

# Sanctions for AI Misuse Among Lawyers and Judges

James I. Jackson\*

## Abstract

*The use of artificial intelligence (AI) in legal work offers clear advantages but also serious risks. This paper reviews how lawyers and judges in the United States have been disciplined for misusing AI and how existing professional standards are being tested by new technology. It examines how U.S. courts and disciplinary authorities have responded to AI misuse and compares those responses to international efforts, particularly the Chartered Institute of Arbitrators Guideline on the Use of AI in Arbitration (2025). The paper argues that current U.S. approaches are reactive and fragmented and advocates for proactive reform, including clearer rules that define responsible use, required disclosure of AI involvement, and expanded professional education on the subject. To address these concerns, this paper recommends reforms aimed at ensuring that the use of AI in legal practice supports, rather than weakens, the integrity of the legal system.*

**Keywords:** AI misuse; professional discipline; ethical duties; hallucinated citations; regulatory reform

## A Introduction

Artificial intelligence (AI) is reshaping the practice of law.<sup>1</sup> Tools capable of searching large data sets, drafting text, or managing case materials can increase efficiency, but they also introduce significant risks.<sup>2</sup> Lawyers and judges have begun relying on systems that sometimes generate fabricated citations, expose confidential information, or embed hidden biases.<sup>3,4,5</sup> As these tools become more common, the central question is not whether the legal profession will use AI, but how its use should be governed.

\* James I. Jackson, Indiana University Robert H. McKinney School of Law, Indianapolis.

1 Alarie, 'The Path of the Law: Towards Legal Singularity', 66 *U. Toronto L.J.* (2016) 1, at 3-7.

2 Coglianese and Lehr, 'Regulating by Robot: Administrative Decision Making in the Machine-Learning Era', 105 *Geo. L.J.* (2017) 1145, at 1147-1152.

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

4 Lieberknecht, 'From Innovation to Exposure: Artificial Intelligence Risks for Legal Professionals', *Reuters* (2025).

5 Ferrer, 'Bias and Discrimination in AI: A Cross-Disciplinary Perspective', *IEEE Technology and Society* (2021), available at <https://technologyandsociety.org/bias-and-discrimination-in-ai-a-cross-disciplinary-perspective/>.

Using AI is not improper on its own. Problems arise when lawyers or judges rely on it without understanding its limits, fail to review what it produces, or conceal its involvement. Such conduct can violate long-standing professional duties of competence, diligence, and candor under the American Bar Association (ABA) Model Rules of Professional Conduct<sup>6</sup> and, in litigation, may also violate the certification and reasonable-inquiry requirements of Rule 11 of the Federal Rules of Civil Procedure (FRCP).<sup>7</sup> Recent disciplinary cases show that unverified AI-generated material can harm clients, burden courts, and undermine public trust in the justice system.<sup>8,9,10</sup>

This article examines how United States (U.S.) courts and disciplinary authorities have responded to AI misuse and compares those responses to international efforts, particularly the Chartered Institute of Arbitrators Guideline on the Use of AI in Arbitration (2025).<sup>11</sup> While the U.S. continues to rely on traditional ethical principles, newer global standards point toward clearer requirements for verification, transparency, and accountability. The paper concludes by recommending reforms aimed at ensuring that the use of AI in legal practice supports, rather than weakens, the integrity of the legal system. To understand why AI misuse is difficult to regulate under existing frameworks, the next section identifies the most common technological and ethical pitfalls that arise when lawyers and judges rely on AI tools.

## B Common AI Pitfalls Affecting Lawyers and Judges

As lawyers and judges increasingly incorporate AI tools into their work, several recurring risks have become apparent. These pitfalls stem not from intentional misuse but from the limitations of current AI systems and the challenges of supervising tools that operate through non-transparent processes.<sup>12</sup> Together, these risks underscore the need for guidance that goes beyond existing standing orders or general ethics rules. As AI becomes more integrated into legal workflows, clearer standards will be necessary to ensure that its use remains transparent, limited, and always subject to human judgment.

