

January 22, 2026

MCLE Blitz

AI Hallucinations: Policies & Practice

TABLE OF CONTENTS

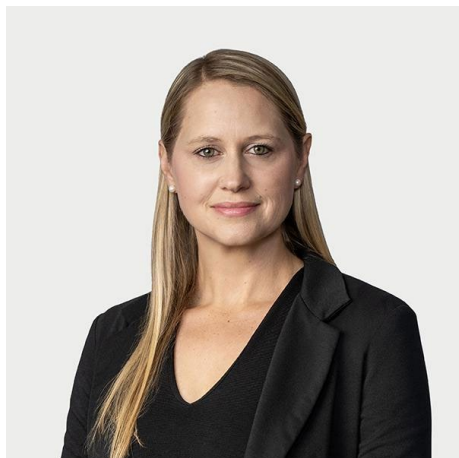
- 01** Background & Technical Framing
- 02** Risks: Where Things Go Wrong in Practice
- 03** Emerging Rules, Enforcement & Sanctions
- 04** Best Practices & Mitigation

Presenters

Abbey A. Barrera

Partner

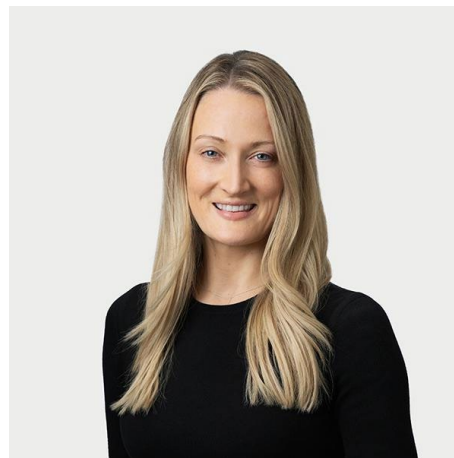
San Francisco



Frances Waldmann

Partner

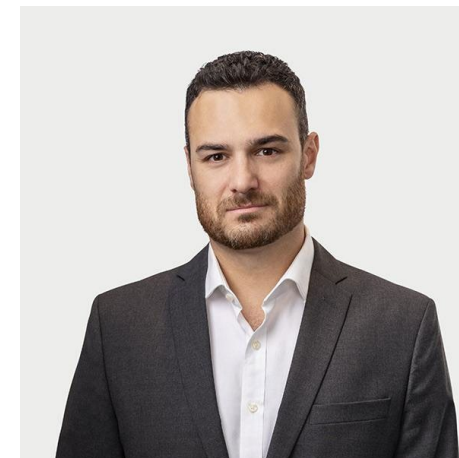
Los Angeles/London



Tim Biché

Associate

Los Angeles



Background & Technical Framing

01

What is a **Hallucination**?

When a generative AI model confidently generates **an answer that isn't true**.

When a language model produces a **plausible yet incorrect statement** instead of **admitting uncertainty**.



Key Definitions

Artificial Intelligence

Capability of a computer system to mimic human cognitive functions such as learning and problem-solving by using math and logic to simulate human reasoning and decision-making.

Machine Learning

Subset of AI that employs advanced statistical methods to enable systems to learn from data, identify patterns, and make predictions without explicit human programming.

Neural Network

Subset of machine learning composed of interconnected layers of mathematical “neurons” that process and transform data to recognize complex patterns and relationships.

Language Model

Type of neural network trained on large volumes of text data to learn language patterns and generate or predict human-like text.

Rapidly Accelerating AI Progress

- **Rapid Advancements:** Frontier models now exceed human benchmarks in reasoning, language, coding, and vision.
- **Accelerated Development:** The pace of AI innovation has advanced quickly, as model development timelines compress with rapid iterations.
- **Scaling as a Driver:** Breakthroughs are increasingly propelled by larger training datasets, specialized GPUs, and optimized inference pipelines.
- **Evolving Focus:** The emphasis is shifting from raw accuracy to the quality of reasoning, robustness, and domain alignment.
- **Policy Shift:** The discourse around AI is moving from "AI safety" towards "innovation and competitiveness."

