry out multi-step tasks on their own, such as searching, picking sources, reading, summarizing, citing, and drafting, with little human input between steps. “Deep research” products are the best-known example.

Duty to verify: the professional obligation to confirm, before relying on or submitting work, that each source exists, says what it is cited for, and is still valid.

Non-delegable duty: a duty that others can help you perform but cannot take off your hands.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

2. Understanding AI Hallucination

2.1 Why Models Make Things Up

Large language models write by predicting which words are likely to come next. Nothing in that process checks a citation against a list of real authorities, and polished prose says nothing about whether a source is real. Researchers who study the problem distinguish intrinsic hallucinations, which contradict the source material, from extrinsic ones, which cannot be checked against it at all (Ji et al., 2023). More recent work argues that the way models are trained and scored rewards a confident guess over an honest admission of uncertainty. If that is right, hallucination is a predictable result of how these systems are built and graded, and it will not simply fade away as models improve (Kalai et al., 2025).

Judges have reached much the same conclusion. The Divisional Court of England and Wales found that freely available generative tools cannot be relied on for legal research, because they can produce answers that are coherent and plausible and still entirely wrong (*R (Ayinde) v London Borough of Haringey*, 2025). In February 2026 the U.S. Court of Appeals for the Fifth Circuit noted that the problem showed no sign of easing and that it now often takes the form of false quotations or misstatements of law attributed to real cases. Those errors are far harder to spot than a case that plainly does not exist (*Fletcher v. Experian Information Solutions, Inc.*, 2026).

2.2 How Often: The Peer-Reviewed Evidence

The most careful measurement in the legal field comes from a Stanford team that asked general-purpose models specific, checkable questions about randomly selected federal cases. Depending on the model, between 58% (GPT-4) and 88% (Llama 2) of the answers were hallucinated, and the models did worst on harder tasks such as identifying a case's core holding (Dahl et al., 2024). The models also could not tell when they were getting things wrong. In practice, a model that sounds sure of itself gives you no information about whether it is right.

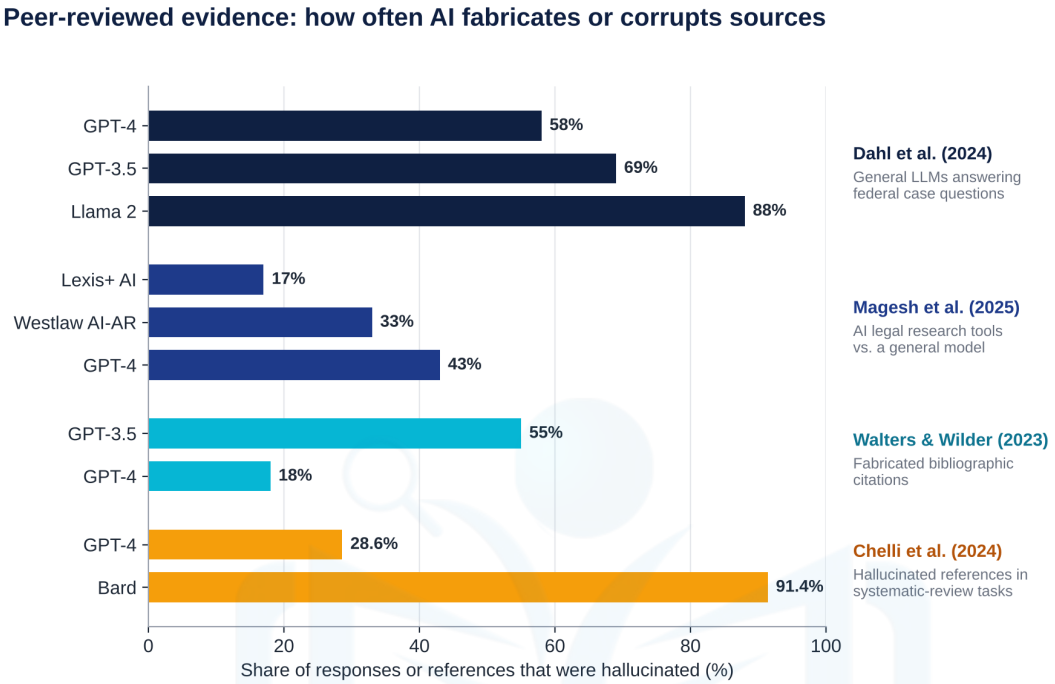
Specialized legal tools do better, but they are a long way from the hallucination-free performance some vendors advertised. In the first preregistered test of commercial legal research products built on retrieval, Lexis+ AI hallucinated on about 17% of queries and Westlaw AI-Assisted Research on about 33%, against 43% for GPT-4 (Magesh et al., 2025). The authors also found that these tools sometimes cite a real authority for a point it does not support. That is arguably the more dangerous failure, since anyone who only checks that the case exists will come away satisfied.

Law is not a special case. When researchers asked ChatGPT to write literature reviews on 42 topics, 55% of the GPT-3.5 citations and 18% of the GPT-4 citations were fabricated, and many of the genuine citations contained errors (Walters & Wilder, 2023). In a medical systematic-review task, the share of hallucinated references ranged from 28.6% for GPT-4 to 91.4% for Bard (Chelli et al., 2024). Figure 1 brings these results together.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

Figure 1

Hallucination and Fabrication Rates Reported in Peer-Reviewed Studies



Note. Studies used different tasks, models, and definitions, so rates are not directly comparable across groups. Model versions reflect the periods tested (2023–2024); newer systems generally perform better but have not eliminated the problem.

Table 1 sorts these errors into types that turn up both in court filings and in online research. The point of separating them is practical, because each type is caught by a different check. Looking a source up will expose one that was invented, but only reading it will expose a false quotation or a misstated finding.

Table 1

A Typology of AI Hallucination Across Legal and Online Research Contexts

Error type	How it appears in court filings	Online-research analogue	Detection difficulty
Fabricated source	An invented case, statute, or article with a plausible citation and real judges' names	An invented study, report, news story, URL, or DOI	Low to moderate: looking it up catches it
False quotation	Words attributed to a real decision that never appear in it	Invented or altered quotations or statistics attributed to a real document	High: the source exists, so quick checks pass
Misrepresented authority	A real case cited for a point it does not support	A real study cited for a conclusion it does not reach	High: requires reading and judgment
Outdated or overturned material	Reversed decisions or repealed laws presented as current	Retracted papers, corrected stories, superseded data, changed web pages	Moderate: requires a currency check
Fabricated facts or exhibits	Invented record facts or documents presented as evidence	Invented company facts, biographical details, or events	Moderate to high

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

Error type	How it appears in court filings	Online-research analogue	Detection difficulty
Fabricated URL	Rarely tracked separately in case law	Plausible but non-existent links in AI answers and deep-research reports	Low with automated link and archive checks

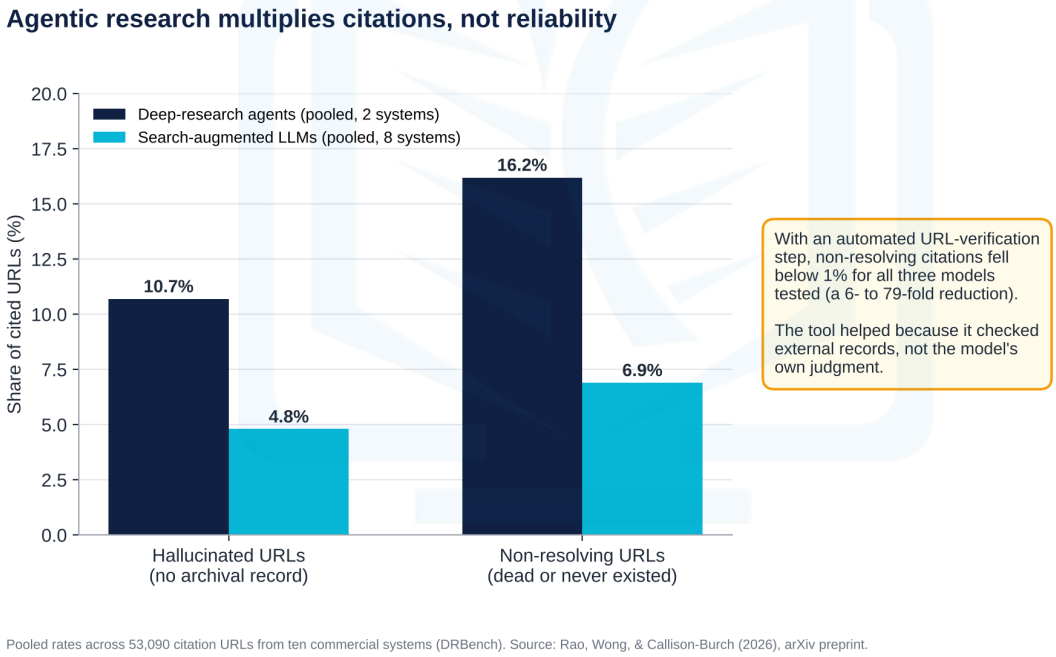
Note. Developed for this report from the error categories used by Charlotin (2026) and the findings of Dahl et al. (2024), Magesh et al. (2025), and Rao et al. (2026).

2.3 Agentic Research Tools: More Citations, Same Problem

The newest research tools are agentic. A deep-research system plans a series of searches, opens pages, writes a long report, and attaches dozens of citations. The result looks more authoritative than a chatbot reply, but more citations are not the same as better ones. A 2026 study of more than 53,000 citation URLs from ten commercial systems found that between 3% and 13% of URLs were hallucinated, meaning they had no archival record and probably never existed, and that 5% to 18% did not resolve at all. The two deep-research agents in the study produced hallucinated URLs at a pooled rate of 10.7%, compared with 4.8% for the search-augmented models (Rao et al., 2026).

Figure 2

Citation Failure in Deep-Research Agents Versus Search-Augmented Models



Note. Pooled rates from DRBench (53,090 URLs). “Hallucinated” URLs have no Wayback Machine record; “non-resolving” URLs return errors and include ordinary link rot. Data from Rao et al. (2026), a preprint under review.

The same study also points toward a remedy. When models were given an automated tool that checked each URL’s live status and archive history, the share of non-resolving citations fell below 1% for all three models tested, a 6- to 79-fold improvement (Rao et al., 2026). What made the difference was checking against records that sit outside the model. A second pass through the same model would have had no such information to work with.