### I *Hallucination and Fabrication*

One of the most common dangers associated with AI is the risk of inaccurate or fabricated information. Generative AI systems may produce text that sounds authoritative but has no basis in actual legal sources. These “hallucinations” include invented cases, quotations, and procedural histories that appear credible but are

6 Model Rules of Pro. Conduct (Am. Bar Ass’n 2025).

7 Fed. R. Civ. P. 11.

8 *Mata*, supra note 3.

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

10 *People v. Crabill*, No. 23PDJ067 (Colo. 2023).

11 Chartered Institute of Arbitrators, *Guideline on the Use of AI in Arbitration* (2025).

12 Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), OJ 2024 L, 2024/1689, available at <https://artificialintelligenceact.eu/>.

entirely false.<sup>13</sup> Such errors occur because large language models generate plausible-sounding text without verifying accuracy against authoritative sources. Hallucination and fabrication errors have appeared in multiple disciplinary cases involving lawyers and judges who have submitted AI-generated data without verifying its accuracy.<sup>14,15,16</sup> As illustrated in Section 4, reliance on unverified AI output can violate duties of candor, competence, and diligence.

## *II Lack of Explainability and Transparency*

Another key concern is explainability. Most AI tools cannot reliably show how they arrive at a specific answer, relying instead on statistical prediction and internal reasoning that cannot be examined or replicated.<sup>17</sup> For lawyers, this uncertainty makes it difficult to assess whether an AI-generated suggestion is analytically sound. For judges, the risks are even more serious. When AI is used to review the record or draft portions of an opinion, litigants may have no indication that the analysis was influenced by a tool outside the evidentiary process. This raises due process concerns, since parties are entitled to know what materials and reasoning underlie a judicial decision. Unlike humans, AI systems are not bound by judicial ethics rules and cannot be questioned about omissions or errors. Transparency is essential to maintaining public trust. When AI-related mistakes come to light, undisclosed use of AI can undermine confidence in the integrity and independence of the judiciary.<sup>18</sup>

The explainability problem also implicates a broader principle recognized in international standards. The Chartered Institute of Arbitrators' Guideline on the Use of AI in Arbitration describes the "paternity of decision" requirement, which holds that human decision-makers must retain ownership of the reasoning behind their determinations.<sup>19</sup> Although developed in the arbitral context, this principle highlights a fundamental concern for both lawyers and judges. AI-generated analysis cannot substitute for accountable human judgment, especially when the basis for the tool's output cannot be explained or scrutinized.

## *III Confidentiality, Data Security Risks, and Bias*

Many AI platforms store user inputs or use them to train future versions of the model, creating significant confidentiality risks.<sup>20</sup> Uploading client materials or confidential drafts into commercial AI tools can expose privileged or confidential

13 IBM, 'What Are AI Hallucinations?' (2023).

14 *Mata*, *supra* note 3.

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

16 *Benjamin v. Costco Wholesale Corp.*, 779 F. Supp. 3d 341 (C.D. Cal. 2025).

17 Regulation (EU) 2024/1689, *supra* note 12.

18 ABA Journal, 'Federal Judge Withdraws Opinion After Lawyer Points Out Fake Quotes and Misstated Case Outcomes' (2025), available at <https://www.abajournal.com/news/article/federal-judge-withdraws-opinion-after-lawyer-points-out-fake-quotes-and-misstated-case-outcomes>.

19 CI Arb Guideline, *supra* note 11.

20 Lieberknecht, *supra* note 4.

information. Lawyers and judges may inadvertently reveal sensitive data when using public AI systems that lack robust security protections.

Bias presents a related challenge. AI systems trained on historical data may replicate or even amplify existing biases.<sup>21</sup> These risks can affect legal research, evidence evaluation, risk assessments, sentencing analyses, and document review. Lawyers and judges who rely on biased outputs may unknowingly adopt skewed perspectives. Because algorithmic bias is often difficult to detect, careful oversight and skepticism are necessary whenever AI tools inform legal analysis.