Predictive AI



Uses statistical algorithms to analyze data and make predictions about future events based on the historical data (e.g., forecasting outcomes, classifying events, generating actionable insights).

Generative AI



Relies on neural network techniques to autonomously generate new content, data, or outputs that mimic or resemble human-created content.

Why Do Hallucinations Occur?

- Language models **predict** the next most likely token
- Outputs are generated **probabilistically**, not retrieved
- Models lack an internal concept of **truth**
- **Confidence** is optimized over accuracy

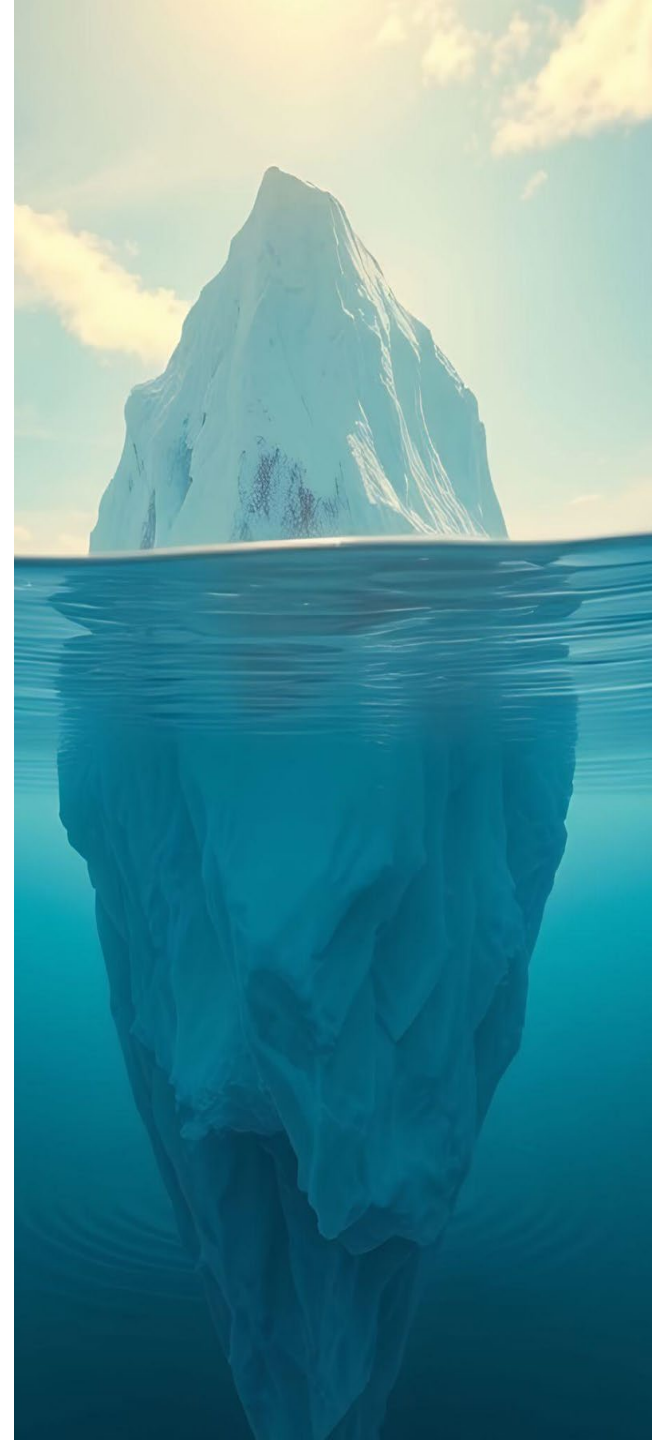
Risks:

Where Things Go Wrong in Practice

02

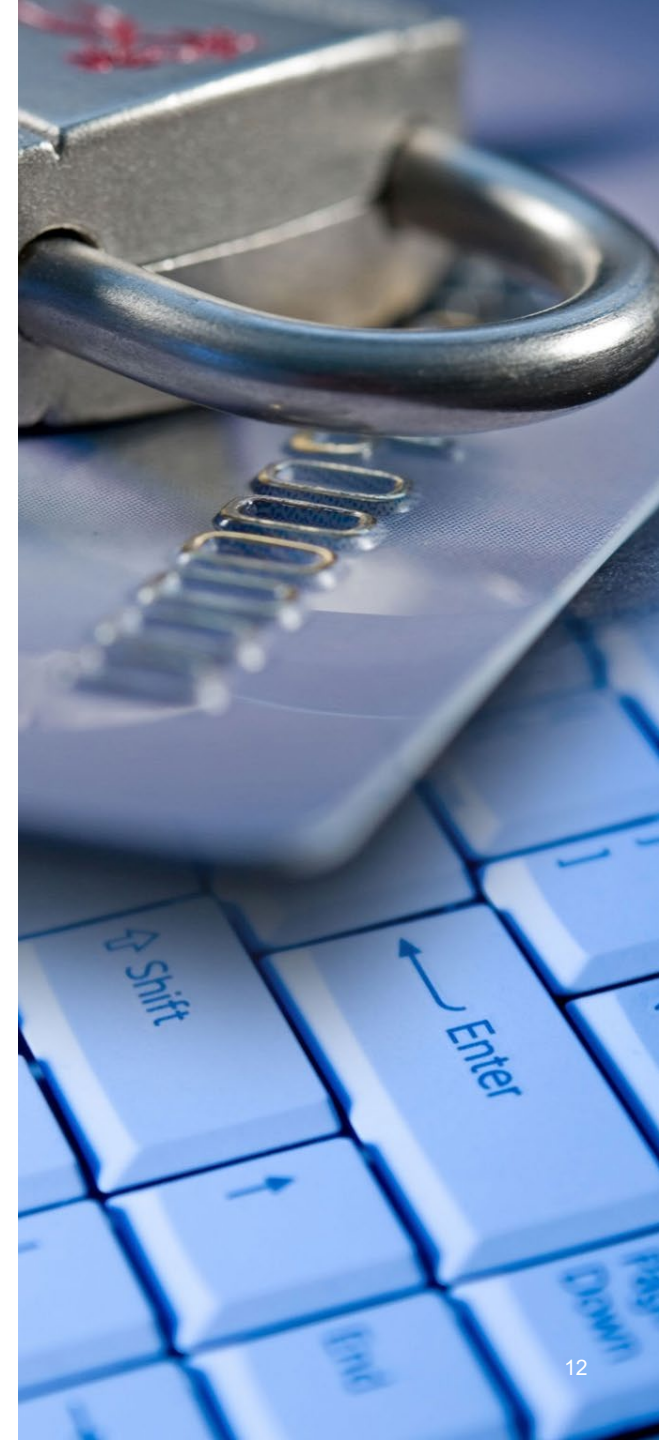
AI Use Risks: Transparency

- Lack of insight into how AI systems are trained, learn, and make decisions
- The more complex and independent an AI system is, the more risk it creates
- AI systems can be:
 - **Deterministic systems**, which involve clear logic and specified goals, limited data flow, and are relatively simple to validate, but still create some baseline employment and privacy law risks.
 - **Autonomous/agentic AI systems**, which exhibit more agency, autonomy, capability, and/or generality and may be connected to “scaffolding” software or tools, multiplying those baseline risks. These systems are typically directly regulated by new and emerging AI laws, more opaque and difficult to explain or reverse-engineer, and more likely to result in algorithmic bias.