Two other findings help explain why people stop checking when the output looks polished. An audit of four generative search engines found that only about 51.5% of the sentences they produced were fully supported by the citations attached to them, even though the answers read as fluent and informative

(Liu et al., 2023). And a 2026 survey of researchers found a wide gap between what people say and what they do: 87% said they always verify AI-generated citations, yet 42% admitted pasting AI-supplied references into their work without checking them (Xu et al., 2026, as cited in Rao et al., 2026). The same body of work reports that hallucinated references in computational-linguistics conference papers rose more than tenfold between 2024 and 2025 (Sakai et al., 2026, as cited in Rao et al., 2026).

What “agentic dependence” means in this report

By agentic dependence I mean handing a chain of research steps (searching, choosing sources, reading, summarizing, citing, drafting) to an AI system and then using what it produces with little or no independent checking. In the case law it shows up as chatbots drafting whole briefs, in-house AI platforms producing motions, research outlines built with several AI tools and passed from one firm to another, and AI-written case summaries used in place of the opinions themselves.

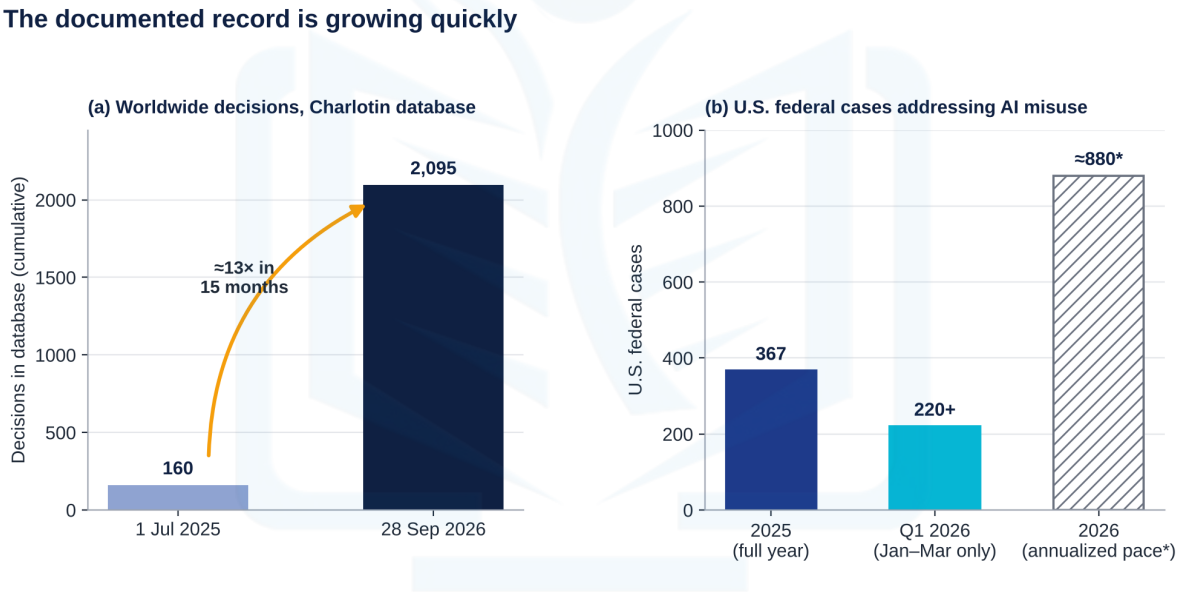
© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

3. The Scale of the Courtroom Problem

3.1 Growth of the Documented Case Record

In fifteen months the documented record grew about thirteenfold, from 160 decisions in July 2025 (Manach, 2025) to 2,095 by 28 September 2026 (Charlotin, 2026). It is still growing quickly. The first page of the database's results listed about fifty new decisions dated between 8 and 25 September 2026, close to three a day. Looking only at federal courts, a law-library review counted 367 U.S. cases involving AI-induced misrepresentations in 2025 and more than 220 in the first quarter of 2026 alone (University of Illinois Chicago School of Law Library, 2026). When the Fifth Circuit decided *Fletcher* in February 2026, it noted that the database already listed 239 U.S. decisions involving lawyers, not counting self-represented litigants and judges, and concluded that the problem was getting worse rather than better (*Fletcher v. Experian Information Solutions, Inc.*, 2026).

Figure 3
Growth of Documented AI-Hallucination Decisions



*Annualized pace = Q1 2026 count × 4; illustrative, not an observed total. Sources: Manach (2025); Charlotin (2026); University of Illinois Chicago School of Law Library (2026).

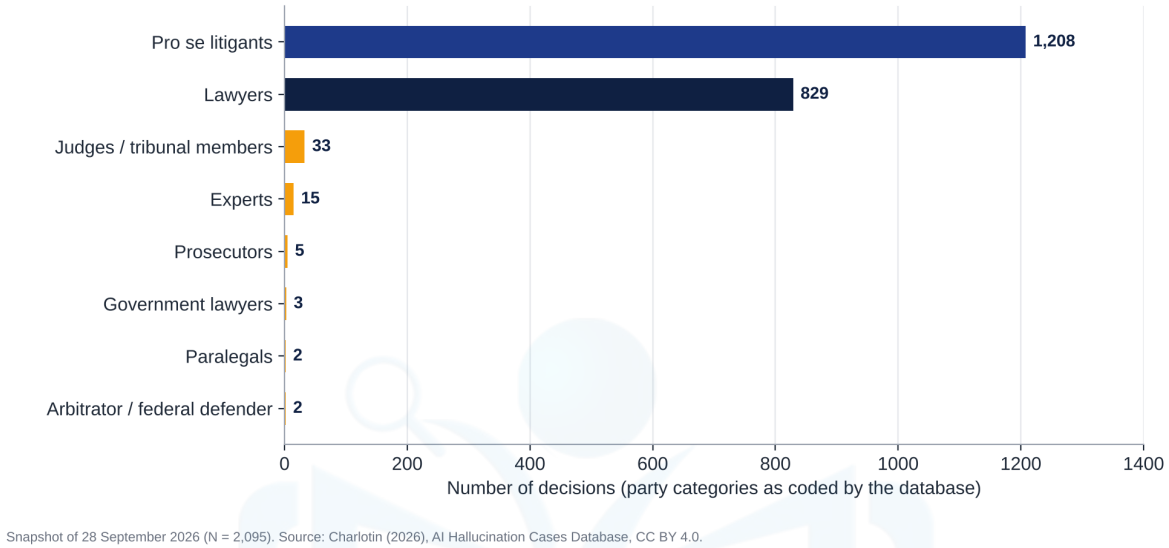
Note. Panel (a) compares two dated counts of the worldwide database. Panel (b) uses federal-docket counts compiled by the University of Illinois Chicago School of Law Library (2026); the 2026 bar is an illustrative annualization of first-quarter data, not an observed total.

3.2 Who Submits Hallucinated Material

Self-represented litigants account for 1,208 of the decisions (58%), which is hardly surprising now that anyone with a chatbot can produce something that reads like a legal brief. Lawyers account for another 829 (40%), and their cases carry the heaviest professional consequences. The rest are few in number but hard to ignore: 33 involve judges or tribunal members, 15 involve experts, and a handful involve prosecutors, government lawyers, and paralegals (Figure 4). In September 2026, for instance, the Supreme Court of India set aside a customs penalty order after finding that it rested on what appeared to be AI-generated judgments, some of which did not exist (Charlotin, 2026).

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

Figure 4
Parties Responsible for Hallucinated Submissions



Note. Snapshot of 28 September 2026 (N = 2,095). Party categories follow the database coding.

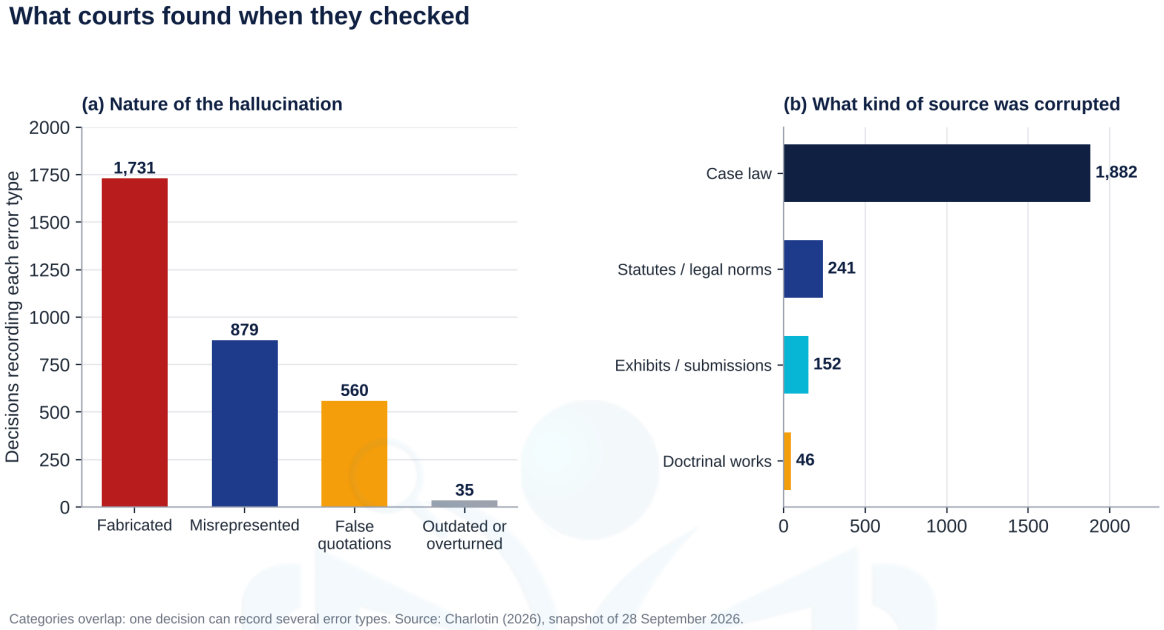
3.3 What Kinds of Errors Courts Find

Invented authorities are still the most common finding, recorded in 1,731 decisions. But misrepresented authorities (879 decisions) and false quotations attributed to real sources (560) together come close to that figure (Figure 5). Most of the problems involve case law (1,882 decisions), followed by statutes and other legal rules (241), exhibits and submissions (152), and scholarly or doctrinal works (46). For researchers, the takeaway is that showing a source exists is only the start. The harder and more important job is making sure it says what you claim it says.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

Figure 5

Nature of Hallucinations Identified by Courts



Note. Categories overlap because a single decision may record several error types.