#### *IV Inequality of Platforms*

Disparities in access to advanced AI tools can also affect the fairness of legal proceedings. Larger firms with substantial resources may use sophisticated AI tools, while smaller firms or pro se litigants may rely on free tools that potentially produce more errors.<sup>22</sup> Unequal access to reliable AI can influence briefs, litigation strategies, and even judicial opinions. This “inequality of arms” raises concerns about substantive and procedural fairness, particularly when litigants do not know whether the opposing party, or even the judge, relied on tools of different capabilities.

#### *V Disclosure and Public Trust*

Undisclosed use of AI by lawyers and judges creates additional concerns. When lawyers rely on AI without disclosure, clients and opposing parties have no way to evaluate whether key arguments are supported by authority or whether the filing contains incomplete or flawed analysis. The same problem arises when judges use AI tools without informing litigants who have no way to evaluate how the tool may have influenced the court’s understanding of the issues. To address these fairness concerns, some courts have adopted standing orders requiring parties to disclose the use of generative AI, though few have addressed whether judges must reveal their own AI use.

Transparency is not meant to compromise advocacy or judicial independence; instead, it reassures litigants that the reasoning in a brief or judicial opinion is ultimately human, and that any AI-assisted text was reviewed and verified. Limited disclosure, such as noting that AI materially assisted with legal research, drafting, or analysis, together with confirmation that resulting content was independently reviewed and verified, can help preserve public trust without imposing significant burdens on lawyers and courts.

### **C Legal and Ethical Framework for AI Misuse**

In the United States, AI-related misconduct in legal practice falls under two systems of oversight: the Federal Rules of Civil Procedure (FRCP) and the American Bar Association

21 Ferrer, *supra* note 5.

22 National Center for State Courts, *Artificial Intelligence in State Courts* (2025), available at <https://www.ncsc.org/resources-courts/ai-state-courts>.

(ABA) Model Rules of Professional Conduct. Each governs a different dimension of professional conduct, and both may apply to the same incident. State courts maintain their own rules of civil procedure, although many states have modeled their procedural rules, to varying degrees, on the Federal Rules of Civil Procedure.<sup>23</sup>

AI-related misconduct often triggers both the FRCP and Model Rule frameworks. A lawyer who files a brief containing fabricated cases may face a sanction under the FRCP in federal court, while also being investigated by state disciplinary authorities for violating ethical duties under the Model Rules. This combined response reflects two forms of harm: damage to the functioning of the court system and damage to the integrity of the legal profession. The subsections below outline how each framework governs AI use.

### *I Federal Rules of Civil Procedure*

The FRCP provide procedural standards governing federal litigation, and Rule 11 is the primary vehicle for addressing AI-related misconduct.<sup>24</sup> Courts apply Rule 11 when lawyers submit filings containing unsupported legal or factual content, regardless of whether the errors originated with an AI tool or with human oversight.

Several provisions are especially relevant.

- Rule 11(b)(2) requires that legal arguments be grounded in existing law or in a non-frivolous argument for modifying it.<sup>25</sup> When AI tools generate nonexistent cases or misstate the law and lawyers submit them without verification, this rule is violated.
- Rule 11(b)(3) applies when factual assertions lack evidentiary support, including when AI produces distorted or fabricated procedural histories or summaries.<sup>26</sup>
- Rule 11(b)(1) prohibits filings for improper purposes, which courts sometimes cite when careless AI use wastes judicial resources or causes avoidable delays.<sup>27</sup>
- Rule 11(c) covers the sanctions themselves. Courts may impose fines, order corrections of the record, or limit an attorney's ability to appear, depending on the severity of the misconduct.<sup>28</sup> Law firms may also be held jointly responsible for violations under Rule 11(c)(1).<sup>29</sup>

At its core, Rule 11 makes clear that the responsibility to ensure accuracy cannot be delegated to a digital tool and that attorneys must independently verify all information submitted to a court.

23 John B. Oakley & Arthur F. Coon, 'The Federal Rules in State Courts: A Survey of State Court Systems of Civil Procedure', 61 Wash. L. Rev. (1986) 1367, available at: <https://digitalcommons.law.uw.edu/wlrvol61/iss4/7>.