AI Use Risks: Confidentiality & Privacy

- **Confidential data**
 - Contractual requirements for handling of confidential data
 - Potential liability if confidential data in your care is exposed through use of AI tools
 - Employees may use unapproved AI tools and share sensitive data
- **Personal data**
 - Contractual requirements, including as required by state privacy laws
 - Regulatory risk, including if personal data uses are not clearly disclosed
- **Employee monitoring & surveillance**
 - Use of AI monitoring or recording tools



AI Use Risks: **Privilege & Litigation Holds**

- **Privilege**

- Using AI can inadvertently waive attorney-client privilege
- AI tools may be treated as another participant in the conversation
- Factors to consider:
 - Is AI being used to support a legal purpose (e.g., to analyze legal documents or create summaries of privileged meetings or documents at the direction of counsel)?
 - Are appropriate confidentiality protections in place (e.g., access restrictions)?

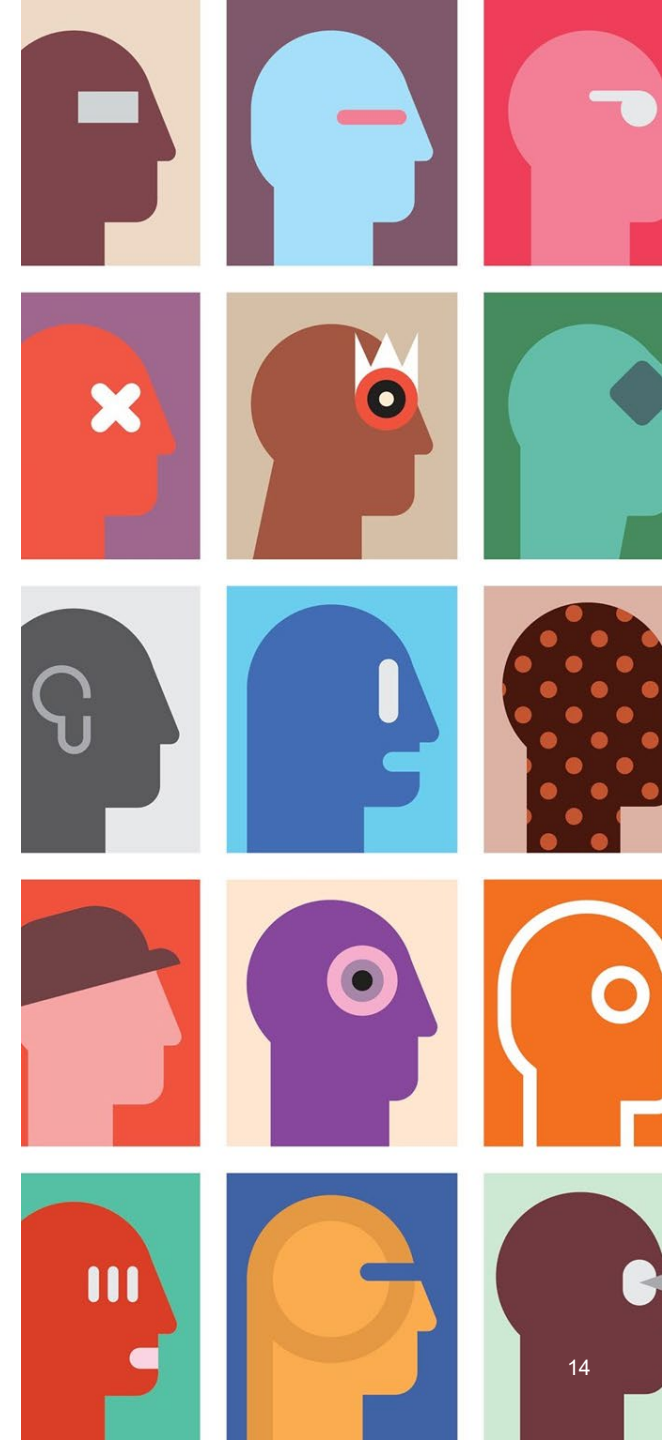
- **Litigation Holds**

- Litigants are demanding that AI-generated transcripts or notes be retained



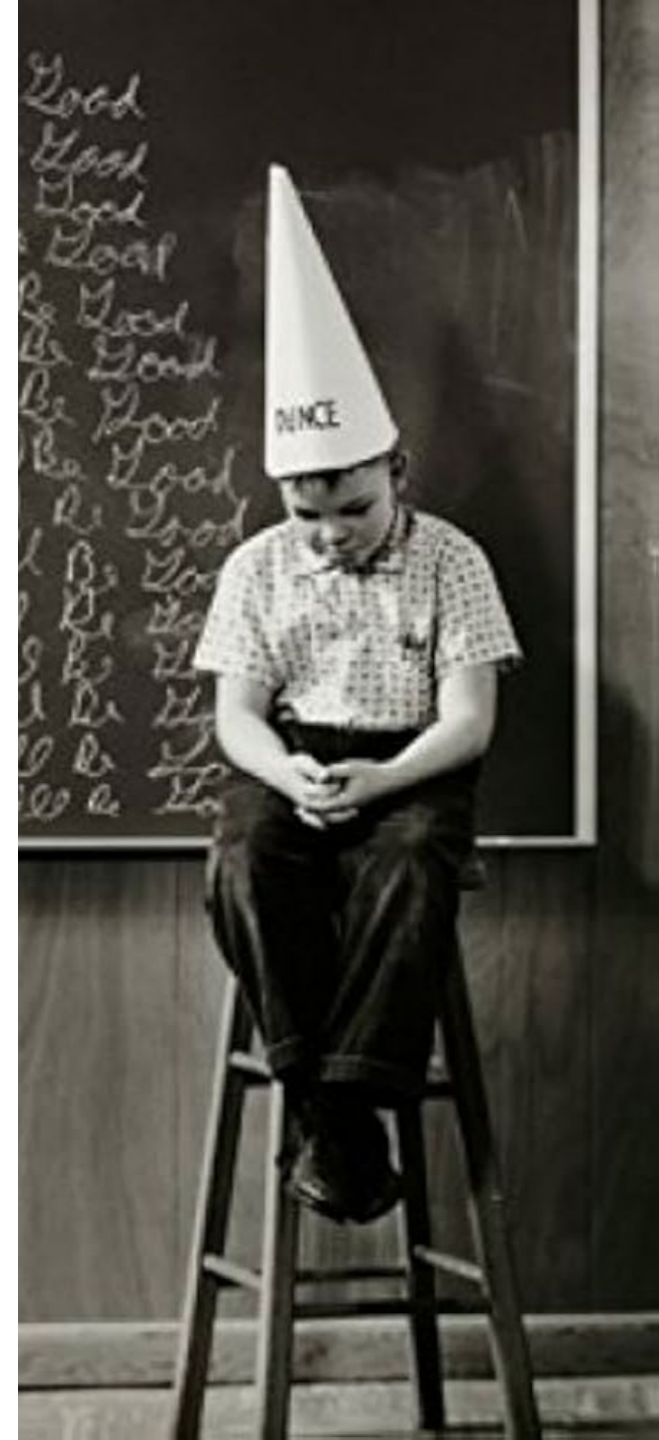
AI Use Risks: Bias and Discrimination

- **Several causes can contribute to bias in AI systems**
 - **Historical Data:** underlying data used to train AI systems may contain biases
 - **Feedback Loops:** May lead to biased outcomes that reinforce existing stereotypes
 - **Design Flaws:** Use of inaccurate, biased, manipulated, or incomplete data in an algorithm can lead to inaccuracies or replicated bias in the output
- **Lack of transparency about inputs used to train the model may impact reliability and exacerbate risk**