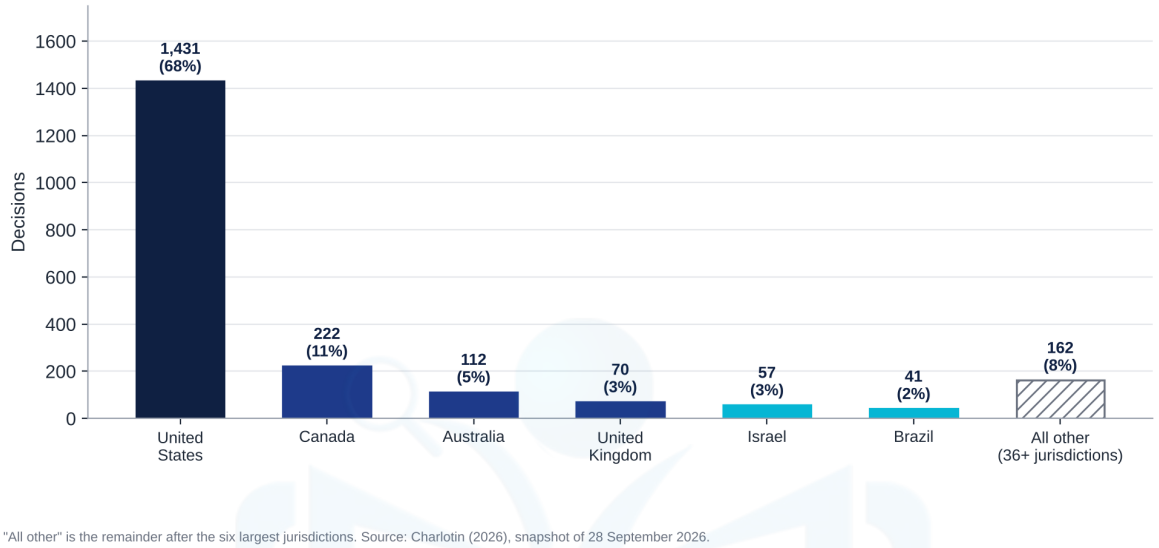
3.4 Where Courts Confront the Problem

The United States accounts for 1,431 decisions (68%), followed by Canada (222), Australia (112), the United Kingdom (70), Israel (57), and Brazil (41). Altogether the database covers more than 40 jurisdictions, plus domestic and international arbitration (Figure 6). The most affected areas of law are contract (450 decisions), administrative law (259), civil rights (203), employment (189), tort (183), and family law (136). In these areas many parties represent themselves, and filings lean heavily on research.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

Figure 6

Decisions Addressing Hallucinated Material by Jurisdiction



Note. The U.S. share partly reflects the volume and accessibility of American court records. “All other” was computed as the remainder of the 2,095 decisions.

4. Landmark Cases: From Warning to Sanction

I chose the cases below because each one set or sharpened a principle that later courts relied on. Read in order, they trace a clear path, from modest fines and stern words in 2023 to dismissed cases, disqualified lawyers, and suspended licenses in 2026. Table 2 at the end of the section summarizes them, and Figure 7 charts the money.

4.1 Mata v. Avianca (2023): The Case That Verified Itself

The case began as an ordinary personal-injury claim against an airline. Plaintiff's counsel filed an opposition brief built on decisions that ChatGPT had made up, including a fictitious *Varghese v. China Southern Airlines*. When neither the court nor opposing counsel could find the cases, counsel went back to the same chatbot, which assured them the cases were real, and then filed what it offered as excerpts. Judge P. Kevin Castel imposed a $5,000 penalty, ordered counsel to write to their client and to each judge falsely named as an author, and called the episode "an abuse of the adversary system" (*Mata v. Avianca, Inc.*, 2023). It remains the textbook example of what I call circular verification, discussed in Section 7.

4.2 Park v. Kim (2024): The Appellate Floor

The Second Circuit was the first federal appeals court to deal with the problem. An attorney admitted that a state-court decision cited in her reply brief had come from ChatGPT and did not exist. The court referred her to its Grievance Panel and ordered her to give her client a copy of its ruling. Along the way it stated the minimum standard that later courts have quoted again and again: Rule 11 requires attorneys to "read, and thereby confirm the existence and validity of, the legal authorities" they rely on (*Park v. Kim*, 2024, p. 615).

4.3 Wadsworth v. Walmart (2025): The Signature Is the Verification

Lawyers from a national plaintiffs' firm filed motions in limine citing cases that the firm's in-house AI platform had invented. The attorney who drafted the motions was fined $3,000 and lost his pro hac vice admission. The two colleagues who co-signed them were fined $1,000 each, even though they had not drafted the motions, because signing a filing certifies that the signer personally made a reasonable inquiry (*Wadsworth v. Walmart Inc.*, 2025). An enterprise AI tool, however well built into a firm's systems, does not change who answers for its output.

4.4 Lacey v. State Farm (2025): The Multi-Tool, Multi-Firm Pipeline

In an insurance dispute, lawyers used several AI tools to build a research outline, which went to co-counsel at a second firm and into a brief without anyone checking it. The special master, former Magistrate Judge Michael Wilner, flagged two citations. The lawyers removed those two and filed a revised brief that still contained several AI errors and still made no mention of AI. He struck the briefs, denied the discovery relief they had asked for, and ordered the firms to pay about $31,000 (*Lacey v. State Farm General Insurance Co.*, 2025). His main complaint was that the lawyers had been told about the problem, had every chance to fix it, and did not.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

4.5 Johnson v. Dunn (2025): When Fines Stopped Being Enough

Lawyers at a large regional firm defending Alabama corrections officials filed motions with citations that ChatGPT had fabricated, despite the firm's own warnings and policies about AI. A partner whose name was in the signature block said he had relied on his team. Judge Anna Manasco decided that a fine would not be enough, reasoning that "fines do not account for the extreme dereliction of professional responsibility" involved ("Federal Judge Sanctions Three Butler Snow Attorneys," 2025). Instead she publicly reprimanded the lawyers, ordered them to share her decision with their clients, opposing counsel, and the judges in their other cases, removed them from the case, and referred them to licensing authorities (*Johnson v. Dunn*, 2025).

4.6 Ayinde and Al-Haroun (2025): England's Profession-Wide Warning

Two referrals reached a Divisional Court of the King's Bench Division, presided over by the President of the King's Bench Division, and were heard together. In *Ayinde*, grounds for judicial review cited five cases that do not exist and misstated a housing statute; the barrister and the law centre that instructed her had already been ordered to pay wasted costs. In *Al-Haroun*, 18 of the 45 citations in materials filed for the claimant did not exist, and many of the rest were misquoted or beside the point. The solicitor had relied on legal research done by his own client, a non-lawyer (*R (Ayinde) v London Borough of Haringey*, 2025).

The court used the judgment to speak to the whole profession. Anyone who uses generative tools for research, it said, must check the output against authoritative sources before relying on it, and that duty is no lighter when a lawyer is relying on someone else's work. The court also set out the full range of responses open to it, from public admonishment and costs orders to regulatory referral, contempt proceedings, and referral to the police, and it referred the lawyers involved to their regulators. Its core instruction fits in a single sentence:

> *"The critical safeguard is to check any output by reference to an authoritative source."*
>
> *R (Ayinde) v London Borough of Haringey* (2025, para. 87)

4.7 Kohls v. Ellison (2025): The Expert Who Trusted the Machine

This case came out of litigation over a state law on political deepfakes. An expert declaration from a well-known misinformation scholar cited academic articles that did not exist; the errors crept in while an AI tool was being used to help draft it. The expert owned up to the mistakes. Judge Laura Provinzino excluded the whole declaration anyway, remarking on the irony of an expert on AI-enabled misinformation falling into the same trap and noting that an expert's credibility is hard to repair once it has been damaged (*Kohls v. Ellison*, 2025). Of all the cases here, this is the one closest to everyday research work. The mistake was a researcher's, not a lawyer's, and it cost the researcher the entire work product.

4.8 Couvrette v. Wisnovsky (2025–2026): The Six-Figure Outlier

This was an intrafamily dispute in Oregon. Summary-judgment briefs cited fifteen cases that do not exist and eight fabricated quotations attributed to real decisions. After the defense pointed out the problems, lead counsel removed some of the citations and refiled without fully explaining why, which the court saw as an attempted cover-up. Magistrate Judge Mark Clarke described the matter as "a notorious outlier in both degree and volume" (Robert, 2026). Across two orders the sanctions came to

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

roughly $110,000: about $15,500 in monetary sanctions plus roughly $80,000 of the defendants' fees against lead counsel, and about $14,000 against local counsel for failing to take a meaningful part in the filings. The court also dismissed the claims with prejudice (*Couvrette v. Wisnovsky*, 2025, 2026; University of Illinois Chicago School of Law Library, 2026).

4.9 Fletcher and Whiting (2026): Appellate Courts Close the Door

In *Fletcher v. Experian Information Solutions, Inc.* (2026), the Fifth Circuit found sixteen fabricated quotations and other misrepresentations in a reply brief. Counsel first said she had worked from publicly available versions of the cases. The court found that the services she named were generative-AI research products and that she had not been open about using them. It imposed a $2,500 sanction in a published opinion and brushed aside any suggestion that new technology called for new rules: "Modern generative AI may be a new technology, but the same sanctions rules apply."

A month later, in *Whiting v. City of Athens* (2026), the Sixth Circuit found more than two dozen fake citations and misstatements of the record across three consolidated appeals. It ordered the two attorneys responsible to pay the appellees' full appellate fees and double costs, fined each of them $15,000 in punitive sanctions, and said plainly that it did not matter whether AI had caused the errors. Any citation a lawyer has not personally read and checked breaches the duty, wherever it came from (Savino, 2026).