24 Fed. R. Civ. P. 11.

25 Fed. R. Civ. P. 11(b)(2).

26 Fed. R. Civ. P. 11(b)(3).

27 Fed. R. Civ. P. 11(b)(1).

28 Fed. R. Civ. P. 11(c).

29 Fed. R. Civ. P. 11(c)(1).

## II American Bar Association Model Rules of Professional Conduct

The ABA Model Rules, in contrast, set out the broader ethical duties that apply across all areas of legal practice, not just litigation. State disciplinary bodies enforce these rules, and violations may lead to reprimand, suspension, or disbarment. AI misuse can implicate several ethical obligations, including the duty of competence (Rule 1.1), confidentiality (Rule 1.6), candor toward the tribunal (Rule 3.3), supervision of nonlawyers and tools used in practice (Rule 5.3), and professional misconduct (Rule 8.4).<sup>30</sup> These duties require lawyers to understand the tools they employ, verify the accuracy of their work, protect client information, and avoid misleading courts or clients, regardless of whether the errors result from a human mistake or an AI system. These rules address concerns such as honesty, diligence, confidentiality, and the need to oversee tools and assistants, duties that remain the lawyer's responsibility even when technology is involved.

Although the Model Rules provide a flexible foundation for evaluating attorney conduct involving AI, the rapid spread of generative tools has exposed areas where existing duties lack clarity. Several rules are being tested in ways the drafters did not anticipate, creating uncertainty about how far traditional doctrines extend and whether new commentary or amendments are necessary. The ABA recently issued *Formal Opinion 512*, emphasizing that lawyers must supervise AI under Rule 5.3, verify the accuracy of AI-generated content, and ensure AI tools do not compromise client confidentiality.<sup>31</sup>

Rule 1.1 (Competence) has taken on new significance, particularly in light of Comment 8, which instructs lawyers to maintain technological competence — increasingly understood to include the benefits and risks of AI tools.<sup>32</sup> These considerations are not optional. The ABA's 2024 ethics guidance confirms that technological competence includes understanding the limitations of AI tools, the risks of inaccuracy, and the need to independently verify all AI-assisted work.<sup>33</sup> Technological competence may now include an affirmative duty to evaluate whether an AI system is reliable, whether its training data may embed bias, and whether its confidentiality practices are adequate. As AI becomes more common, disciplinary authorities may expect lawyers to test AI outputs, verify their sources, and avoid tools that cannot be independently validated. Some bar associations have begun considering whether Comment 8 should be formally expanded to include explicit references to generative AI.

Rule 1.6 (Confidentiality) raises particularly significant concerns.<sup>34</sup> Many widely used AI platforms, including consumer-facing chatbots, store user inputs,

30 ABA Model Rules of Professional Conduct rr. 1.1, 1.6, 3.3, 5.3, 8.4 (2025).

31 American Bar Association, *Formal Opinion 512: Use of Generative Artificial Intelligence in the Practice of Law* (2024), available at [https://www.americanbar.org/content/dam/aba/administrative/professional\\_responsibility/ethics-opinions/aba-formal-opinion-512.pdf](https://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/ethics-opinions/aba-formal-opinion-512.pdf).

32 ABA Model Rules of Professional Conduct r. 1.1 cmt. 8 (2025).

33 American Bar Association, 'ABA Issues First Ethics Guidance on Use of AI Tools' (2024), available at <https://www.americanbar.org/news/abanews/aba-news-archives/2024/07/aba-issues-first-ethics-guidance-ai-tools/>.

34 ABA Model Rules of Professional Conduct r. 1.6 (2025).

reuse them for model training, or transmit them across servers in ways lawyers cannot control. When attorneys upload discovery materials, client summaries, medical records, financial data, or even factual narratives into an AI tool, they risk a breach of confidentiality. In some instances, the lawyer may unintentionally waive privilege. These risks are not hypothetical. *Formal Opinion 512* specifically warns that lawyers must not input client confidential information into AI tools unless they understand and can control how the data will be stored, transmitted, or reused.<sup>35</sup> Several jurisdictions have issued warnings or opinions indicating that lawyers must avoid entering client information into public AI systems unless they have reasonably verified the platform's privacy protections.<sup>36,37</sup> This presents a difficult challenge because many systems do not disclose their data retention or training practices. While Rule 1.6 clearly prohibits disclosures that compromise client confidentiality, it does not yet provide guidance on how attorneys should evaluate AI systems, nor does it resolve whether informed client consent is required when using such tools.