AI Use Risks: Hallucinations in Court

- **Hallucinations can show up in briefs prepared with AI tools**
 - Possible that AI makes up cases, citations, or quotations
 - Real case names with fictional holdings
 - AI citations to or quotations of real cases might misrepresent the law
- **Common Hallucination Triggers**
 - Niche or recent case law
 - Requests for quotes or pincites
 - Overly broad or underspecified prompts
- **Potential for sanctions for unreliable AI**
 - Monetary sanctions
 - Bar referral



GenAI in Legal Practice: **New Opportunities & Risks**

Capability and Modality

Models now reason across entire documents, datasets, and web content

Legal Tech Tools / Broad Adoption

Rapid proliferation of AI tools purpose-built for legal work

Perception

Outputs appear fluent and authoritative even when incorrect

Expectations

Courts now expect AI literacy; hallucinations are treated as foreseeable, not novel

Emerging Rules, Enforcement & Sanctions

03

California issues historic fine over lawyer's ChatGPT fabrications

NY judge sanctions lawyer for fake AI citations

Miss. Attys Sanctioned Over AI Misuse In Age Bias Case

AI hallucinated made-up citations.
Kansas judge may sanction lawyers

Fla. Atty Faces Bar Referral Over 'Hallucinated' Case In Filing

'Just wrong': Baldwin County lawyer fined, reprimanded by federal judge after using AI to create draft court filings



Emerging Rules & Enforcement Landscape

Trends:

- Patchwork of specific rules—many require disclosure of AI use and certification of human verification
- Courts increasingly scrutinize AI-generated filings
- Sanctions for fabricated citations and misrepresentations
- Judges emphasizing lawyer—not tool—responsibility

Risks Beyond Sanctions:

- Media scrutiny
- Client confidence erosion
- Long-term reputational damage

Emerging Rules & Guidance

Key Sources:

- ABA Formal Opinion 512
- California Bar Practical Guidance on Generative AI
- California Rule 10.430
- Local rules and judge-specific standing orders

Emerging Rules & Guidance:

ABA Formal Opinion 512

Focus on generative AI in the practice of law and emphasis on **existing ethical obligations** and model rules, including:

- Model Rule 1.1 (Competence)
- Model Rule 1.6 (Confidentiality of Information)
- Model Rule 1.4 (Communications)
- Model Rule 1.5 (Fees)

AMERICAN BAR ASSOCIATION

STANDING COMMITTEE ON ETHICS AND PROFESSIONAL RESPONSIBILITY

Formal Opinion 512

July 29, 2024

Generative Artificial Intelligence Tools

To ensure clients are protected, lawyers using generative artificial intelligence tools must fully consider their applicable ethical obligations, including their duties to provide competent legal representation, to protect client information, to communicate with clients, to supervise their employees and agents, to advance only meritorious claims and contentions, to ensure candor toward the tribunal, and to charge reasonable fees.

Emerging Rules & Guidance:

California Bar Practical Guidance

Similar focus on **existing obligations**, including those set forth in the Rules of Professional Conduct and the State Bar Act:

- Duty of Confidentiality
- Duties of Competence and Diligence
- Duty to Comply with the Law
- Duty to Supervise Lawyers and Nonlawyers, Responsibility of Subordinate Lawyers
- Communication Regarding Generative AI Use
- Charging for Work Produced By Generative AI and Generative AI Costs
- Candor to the Tribunal; and Meritorious Claims and Contentions
- Prohibition on Discrimination, Harassment, and Retaliation
- Professional Responsibilities Owed to Other Jurisdictions

Emerging Rules & Guidance:

California Rules of Court Rule 10.430

- Applies to **California courts** (including superior courts, Courts of Appeal, and the Supreme Court)
- Required any court that does not prohibit the use of generative AI by court staff or judicial officers to adopt a generative AI use policy by December 15, 2025
- Policy requirements focus on:
 - Protecting confidentiality
 - Preventing discrimination
 - Verifying content
 - Removing biased, offensive, or harmful content
 - Disclosing use of AI if final work product provided to the public “consists entirely of generative AI outputs”
 - Complying with “all applicable laws, court policies, and ethical and professional conduct rules, codes, and policies”

Emerging Rules & Guidance: Judge-Specific Requirements

h. Artificial Intelligence

Any party who uses generative artificial intelligence (such as ChatGPT, Harvey, CoCounsel, or Google Bard) to generate any portion of a motion, brief, pleading, or other filing must attach to the filing a separate declaration disclosing the use of artificial intelligence and certifying that the filer has reviewed the source material and verified that the artificially generated content is accurate and complies with the filer's Rule 11 obligations.