4.10 From Fines to Licenses

In April 2026 the Nebraska Supreme Court temporarily suspended an Omaha attorney after 57 of the 63 citations in his divorce-appeal brief turned out to be defective. He first denied using AI, then admitted it, calling his lack of candor a "grave error of judgment" (WOWT, 2026). In the same quarter, a federal court in Alabama barred an attorney from practicing in the district for three months, removed him from the case, and awarded more than $47,000 in fees (*Rivera v. Triad Properties Corp.*, 2026). Another federal court in the state awarded more than $55,000 in fees, ordered its sanctions decision filed in the attorney's other cases for twelve months, and recommended that he be found incompetent to practice (*Heimkes v. Fairhope Motorcoach Resort Condominium Owners Ass'n*, 2026). Outside the United States, the Solicitors Disciplinary Tribunal in England struck a lawyer off the Register of Foreign Lawyers in August 2026, and a court in Tanzania imposed a six-month suspension (Charlotin, 2026).

4.11 The Bench Is Not Exempt

Judges have made the same mistake. In mid-2025, federal judges in New Jersey and Mississippi withdrew orders that contained misquotations and named parties who had nothing to do with the cases. Asked to explain by the chair of the Senate Judiciary Committee, one judge said a law-school intern had used ChatGPT without permission. The other said a law clerk had used Perplexity to help draft and that the order went out before review was finished, calling it "a lapse in human oversight" (Reuters, 2025). Both judges said they had since tightened their review procedures, and one adopted a written AI policy. Like lawyers, judges can hand off drafting. They cannot hand off responsibility for what goes out under their names.

4.12 Many Smaller Cases, Same Direction

Plenty of less famous decisions point the same way. A Texas federal court fined an attorney $2,000 and ordered training on generative AI (*Gauthier v. Goodyear Tire & Rubber Co.*, 2024). A Colorado court fined two lawyers $3,000 each in a high-profile defamation case (*Coomer v. Lindell*, 2025). A

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

Florida federal court ordered more than $85,000 in fees and referred the lawyer for discipline (*ByoPlanet International, LLC v. Johansson*, 2025). An Illinois state court ordered a lawyer to pay $10,000 and his firm $49,500 (*Jordan v. Chicago Housing Authority*, 2025). In September 2026 a New York trial court struck petitions drafted with a contract-drafting AI tool that contained fabricated material and imposed $15,000 in sanctions and costs, and in Australia a party that had relied on hallucinated authorities was ordered to pay AUD 9,600 in costs (Charlotin, 2026). The tools involved range from research-grade legal AI to general chatbots to combinations of several assistants, a point Section 7 returns to.

Table 2

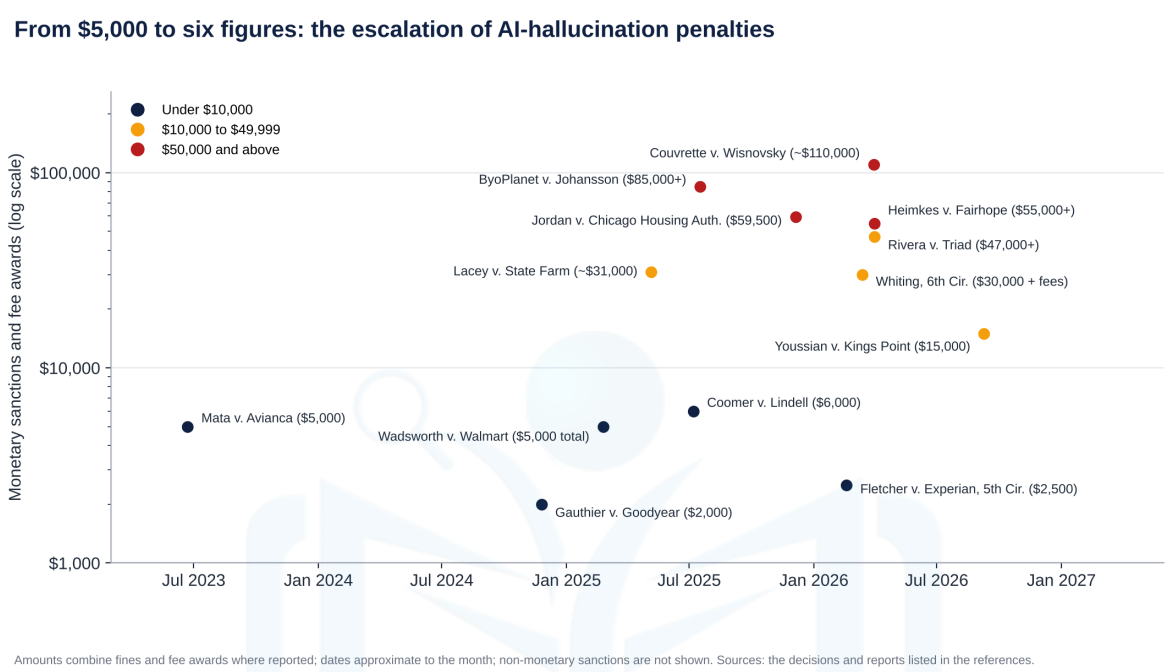
Landmark AI-Hallucination Decisions and the Verification Lessons They Establish

Case (court, year)	What went wrong	Consequence	Verification lesson
Mata v. Avianca (S.D.N.Y., 2023)	ChatGPT-generated non-existent decisions; counsel asked ChatGPT to confirm them	$5,000; letters to the client and to the judges falsely named	A tool vouching for itself is not verification
Park v. Kim (2d Cir., 2024)	Reply brief cited a non-existent decision drawn from ChatGPT	Referral to Grievance Panel; ruling sent to client	The minimum: read and confirm every authority
Wadsworth v. Walmart (D. Wyo., 2025)	In-house AI platform produced citations to non-existent cases	$3,000 and loss of pro hac vice status (drafter); $1,000 each for co-signers	Signing means you checked, even if you did not draft
Lacey v. State Farm (C.D. Cal., 2025)	AI-generated outline passed between firms; revised brief still wrong; AI use undisclosed	Briefs struck; relief denied; about $31,000	Responsibility does not get lost in a pipeline
Johnson v. Dunn (N.D. Ala., 2025)	Fabricated citations despite firm AI policies; reliance on the team	Reprimand, publication order, disqualification, bar referral	A policy nobody enforces protects nobody
Ayinde / Al-Haroun (Div. Ct., 2025)	Five fictitious cases; 18 of 45 citations non-existent, drawn from client research	Wasted costs; regulatory referrals; profession-wide warning	Relying on others, even the client, does not shift the duty
Kohls v. Ellison (D. Minn., 2025)	Expert declaration cited non-existent academic articles	Entire declaration excluded	Experts and researchers are held to the same standard
Couvrette v. Wisnovsky (D. Or., 2025–26)	15 non-existent cases, 8 fabricated quotations; quiet refiling	About $110,000; dismissal with prejudice	Hiding an error turns it into a six-figure problem
Fletcher v. Experian (5th Cir., 2026)	16 fabricated quotations; AI output described as public case text	$2,500 in a published opinion	An AI summary is not the source; candor counts
Whiting v. City of Athens (6th Cir., 2026)	More than 24 fake citations and record misstatements	Full appellate fees, double costs, $15,000 each	The duty applies whatever tool was used
Rivera v. Triad Properties (N.D. Ala., 2026)	AI-drafted filings with false statements of law	Three-month suspension in the district; disqualification; >$47,000	Sanctions now reach the right to practice
Nebraska Supreme Court (2026)	57 of 63 citations defective; AI use initially denied	Temporary license suspension	Denial makes it worse; candor helps
Federal judges, N.J. and Miss. (2025)	Staff-drafted orders using ChatGPT and Perplexity issued without full review	Orders withdrawn; written policy and added review	Judges cannot delegate the duty either

Note. Monetary amounts include fee-shifting awards where reported. Sources: the decisions and reports listed in the references.

Figure 7

Escalation of Monetary Consequences for AI-Hallucinated Filings, 2023–2026



Note. Log scale. The chart leaves out non-monetary sanctions, which are often more serious, and reflects reported decisions only. Dates are approximate to the month.

5. The Sanctions Architecture

5.1 Legal Bases for Sanctions

Courts have not needed new law to deal with any of this. As the Fifth Circuit put it, existing procedural rules, professional-conduct rules, and the court's inherent powers already require accuracy (*Fletcher v. Experian Information Solutions, Inc.*, 2026). Table 3 sets out the main sources of that authority.

Table 3

Principal Legal Bases for Sanctioning Reliance on Hallucinated AI Output

Basis	Jurisdiction	What it requires or permits	Illustrative decisions
Fed. R. Civ. P. 11(b)–(c)	U.S. federal trial courts	Signature certifies that legal contentions are warranted by existing law after a reasonable inquiry; sanctions on signer and firm	*Mata* (2023); *Wadsworth* (2025); *Gauthier* (2024)
Inherent authority	U.S. courts generally	Sanctions for bad faith and abuse of the judicial process, including misleading the court about where errors came from	*Johnson* (2025); *Heimkes* (2026)
Fed. R. App. P. 38 and 46(c)	U.S. courts of appeals	Damages and double costs for frivolous appeals; discipline for conduct unbecoming a member of the bar	*Whiting* (2026); *Park* (2024)
28 U.S.C. § 1927	U.S. federal courts	Fees for unreasonably and vexatiously multiplying proceedings	*Couvrette* (2025–26)
Professional-conduct rules	U.S. states (Model Rules 1.1, 3.1, 3.3, 5.1, 5.3, 8.4)	Competence, meritorious claims, candor, supervision of lawyers and non-lawyer assistance, honesty	ABA Formal Op. 512 (2024); Nebraska (2026)
Hamid jurisdiction, wasted costs, contempt	England and Wales	Enforcement of lawyers' duties to the court; costs; regulatory referral; contempt or police referral	*Ayinde / Al-Haroun* (2025)
Costs and abuse of process	Canada; Australia	Personal costs against counsel; findings that fake citations amount to misleading the court; regulator referral	*Zhang v. Chen* (2024); *Ko v. Li* (2025)