Rule 3.3 (Candor Toward the Tribunal) is implicated when attorneys fail to verify AI-generated citations, summaries, or legal propositions.<sup>38</sup> The Rule prohibits lawyers from knowingly making false statements of law or fact, but emerging disciplinary cases reveal that courts may treat an attorney's failure to check AI-generated content as the functional equivalent of "knowing" misconduct. Several courts have emphasized that it is no defense to say the lawyer believed the AI output was accurate. The duty of candor requires independent verification. This evolving interpretation suggests that, as AI tools become more capable of fabricating detailed but entirely fictitious authority, the standard for "reasonable inquiry" under Rule 11 and Rule 3.3 will rise accordingly. Lawyers may be expected not only to confirm the existence of cited cases but to review the full text to ensure the propositions are correct.

Rule 5.3 (Supervision of Nonlawyers) presents another unresolved issue, whether AI systems should be treated as "assistants" for purposes of supervision.<sup>39</sup> The rule requires lawyers to supervise nonlawyer employees and to ensure that their work is compatible with the lawyer's ethical obligations. Although AI is not a person, it performs tasks traditionally carried out by paralegals, clerks, or junior associates. As a result, several commentators and ethics committees have argued that Rule 5.3 extends to AI tools. This interpretation means that lawyers must adopt reasonable measures to supervise the output of AI systems, implement office-wide policies explaining when and how AI may be used, and ensure that AI does not perform tasks that require legal judgment. The ABA has emphasized that lawyers must establish appropriate supervisory policies and procedures when generative AI is used in legal practice. Some firms have already introduced internal AI-use policies requiring attorneys to document their review of AI-generated

35 ABA Formal Opinion 512, *supra* note 31.

36 ABA Model Rules of Professional Conduct r. 1.6 (2025).

37 Florida Bar, Ethics Opinion 24-1 (2024), available at <https://www.floridabar.org/etopinions/opinion-24-1/>.

38 ABA Model Rules of Professional Conduct r. 3.3 (2025).

39 ABA Model Rules of Professional Conduct r. 5.3 (2025).

content or to refrain from using AI with confidential data. These evolving practices illustrate how Rule 5.3 may be applied more rigorously as AI becomes further embedded in legal workflows.

Rule 8.4 (Misconduct) also plays an increasingly important role. Subsection (c) prohibits conduct involving dishonesty, fraud, deceit, or misrepresentation.<sup>40</sup> When attorneys submit AI-generated content that they know, or should have known, was inaccurate, they risk violating this provision. Courts in recent AI-related sanctions cases have stressed that misrepresenting the origin of legal research or concealing one's reliance on AI can amount to dishonesty. Additionally, Rule 8.4(d), which prohibits conduct prejudicial to the administration of justice, may apply when AI misuse causes substantial delays, wastes judicial resources, or undermines confidence in the court system.<sup>41</sup> As AI tools grow more sophisticated, disciplinary authorities may consider whether reckless or unsupervised AI use constitutes conduct prejudicial to justice even when a lawyer does not intend to deceive.

Finally, the growing overlap between these rules raises broader questions about whether AI-specific amendments are necessary. Some commentators advocate for the adoption of a new rule directly addressing AI use, including requirements for disclosure, verification, confidentiality safeguards, and training. Others argue that existing rules already impose sufficient obligations, and that the profession should avoid over-regulation of AI. What is clear, however, is that AI has already begun shaping the interpretation and application of the rules. As disciplinary decisions accumulate, the profession will need to decide whether to clarify these duties through formal amendments or allow courts and ethics committees to continue shaping the doctrine case by case.