Hon. Fred Slaughter (C.D. Cal.)

4. Artificial Intelligence (AI). Counsel is responsible for providing the Court with complete and accurate representations in any submission (including filings, demonstratives, evidence, or oral argument), consistent with Federal Rule of Civil Procedure 11, the California Rules of Professional Conduct, and any other applicable legal or ethical guidance. Use of ChatGPT or other such tools is not prohibited, but counsel must at all times personally confirm for themselves the accuracy of any content generated by these tools. At all times, counsel—and specifically designated lead trial counsel—bears responsibility for any submission made by the party that the attorney represents. Any submission containing AI-generated content must include a certification that lead trial counsel has personally verified the content's accuracy. Failure to include this certification or comply with this verification requirement will be grounds for sanctions. Counsel is responsible for maintaining records of all prompts or inquiries submitted to any generative AI tools in the event those records become relevant at any point.

Hon. Araceli Martínez-Olguín (N.D. Cal.)

Emerging Rules & Guidance:

Court-Specific Requirements

Any attorney or self-represented litigant who signs a pleading, written motion, or other paper submitted to the Court will be held responsible for the contents of that filing under Rule 11, regardless of whether generative artificial intelligence drafted any portion of that filing. *See* Fed. R. Civ. P. 11(c) (providing for imposition of an “appropriate sanction”—including nonmonetary directives, a penalty payable to the court, or payment to the opposing party of attorney's fees and expenses directly resulting from the violation—if, after notice and a reasonable opportunity to respond, the Court determines that Rule 11(b) has been violated).

General Order, U.S. Dist. Ct., S.D. Tex.

Therefore, all parties are on notice that the Court has a no-tolerance policy for any briefing (AI-assisted or not) that hallucinates legal propositions or otherwise severely misstates the law. Such filings will often result in sanctions absent reasonable excuse. *See generally Willis v. U.S. Bank Nat'l Ass'n et al*, No. 3:25-CV-516-BN, 2025 WL 1408897 (N.D. Tex. May 15, 2025).

AI Notice, U.S. Dist. Ct., Conn.

Sanctions & Enforcement:

Case Examples

California Court of Appeal issued **\$10,000 in sanctions** where plaintiff's briefs were replete with fabricated legal authorities created by generative AI:

Although the generation of fake legal authority by AI sources has been widely commented on by federal and out-of-state courts and reported by many media sources, no California court has addressed this issue. We therefore publish this opinion as a warning. Simply stated, no brief, pleading, motion, or any other paper filed in any court should contain *any* citations—whether provided by generative AI or any other source—that the attorney responsible for submitting the pleading has not personally read and verified. Because plaintiff's counsel's conduct in this case violated a basic duty counsel owed to his client and the court, we impose a monetary sanction on counsel, direct him to serve a copy of this opinion on his client, and direct the clerk of the court to serve a copy of this opinion on the State Bar.

Noland v. Land of the Free, L.P., 114 Cal. App. 5th 426, 449 (2025)

Sanctions & Enforcement:

Case Examples

Special Master imposed **discovery sanctions and awarded costs** where attorneys submitted briefs “that contained bogus AI-generated research”:

A final note. Directly put, Plaintiff's use of AI affirmatively misled me. I read their brief, was persuaded (or at least intrigued) by the authorities that they cited, and looked up the decisions to learn more about them – only to find that they didn't exist. That's scary. It almost led to the scarier outcome (from my perspective) of including those bogus materials in a judicial order. Strong deterrence is needed to make sure that attorneys don't succumb to this easy shortcut.