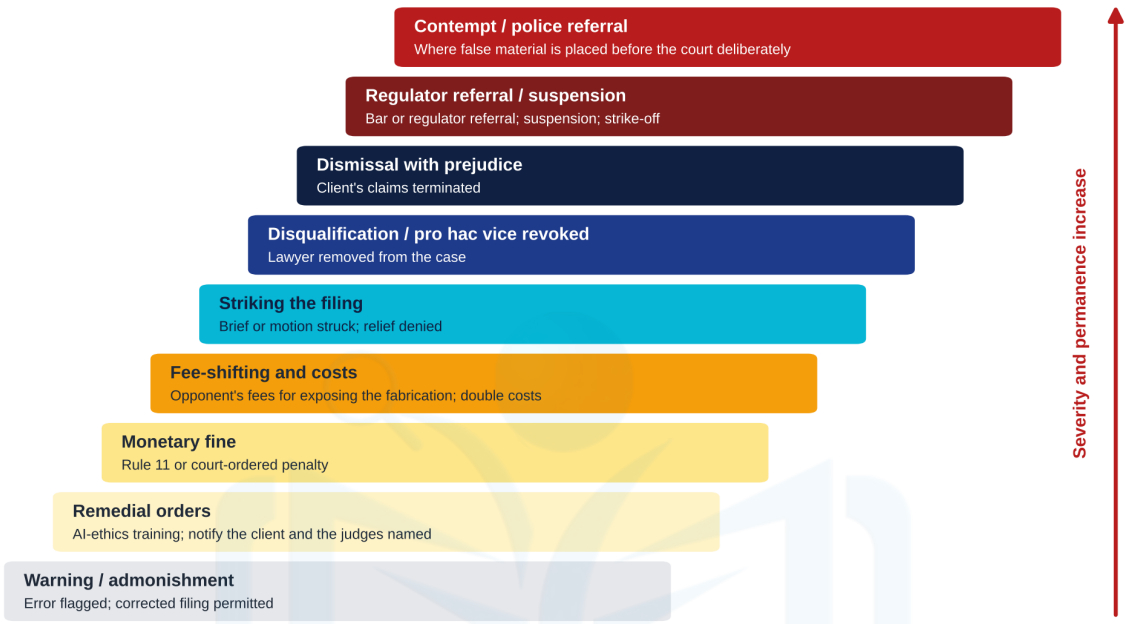
5.2 The Escalation Ladder

Looking across jurisdictions, the responses form something like a ladder (Figure 8). The bottom rungs, such as warnings, admonishments, and orders to take AI-ethics training, are usually kept for first offenses that were small and quickly corrected. Further up, courts impose fines and make the offending side pay what it cost the other side to uncover the problem. Higher still, they strike filings, remove counsel, dismiss claims with prejudice, refer lawyers to regulators, and suspend licenses. At the very top, the English Divisional Court confirmed that knowingly putting false material before a court can amount to contempt, or even a crime (*R (Ayinde) v London Borough of Haringey*, 2025).

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

Figure 8

The Sanctions Escalation Ladder for Unverified AI Output



Note. Synthesized from the decisions in Table 2 and the range of court powers described in *R (Ayinde) v London Borough of Haringey* (2025).

5.3 Why the Response Matters More Than the Mistake

Across the case law, how a professional responds once an error is found matters more than the error itself. The Divisional Court listed the factors it would weigh, including the need to set and enforce standards, how the false material came to be filed, whether a prompt, full, and truthful explanation was given, what was done to limit the harm, the time and money others lost, and the effect on the case (*R (Ayinde) v London Borough of Haringey*, 2025, para. 24). American courts reason the same way. The Fifth Circuit said it would probably have imposed a lighter sanction if counsel had been forthcoming (*Fletcher v. Experian Information Solutions, Inc.*, 2026), while the Oregon court went to six figures after an attempted cover-up (Robert, 2026).

Regulators see it similarly. An Oregon State Bar official told reporters that being open about how AI was used may earn a light response, while hiding it can lead to fines, suspension, or disbarment. The same official said that using AI well means “there has to be that human element who checks the work” (Martinez & Frost, 2026). Table 4 sets the conduct that tends to make things worse beside the conduct that tends to help.

Table 4

Aggravating and Mitigating Conduct After Hallucinated Material Is Discovered

Makes it worse	Helps
Denying or playing down AI use when asked (Nebraska, 2026; *Fletcher*, 2026)	Explaining promptly and voluntarily how the error happened

Makes it worse	Helps
Quietly removing flagged items and refiling without explanation (*Couvrette*, 2025–26; *Lacey*, 2025)	Withdrawing right away, filing a corrected version, and telling the court and the other side
Describing fabrications as minor citation slips	Apologizing to the court, the opponent, and the client; self-reporting to the regulator
The same problem repeated across filings or matters	A single, limited error that did not affect the outcome
Ignoring firm policies or supervisory duties (*Johnson*, 2025)	Fixing the process: training, a written policy, second-person review
Errors that waste the court's and opponents' time and money	Paying back the costs caused and cooperating fully

Note. Compiled from *R (Ayinde) v London Borough of Haringey* (2025, para. 24), the decisions discussed in Section 4, and Martinez and Frost (2026).

5.4 Consequences Beyond Money

Fines get the headlines, but they are rarely the worst of it. Clients lose motions, appeals, and sometimes whole cases when briefs are struck or claims dismissed. Publication orders make lawyers file the sanctions decision in their other cases or send it to every judge they appear before. Because opinions are published and the database is public, the record never goes away, and the damage to a reputation lasts far longer than any fine. A professional researcher's livelihood depends on being trusted, so for researchers the equivalent losses are clients, credentials, and the credibility of every report they have ever delivered.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

6. The Non-Delegable Duty to Verify

6.1 Doctrinal Roots

The duty to verify did not arrive with AI. Rule 11 of the Federal Rules of Civil Procedure has long provided that signing a filing certifies that the signer made a reasonable inquiry and that its legal arguments are supported by existing law. In *Pavelic & LeFlore v. Marvel Entertainment Group* (1989), the Supreme Court read the rule as aimed at the individual who signs. Two years later it held that the signer's duty is judged by an objective standard of reasonable inquiry and that it applies to represented parties who sign as well (*Business Guides, Inc. v. Chromatic Communications Enterprises, Inc.*, 1991). The 1993 amendments let courts hold law firms jointly responsible, but they left the individual signer's obligation intact. *Park v. Kim* (2024) simply carried that obligation into the AI era: a lawyer has to read and confirm every authority relied on.

6.2 Delegation to People Was Never a Defense

Ethics committees were dealing with outsourced research long before chatbots. New York's bar concluded that a lawyer who uses an outside research service must review the work and stays responsible for it (New York State Bar Association Committee on Professional Ethics, 1999). A Los Angeles opinion reached the same conclusion about work prepared by an outside service and filed under the lawyer's name (Los Angeles County Bar Association Professional Responsibility and Ethics Committee, 2006). The Divisional Court said the duty to check AI-assisted research applies just as much when a lawyer relies on others, such as trainees or instructed researchers, and held that a solicitor could not rely on his client for the accuracy of the citations (*R (Ayinde) v London Borough of Haringey*, 2025). Table 5 shows how consistently courts have turned down the argument that *someone else did the research*.

Table 5

Delegation Did Not Transfer Responsibility: Who Relied on Whom

Decision	Who relied on whom	Court's response
Al-Haroun (2025)	Solicitor relied on his client's legal research	Lawyer not entitled to rely on the client for accuracy; referral to regulator
Ayinde (2025)	Law centre relied on instructed counsel, whose sources were never explained	Both held responsible; adequacy of supervision questioned
Wadsworth (2025)	Co-signing attorneys relied on the drafting colleague	Co-signers sanctioned for signing without a reasonable inquiry
Johnson v. Dunn (2025)	Partner relied on his team and the firm's AI policies	Disqualification and referral; the policies were no defense
Lacey (2025)	Firm relied on co-counsel's AI-assisted research outline	Briefs struck; joint financial sanction
Couvrette (2025–26)	Local counsel relied on lead counsel	Local counsel ordered to pay fees for failing to take a meaningful part
Federal judges (2025)	Judges relied on an intern and a law clerk using AI tools	Orders withdrawn; judges took responsibility and changed their review process
Kohls v. Ellison (2025)	Expert relied on AI drafting help for references	Entire declaration excluded

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

6.3 Delegation to Machines Is No Different

Professional guidance treats AI as help that has to be supervised, not as a colleague who can share the blame. ABA Formal Opinion 512 applies the familiar duties of competence, confidentiality, communication, candor, supervision, and reasonable fees to generative AI. It expects lawyers to understand what a tool can and cannot do and to carry out "independent verification or review" suited to the task (American Bar Association Standing Committee on Ethics and Professional Responsibility, 2024, p. 4). An Ontario judge described the same idea as a lawyer's duty to make sure a human reviews material produced by non-human technology such as generative AI (*Ko v. Li*, 2025). The Sixth Circuit's position fills in the rest. It makes no difference whether the error came from a chatbot, a legal research platform, a template, or a colleague; the person who signed is answerable (Savino, 2026).

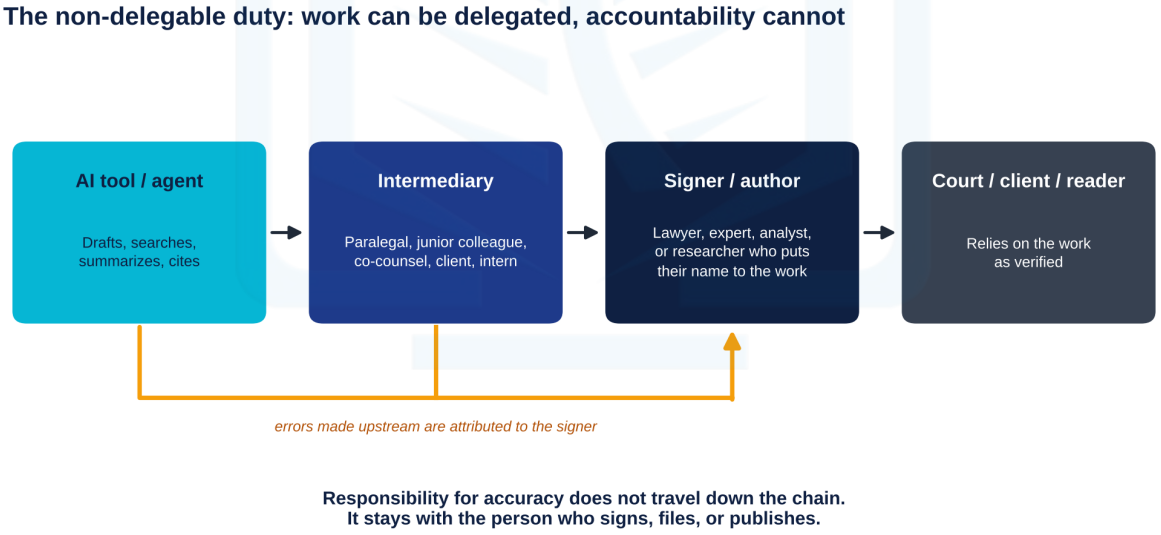
The reason is simple. Software cannot take an oath, be cross-examined, be sanctioned, or lose a license. Every system of professional accountability has to land on a person in the end, and it lands on the person who put the work forward as reliable.