For these reasons, Plaintiff's supplemental briefs are struck, and no further discovery relief will be granted on the disputed privilege issue. Additionally, Plaintiff's law firms are ordered (jointly and severally) to pay compensation to the defense in the aggregate amount of \$31,100.

Lacey v. State Farm Gen. Ins. Co., 2025 WL 1363069, at *5 (C.D. Cal. May 5, 2025)

Sanctions & Enforcement:

Case Examples

Magistrate Judge imposed **personal sanctions of \$1,500** where counsel submitted fictitious case and Court found misrepresentations were made knowingly with intent to mislead, demonstrating bad faith:

For the reasons stated above, the Court orders the following sanctions for violation of Local [Rule 180\(e\)](#) and pursuant to its inherent authority:

1. Assistant Federal Defender Andrew Francisco, defense counsel, is personally sanctioned in the amount of \$1,500. Within 21 days of the date of this Order, Mr. Francisco shall pay these sanctions to the Clerk of Court. The case number and a copy of this Order should be included with payment.
2. The Court orders the Clerk of Court to serve a copy of this order on the District of Columbia Bar, of which Mr. Francisco is a member (DC Bar No. 1619332), and the State Bar of California.
3. The Court orders the Clerk of Court to serve a copy of this order on all the district judges and magistrate judges in this district.

United States v. Hayes, 763 F. Supp. 3d 1054, 1073 (E.D. Cal. 2025)

Best Practices & Mitigation

04

Best Practices: Overview

Practical Verification Techniques

- Require citations with links or source snippets
- Use precise prompts
- Independently confirm
- Use secondary sources for cross-checking
- Treat unfamiliar authority as suspect

Tool Configuration & Controls

- Enterprise tenants
- No training on user data
- Audit logs and retention controls
- Restricted inputs for confidential material

Policies & Governance

- Approved vs. prohibited tools
- Permitted use cases
- Mandatory review requirements
- Disclosure and escalation protocols
- Training and supervision
- Retention of AI-generated drafts
- Transcripts and meeting summaries
- Audit trails for decision-making

When Should You Disclose

- AI use materially shaped arguments
- Required by court order or local rule
- Necessary to avoid misleading the tribunal or client

Future-Proofing

- **How to set standards today that withstand regulatory evolution**
- Be flexible & build room for policy modifications as laws evolve

Key Mitigation: **Human-In-The-Loop**

- Having a “human-in-the-loop” is not a complete defense or necessarily effective
 - **Rubber-stamping complex decisions or lengthy outputs without adequate review is not enough**
 - Meaningful review requires authority, training/subject-matter competence, and time



Key Takeaways

- Hallucinations are a **predictable consequence** of model design
- The risk is **foreseeable and preventable**
- Courts expect lawyers to **understand** these mechanics
- **Verification, governance, and training** are essential

Questions?

Abbey A. Barrera



415.393.8262

abarrera@gibsondunn.com

Frances Waldmann



213.229.7914

fwaldmann@gibsondunn.com

Tim Biché



213.229.7434

tbiche@gibsondunn.com

GIBSON DUNN

Attorney Advertising: These materials were prepared for general informational purposes only based on information available at the time of publication and are not intended as, do not constitute, and should not be relied upon as, legal advice or a legal opinion on any specific facts or circumstances. Gibson Dunn (and its affiliates, attorneys, and employees) shall not have any liability in connection with any use of these materials. The sharing of these materials does not establish an attorney-client relationship with the recipient and should not be relied upon as an alternative for advice from qualified counsel. Please note that facts and circumstances may vary, and prior results do not guarantee a similar outcome. © 2025 Gibson, Dunn & Crutcher LLP. All rights reserved. For contact and other information, please visit us at [gibsondunn.com](https://www.gibsondunn.com).