6.4 The Chain of Accountability

Figure 9 shows how this works in practice. Work can pass from an AI tool to a paralegal, a junior colleague, co-counsel, a client, or an intern, but responsibility for getting it right always comes back to whoever signs, files, or publishes it.

Figure 9

The Non-Delegable Duty: Work Can Be Delegated, Accountability Cannot



Note. Synthesized from *Pavelic & LeFlore* (1989), *Business Guides* (1991), *Park v. Kim* (2024), *Wadsworth* (2025), *Ayinde / Al-Haroun* (2025), and ABA Formal Opinion 512 (2024).

What “non-delegable” does and does not mean

It does not rule out AI tools, research assistants, or splitting up the work. Courts, bar regulators, and the ABA all accept that AI can be useful when it is used carefully.

It does mean that whoever certifies the work has to make sure the checking actually happened, to a standard that matches how much others will rely on it.

It does mean that checking can be shared but not handed off entirely. A tool or a colleague can help, but the person who certifies needs a real basis for confidence.

It does mean that the more is riding on the output, the more direct the check should be. Anything that will be quoted, relied on, or published should be checked against the primary source.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

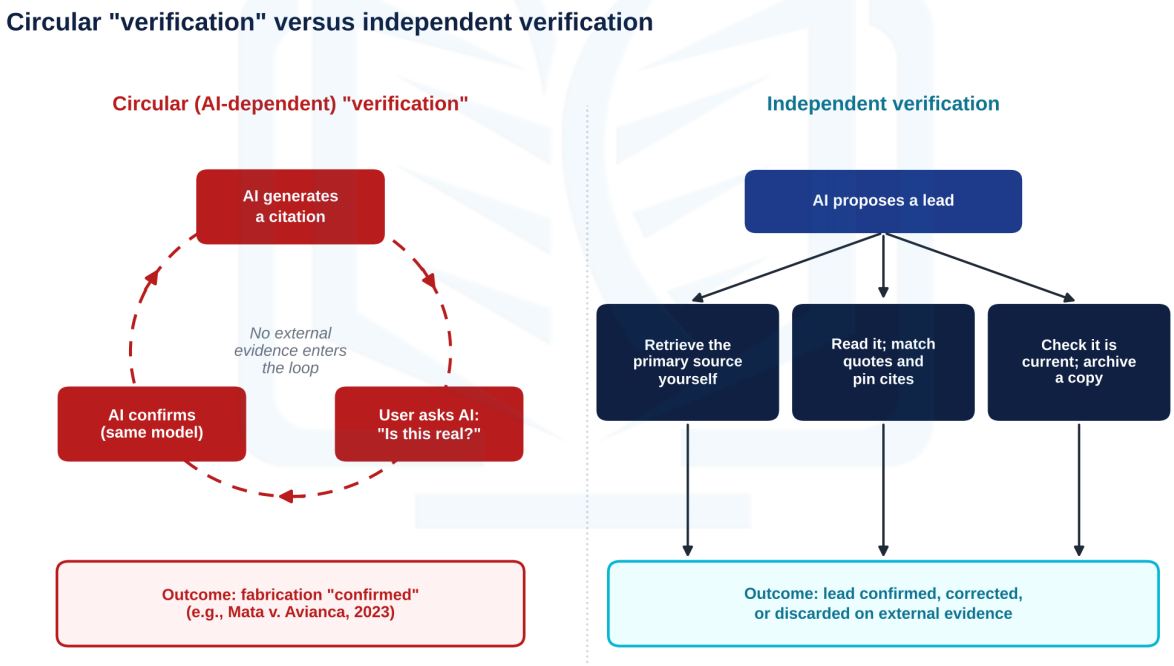
7. When Verification Itself Depends on AI

Not checking at all is one failure. A subtler one is checking with the wrong tool. In several of the most instructive cases, professionals sincerely believed they had verified their work, when their "verification" relied on the same kind of system that had produced the error in the first place.

7.1 Circular Verification: Asking the Model to Vouch for Itself

In *Mata v. Avianca* (2023), counsel asked ChatGPT whether the cases it had supplied were real, and it said they were. I call this circular verification (Figure 10). Asked whether something is real, a model produces another plausible answer from the same learned patterns; it does not look anything up in an outside record. Since models cannot reliably tell when they are hallucinating (Dahl et al., 2024), asking them adds no independent evidence. Using one AI system to check another, without going to the primary sources, carries the same risk, because the two may share training data and the same blind spots.

Figure 10
Circular AI-Dependent "Verification" Versus Independent Verification



Note. Conceptual model developed for this report. Left: the loop closes without external evidence. Right: AI output is treated as a lead, and confirmation comes only from sources outside the model.

7.2 Summaries and "Research-Grade" Tools as Stand-Ins for Sources

A second version of the mistake is treating an AI-written description of a source as if it were the source. In *Fletcher* (2026), counsel traced the fabricated quotations to what she called publicly available versions of the cases. The court found those were generative-AI products, not the opinions. Summaries, snippets, AI overviews, and case digests are descriptions of a source, and they can be wrong in ways the source itself cannot.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

Specialized tools are not exempt. The Stanford evaluation found hallucination rates of 17% to 33% in leading AI legal research products (Magesh et al., 2025), and the database includes sanctions against lawyers who used research-grade legal AI, contract-drafting AI, and several assistants at once (Charlotin, 2026). Even an ordinary web search can let AI in through the back door. In *Ayinde* (2025), the barrister suggested she might have relied on AI-generated summaries at the top of her search results without realizing what they were. The working rule for researchers is straightforward: a summary, snippet, AI overview, or tool output can point you to a source, but it is never the source.

7.3 Prompting Is Not Verification

Careful prompting, such as telling a tool to admit uncertainty or to cite specific paragraphs, may well improve what a model produces. But nothing in the case law treats a good prompt as a substitute for checking. The Divisional Court described the critical safeguard as checking the output against an authoritative source (*R (Ayinde) v London Borough of Haringey*, 2025, para. 87). In the experimental work, the biggest reductions in error came from checking citations against outside records, not from asking the model to be more careful (Rao et al., 2026). A better prompt gives you a better lead, but it is still only a lead.

7.4 Why Capable People Stop Checking

So why do capable professionals skip checks they know they should do? Human-factors research has a well-established answer. Automation bias is the tendency to lean too heavily on automated aids. It produces two kinds of mistakes: following bad advice (errors of commission) and missing problems the system did not flag (errors of omission). Complacency grows as a system becomes more reliable and as workload and time pressure rise (Parasuraman & Manzey, 2010). A systematic review of clinical decision-support tools found automation bias to be common and listed training and greater user accountability among the things that reduce it (Goddard et al., 2012).

Generative AI makes all of this worse. Its output is fluent, neatly formatted, and confident, and a deep-research report full of citations looks like the product of hard work. Answers from generative search engines come across as authoritative even when only about half their sentences are fully supported (Liu et al., 2023), and many researchers who say they always check citations do not (Xu et al., 2026, as cited in Rao et al., 2026). Good intentions have not closed that gap. Routines that are written down and actually followed stand a better chance.

7.5 Consequences of AI-Dependent Verification

Taken together, the cases and the research point to five consequences when verification depends on AI instead of primary sources:

1. **Errors move to where the checks are weakest.** As models get better at producing real-looking citations, the errors shift from invented cases toward false quotations and misstated holdings in real sources, which pass a simple existence check (*Fletcher*, 2026; Charlotin, 2026).
2. **The contamination spreads.** Fabricated authorities have found their way into judicial and administrative orders, and opposing parties, courts, and clients absorb the wasted time and cost (Reuters, 2025; Charlotin, 2026).

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

3. **Other people end up doing the checking.** Once filings cannot be trusted, courts and opponents have to verify everything themselves, and fee-shifting sanctions send that bill back to the responsible professional (*Whiting*, 2026).
4. **Credibility collapses all at once.** A few fabricated references can sink an entire work product, as the excluded expert declaration in *Kohls v. Ellison* (2025) shows.
5. **The person pays, not the tool.** No decision in the record excuses a professional because a tool, vendor, or assistant made the mistake.

Table 6

AI-Dependent Verification Compared With Independent Verification

Dimension	AI-dependent "verification"	Independent verification
Evidence relied on	What the same or another generative model says	The primary document, retrieved from an authoritative index or official source
What actually gets checked	Whether the output sounds plausible and consistent	Existence, exact wording, support for the claim, and currency
Typical failure	Confirms fabrications as confidently as facts (*Mata*, 2023)	Ordinary human error, reduced by checklists and a second reviewer
What it leaves behind	An AI summary, a chatbot confirmation, a tool-generated citation list	The official text or PDF, an archived copy, a verification log
How courts treat it	As no verification at all (*Fletcher*, 2026; *Ayinde*, 2025)	As meeting the reasonable-inquiry standard (*Park*, 2024; ABA, 2024)
Where AI legitimately fits	Never as the final judge	As triage: flagging dead links, missing identifiers, or quotation mismatches for a human to review (Rao et al., 2026)

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

8. Lessons for Online Researchers

8.1 Why the Courtroom Standard Applies to Research

Due-diligence reports, background checks, OSINT attributions, investigative articles, competitive-intelligence briefs, and literature reviews are all relied on by people who usually cannot check them. That reliance is exactly what gives a court filing its weight as a certification, and a research report carries the same weight for the same reason. The stakes are similar too. A fabricated source in a background report can defame the person it describes. A misquoted regulation can steer an investment or policy decision the wrong way. A hallucinated statistic can keep being repeated long after anyone remembers where it came from. In effect, the courts have written down, and started enforcing, rules that careful researchers already follow.

8.2 Seven Transferable Principles

1. **Your name is your certificate.** Whoever delivers the work owns its accuracy, no matter which tools or assistants helped.
2. **Read what you cite.** Confirm for yourself that the source exists and says what you need it to say (*Park v. Kim*, 2024).
3. **Check the primary source, not a description of it.** Summaries, snippets, AI overviews, and tool outputs are leads, not evidence (*Fletcher*, 2026).
4. **Never let a system grade its own homework.** An AI system confirming its own output is not verification (*Mata*, 2023).
5. **Treat "research-grade" labels as lowering risk, not removing it.** Specialized tools still hallucinate at meaningful rates (Magesh et al., 2025).
6. **Supervise in proportion to how much others will rely on the work.** Assistants, contractors, clients, and agents can do the work; you still have to check it (*Ayinde*, 2025).
7. **When you find a mistake, say so and fix it quickly.** Candor limits the damage; concealment multiplies it (*Couvrette*, 2025–26; *Fletcher*, 2026).

Table 7

Translating Courtroom Lessons Into Online Research Practice

Courtroom lesson	What it means for online research
A Rule 11 signature certifies a reasonable inquiry	Treat every report you deliver as a certification, and add an internal sign-off confirming the sources were checked
Read and confirm every authority (*Park*)	Open and read every cited document; never cite from memory, a snippet, or an AI summary
Circular verification is no verification (*Mata*)	Look up AI-supplied sources in independent indexes: publisher sites, DOI resolvers, official registries, court databases
False quotations in real sources (*Fletcher*, *Couvrette*)	Search the source text for every quotation and figure, and note the page or paragraph
Co-signers are liable (*Wadsworth*)	Reviewers and editors who approve a report need to check the sources, not just the writing
Relying on the client's research (*Al-Haroun*)	Check leads and documents from clients independently before repeating them

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

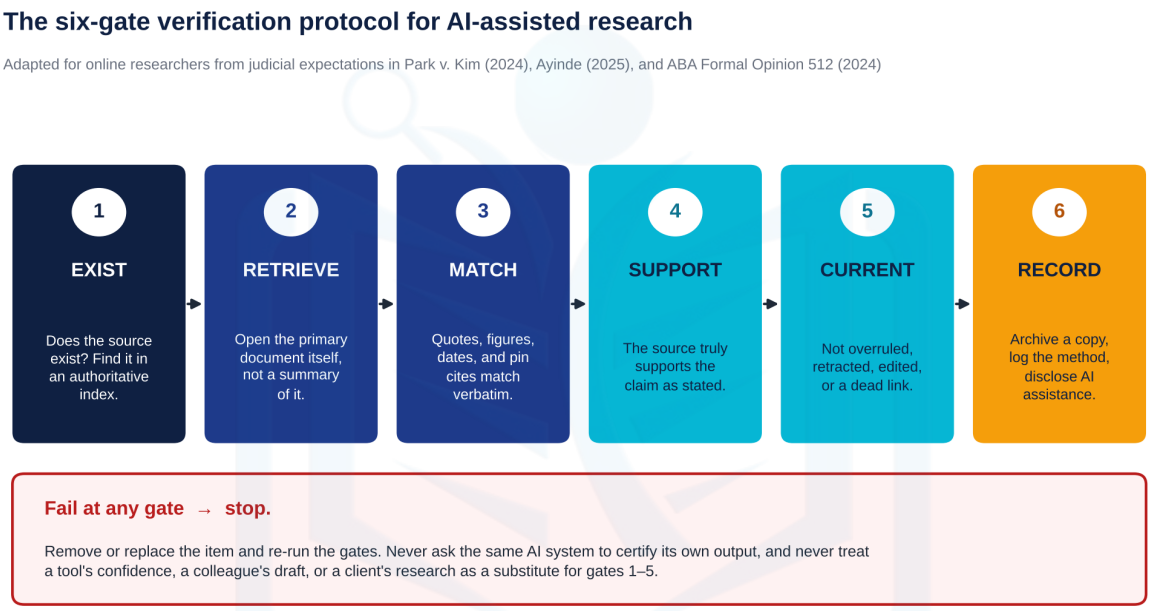
Courtroom lesson	What it means for online research
Candor after discovery (*Ayinde* factors)	Have a correction routine ready: acknowledge, withdraw, correct, and notify recipients

8.3 A Six-Gate Verification Protocol

Figure 11 turns the courts' expectations into a working protocol for AI-assisted research. Anything that will be cited, quoted, or relied on has to clear all six gates. If it fails one, it comes out or gets replaced, and the check starts over.

Figure 11

The Six-Gate Verification Protocol for AI-Assisted Research



Note. Developed for this report from judicial expectations in *Park v. Kim* (2024), *R (Ayinde) v London Borough of Haringey* (2025), and ABA Formal Opinion 512 (2024).

Gate 1: Exist

Find the source in an authoritative index that has nothing to do with the AI tool: a publisher or journal site, a DOI resolver, a library catalog, an official government or company registry, or a court database. A citation that looks right proves nothing about whether the source is real.

Gate 2: Retrieve

Open the document itself. An abstract, search snippet, AI overview, or someone else's description is not enough. If you cannot get hold of the document, you cannot call it verified; either say so plainly or find another source.

Gate 3: Match

Search the document for every quotation, number, date, and name you attribute to it, and make sure each one matches exactly. Note the page, paragraph, or timestamp. This is the gate that catches false quotations, which an existence check will miss.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

Gate 4: Support

Read enough of the surrounding text to be sure the source really supports the point you cite it for, and not a related or even opposite one. This gate deals with misrepresented authority, a mistake retrieval-based research tools are especially prone to.

Gate 5: Current

Make sure the source still stands. It should not have been retracted, corrected, overruled, or superseded, and a web page should still say what it said when you cited it. Link rot alone is a serious problem: one study found that more than 70% of URLs cited in law journals, and half of those in U.S. Supreme Court opinions, no longer led to the information originally cited (Zittrain et al., 2014).

Gate 6: Record

Save a copy of the source in a persistent archive or file, log how it was checked and by whom, and disclose material AI assistance in a methods note. That record lets someone else repeat the check, and it shows your diligence if the work is ever questioned.

Table 8

Pre-Delivery Verification Checklist for AI-Assisted Research Reports

☐	Check	Evidence it was done
☐	Every cited source found in an independent authoritative index (Gate 1)	Index or registry noted in the verification log
☐	Every source opened and read in full, or the relevant section read (Gate 2)	File or archived copy saved
☐	Every quotation, figure, and date matched word for word, with a pinpoint (Gate 3)	Page, paragraph, or timestamp recorded
☐	Each claim confirmed as supported by its source (Gate 4)	Reviewer initials next to each claim
☐	Retraction, correction, and currency checks done (Gate 5)	Date checked; archive link
☐	No AI output used to confirm other AI output	Methods note explains how sources were verified
☐	Material from clients or colleagues checked independently	Origin of each lead recorded
☐	Second-person review done for high-stakes work	Reviewer sign-off
☐	Significant AI assistance disclosed (Gate 6)	Disclosure statement included
☐	Correction routine ready in case an error turns up later	Named owner and notification list

8.4 Governance for Research Teams

Individual discipline only goes so far without support from the organization. The courts' experience suggests five measures:

- **Put the policy in writing.** One federal judge wrote an AI policy after an intern used a chatbot without permission (Reuters, 2025). Unwritten rules do not reliably reach new staff, interns, or contractors.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

- **Back the policy with review.** *Johnson v. Dunn* (2025) shows that an organization can have AI policies and still be sanctioned when supervisors do not check the work. Build verification into the sign-off process.
- **Use a second reviewer for high-stakes work.** Both judges whose staff had used AI responded by adding review before rulings go out (Reuters, 2025). For client-facing reports, an independent second reader does the same job.
- **Make someone in leadership accountable.** The Divisional Court made clear that it expects people with leadership responsibilities in chambers and firms to put practical safeguards in place (*R (Ayinde) v London Borough of Haringey*, 2025). Research organizations should name a person who owns their AI-verification standards.
- **Plan your response before you need it.** Decide ahead of time how you will acknowledge, withdraw, correct, and notify recipients when an error turns up. What you do in the first few hours largely decides how serious the consequences become.

8.5 Implications for Professional Training and Certification

Source evaluation and verification sit at the center of the Certified Internet Research Specialist (CIRS™) program run by the Association of Internet Research Specialists (AOFIRS). The court record makes a strong case that AI-specific verification should be taught and tested as a skill in its own right. That means recognizing the error types in Table 1, applying the six-gate protocol under time pressure, supervising agentic research workflows, auditing AI-generated bibliographies, and documenting verification well enough to stand up as evidence. Published sanctions decisions make excellent teaching material, because each one is a real, well-documented example of a specific failure. More information about the CIRS program is available at aofirs.org.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

9. The Regulatory Horizon

Rule-makers are settling on holding people accountable for results rather than banning particular technologies. From 2023 onward, individual judges issued standing orders requiring lawyers to certify either that they had not used generative AI or that a human had checked any AI-generated text (for example, Starr, 2023). Some appeals courts considered similar rules but decided, as the Fifth Circuit did, that existing duties were enough, and reminded lawyers instead that they must check their filings for accuracy (*Fletcher v. Experian Information Solutions, Inc.*, 2026).

On the evidence side, federal rule-makers released proposed Federal Rule of Evidence 707 for public comment from August 2025 to February 2026. It would require machine-generated evidence offered without an expert witness to meet the reliability standards that apply to expert testimony. Public comment was mixed, and in 2026 the proposal was put on hold for further study and consultation with technical experts (Proskauer Rose LLP, 2026). In England and Wales, judicial guidance on AI, first issued in December 2023 and updated since, makes legal representatives responsible for whatever they put before the court, and the Divisional Court urged professional bodies and leaders to take further practical steps without delay (*R (Ayinde) v London Borough of Haringey*, 2025). In the United States, bar regulators continue to issue formal opinions that build on ABA Formal Opinion 512.

The verification tools are improving too. Automated reference checkers and URL-health tools can flag suspect citations in bulk, and controlled experiments show large drops in error when models have to check against outside records (Rao et al., 2026). These tools make a good first filter. The case law is clear, though, that they support human verification rather than replace it.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

10. Conclusion

Three years of courtroom experience lead to three conclusions, and none of them is limited to law. The first is that the duty to verify is personal and cannot be handed off. A paralegal, co-counsel, client, intern, vendor, or AI agent can help, but responsibility stays with the person who presents the work as reliable. The second is that verification that depends on AI does not count. Courts have treated asking a model to confirm its own output, reading an AI summary in place of the source, and trusting a research-grade label as no different from not checking at all. The third is that what happens after an error is found matters as much as the error. Professionals who disclose quickly, correct fully, and fix their processes are dealt with firmly but fairly. Those who deny, hide, or play down the problem face sanctions that now run to six figures and can cost them the right to practice.

For online researchers, the courts' warning doubles as a practical guide. AI can make research faster and broader, and agentic tools will keep getting better. None of that changes the basic deal between researchers and the people who rely on them: every source cited has been found, read, matched, understood, checked for currency, and recorded. The six-gate protocol is one way to keep that promise. The habit underneath it, reading before relying, is what the sanctions record teaches most clearly.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org

References

American Bar Association Standing Committee on Ethics and Professional Responsibility. (2024, July 29). *Formal opinion 512: Generative artificial intelligence tools*. American Bar Association. https://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/ethics-opinions/aba-formal-opinion-512.pdf

Business Guides, Inc. v. Chromatic Communications Enterprises, Inc., 498 U.S. 533 (1991).

ByoPlanet International, LLC v. Johansson, 792 F. Supp. 3d 1341 (S.D. Fla. 2025).

Charlotin, D. (2026). *AI hallucination cases database* [Data set; snapshot updated September 28, 2026]. Licensed under CC BY 4.0. Retrieved September 29, 2026, from https://www.damiencharlotin.com/hallucinations/

Chelli, M., Descamps, J., Lavoué, V., Trojani, C., Azar, M., Deckert, M., Raynier, J.-L., Clowez, G., Boileau, P., & Ruetsch-Chelli, C. (2024). Hallucination rates and reference accuracy of ChatGPT and Bard for systematic reviews: Comparative analysis. *Journal of Medical Internet Research, 26*, Article e53164. https://doi.org/10.2196/53164

Coomer v. Lindell, No. 22-cv-01129 (D. Colo. July 7, 2025).

Couvrette v. Wisnovsky, No. 1:21-cv-00157-CL, 2025 WL 4109655 (D. Or. Dec. 12, 2025).

Couvrette v. Wisnovsky, No. 1:21-cv-00157-CL, 2026 WL 870387 (D. Or. Mar. 30, 2026).

Dahl, M., Magesh, V., Suzgun, M., & Ho, D. E. (2024). Large legal fictions: Profiling legal hallucinations in large language models. *Journal of Legal Analysis, 16*(1), 64–93. https://doi.org/10.1093/jla/laae003

Federal judge sanctions three Butler Snow attorneys over AI-generated hallucination in defense of ADOC. (2025, July). *1819 News*. https://1819news.com/news/item/federal-judge-sanctions-three-butler-snow-attorneys-over-ai-generated-hallucination-in-defense-of-adoc

Fletcher v. Experian Information Solutions, Inc., No. 25-20086 (5th Cir. Feb. 18, 2026). https://www.ca5.uscourts.gov/opinions/pub/25/25-20086-CV0.pdf

Gauthier v. Goodyear Tire & Rubber Co., 2024 WL 4882651 (E.D. Tex. Nov. 25, 2024).

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association, 19*(1), 121–127. https://doi.org/10.1136/amiajnl-2011-000089

Heimkes v. Fairhope Motorcoach Resort Condominium Owners Ass'n, No. 22-cv-00448, 2026 WL 883120 (S.D. Ala. Mar. 31, 2026).

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), 1–38. https://doi.org/10.1145/3571730

Johnson v. Dunn, 792 F. Supp. 3d 1241 (N.D. Ala. 2025).

Jordan v. Chicago Housing Authority, No. 2022 L 000095 (Ill. Cir. Ct. Cook Cnty. Dec. 5, 2025).

Kalai, A. T., Nachum, O., Vempala, S. S., & Zhang, E. (2025). *Why language models hallucinate* (arXiv:2509.04664) [Preprint]. arXiv. https://arxiv.org/abs/2509.04664

Ko v. Li, 2025 ONSC 2766 (Ontario Superior Court of Justice).

Kohls v. Ellison, No. 24-cv-3754 (D. Minn. Jan. 10, 2025).

Lacey v. State Farm General Insurance Co., No. CV 24-5205 FMO (MAAx) (C.D. Cal. May 6, 2025).

Liu, N. F., Zhang, T., & Liang, P. (2023). Evaluating verifiability in generative search engines. In *Findings of the Association for Computational Linguistics: EMNLP 2023* (pp. 7001–7025). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-emnlp.467

Los Angeles County Bar Association Professional Responsibility and Ethics Committee. (2006). *Formal opinion no. 518: Ethical considerations in outsourcing of legal services*.

Magesh, V., Surani, F., Dahl, M., Suzgun, M., Manning, C. D., & Ho, D. E. (2025). Hallucination-free? Assessing the reliability of leading AI legal research tools. *Journal of Empirical Legal Studies, 22*(2), 216–242. https://doi.org/10.1111/jels.12413

Manach, J.-M. (2025, July 1). Il génère des plaidoiries par IA, et en recense 160 ayant « halluciné » depuis 2023 [He generates pleadings with AI and lists 160 that have "hallucinated" since 2023]. *Next*. https://next.ink/189552/il-genere-des-plaidoiries-par-ia-et-en-recense-160-ayant-hallucine-depuis-2023/

Martinez, R., & Frost, A. (2026, May 10). AI fabrications in legal filings grow in Oregon, US. *Oregon Public Broadcasting*. https://www.opb.org/article/2026/05/10/ai-oregon-law-fake-lawyer-technology-hallucinations/

Mata v. Avianca, Inc., 678 F. Supp. 3d 443 (S.D.N.Y. 2023).

New York State Bar Association Committee on Professional Ethics. (1999). *Opinion 721: Acceding to carrier's instruction to use named legal research firm*.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors, 52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Park v. Kim, 91 F.4th 610 (2d Cir. 2024).

Pavelic & LeFlore v. Marvel Entertainment Group, 493 U.S. 120 (1989).

Proskauer Rose LLP. (2026, July 29). *Machine-generated evidence challenges the Federal Rules: Inside the contested proposed Federal Rule of Evidence 707*. https://www.proskauer.com/blog/machine-generated-evidence-challenges-the-federal-rules-inside-the-contested-proposed-federal-rule-of-evidence-707

R (Ayinde) v London Borough of Haringey; Al-Haroun v Qatar National Bank QPSC [2025] EWHC 1383 (Admin). https://www.judiciary.uk/judgments/ayinde-v-london-borough-of-haringey-and-al-haroun-v-qatar-national-bank/

Rao, D., Wong, E., & Callison-Burch, C. (2026). *Detecting and correcting reference hallucinations in commercial LLMs and deep research agents* (arXiv:2604.03173) [Preprint]. arXiv. https://arxiv.org/abs/2604.03173

Reuters. (2025, October 23). Two federal judges say use of AI led to errors in US court rulings. *Business Insurance*. https://www.businessinsurance.com/two-federal-judges-say-use-of-ai-led-to-errors-in-us-court-rulings/

Rivera v. Triad Properties Corp., No. 24-cv-01802, 2026 WL 915744 (N.D. Ala. Mar. 31, 2026).

Robert, A. (2026, April 17). Federal judge hands down $110K penalty against 2 lawyers for AI errors in court documents. *ABA Journal*. https://www.abajournal.com/news/article/oregon-federal-judge-hands-down-110000-penalty-for-ai-errors

Savino, C. (2026). Fake cases, real sanctions: The dangers of AI. *Litigation News*. American Bar Association Section of Litigation. https://www.americanbar.org/groups/litigation/resources/litigation-news/2026/fake-cases-real-sanctions-dangers-ai/

Starr, B. (2023, May 30). *Mandatory certification regarding generative artificial intelligence* [Judge-specific requirement]. U.S. District Court for the Northern District of Texas.

University of Illinois Chicago School of Law Library. (2026). *Flood of AI-related sanctions persist*. https://library.law.uic.edu/news-stories/flood-of-ai-related-sanctions-persist/

Wadsworth v. Walmart Inc., 348 F.R.D. 489 (D. Wyo. 2025).

Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. *Scientific Reports, 13*, Article 14045. https://doi.org/10.1038/s41598-023-41032-5

Whiting v. City of Athens, 170 F.4th 455 (6th Cir. 2026).

WOWT. (2026, April 16). *Nebraska Supreme Court suspends Omaha attorney over AI use*. https://wowt.com/2026/04/16/nebraska-supreme-court-suspends-omaha-attorney-over-ai-use/

Zhang v. Chen, 2024 BCSC 285 (Supreme Court of British Columbia).

Zittrain, J., Albert, K., & Lessig, L. (2014). Perma: Scoping and addressing the problem of link and reference rot in legal citations. *Harvard Law Review Forum, 127*, 176–199.

Note on legal authorities: Following APA 7th edition style, court decisions are listed alphabetically by case name, with case names italicized in the reference list and in text. Secondary citations (Xu et al., 2026; Sakai et al., 2026) are cited as reported in Rao et al. (2026), which is listed above.

Appendix A. Methodological Note on the Case Data

The figures in Section 3 come from the public AI Hallucination Cases Database maintained by Damien Charlotin. I accessed it on 29 September 2026; it had last been updated on 28 September 2026. The database records legal decisions in which a court or tribunal found, or implied, that a party relied on hallucinated content. It does not record mere allegations. It is licensed under Creative Commons Attribution 4.0 and is used here with attribution.

Counts by party, type of hallucination, subcategory, jurisdiction, and area of law come from the database's filter summaries for the full data set (N = 2,095). The categories for the nature of the hallucination overlap, because one decision can record several kinds of error. The "all other jurisdictions" value in Figure 6 is what remains after subtracting the six largest jurisdictions. The July 2025 figure of 160 decisions is the count reported in media coverage of the database at the time (Manach, 2025). The federal-docket counts for 2025 and the first quarter of 2026 are those reported by the University of Illinois Chicago School of Law Library (2026). Dollar amounts in Figure 7 and Table 2 come from the decisions or from reliable reporting on them and include fee awards where reported, so they show total financial exposure rather than punitive fines alone.

Every count here is a minimum. Many hallucinated filings are never caught, many that are caught are dealt with without a written decision, and coverage of lower courts and non-English decisions is incomplete. The database keeps growing, so check it directly for current figures.

© 2026 Association of Internet Research Specialists (AOFIRS) · aofirs.org