## D AI Misuse in Practice: Disciplinary Cases and Judicial Sanctions

It is challenging to report the exact number of AI-related disciplinary cases filed in the U.S.<sup>42</sup> Most information about AI-related discipline comes from court decisions, press coverage, and state-level ethics commentary rather than national reporting. The National Center for State Courts (NCSC) maintains an 'Artificial Intelligence' section on their website, which provides guidance on AI in state courts, including court orders, rules, policies and an interactive map of state-by-state activity, although it does not report every case of AI-related discipline.<sup>43</sup> Another project, *DAIL - the Database of AI Litigation*, developed by the George Washington Ethical Tech Initiative and the George Washington Center for Law and Technology, is attempting to gather cases of ongoing and completed AI litigation and had amassed more than 250 cases at the time of writing.<sup>44</sup>

40 ABA Model Rules of Professional Conduct r. 8.4(c) (2025).

41 ABA Model Rules of Professional Conduct r. 8.4(d) (2025).

42 NCSC, *supra* note 22.

43 NCSC, *supra* note 22.

44 Ethical Tech Initiative and Center for Law and Technology, *DAIL - the Database of AI Litigation* (George Washington University Law School), available at <https://blogs.gwu.edu/law-eti/ai-litigation-database/>.

Recent AI-related disciplinary actions in the United States show how existing rules are being adapted to address new forms of misconduct. Although the governing rules do not mention artificial intelligence explicitly, they have been flexible enough to cover the most common problems of unverified citations, failure to supervise AI tools, and dishonesty once errors are discovered. Despite the flexibility of these rules, there is a critical need to further define and standardize AI regulation in the U.S. The cases also demonstrate a lack of standardization in discipline and that sanctions vary widely depending on aggravating and mitigating factors.

## *I Attorney Misconduct Cases*

### *1 Mata v. Avianca, Inc. 678 F Supp. 3d 443 (SDNY 2023)*<sup>45</sup>

This landmark case is a widely cited instance of AI misuse in litigation. Attorney Schwartz used a generative AI program to conduct legal research, which produced fabricated judicial opinions containing nonexistent cases, misidentified judges, and inaccurate quotations. Attorney LoDuca, the attorney of record, signed and filed the brief without conducting independent research. Opposing counsel alerted both attorneys to the fabricated cases more than once, but neither retracted nor verified the material, and both stood by it.

The court found that both lawyers acted in subjective bad faith and their conduct constituted an abuse of the judicial system. Both lawyers violated FRCP 11(b)(2) by failing to conduct a reasonable inquiry to ensure their filings were supported by existing law.<sup>46</sup> Schwartz used AI to perform legal research, failed to disclose its use, and failed to check the information it generated. LoDuca signed both an Affidavit and an Affirmation without independently checking the cases. Their law firm was also found to be in violation of Rule 11(c)(1), which holds law firms jointly responsible for violations by their attorneys unless exceptional circumstances exist.<sup>47</sup>

In deciding disciplinary actions, the Court considered the remedial steps of the law firm to order mandatory training on AI programs for all lawyers and staff, the amount of negative publicity the case generated, and the remorse and sincere embarrassment of the Respondents. Sanctions against Schwartz, LoDuca and their law firm included a \$5,000 fine payable to the court registry, mandatory notification to the client and judges mistakenly cited in the fabricated opinions and filing proof of compliance with these orders. Importantly, the court emphasized that technological ignorance could not excuse a failure of professional diligence, and the harms that result from the submission of fake cases include wasted time and

<sup>45</sup> *Mata*, *supra* note 3.

<sup>46</sup> Fed. R. Civ. P. 11(b)(2).

<sup>47</sup> Fed. R. Civ. P. 11(c)(1).

resources, damage to the reputation of the judiciary, and promotion of cynicism regarding the legal system.

2 *Wadsworth v. Walmart Inc.* 348 F.R.D. 489 (D. Wyo. 2025)<sup>48</sup>

In *Wadsworth v. Walmart Inc.*, the plaintiff's attorneys, Ayala, Morgan, and Goody, submitted motions with fake case citations generated by an AI program. Ayala, who primarily drafted the motions, relied on the AI tool to add case law citations but failed to verify their accuracy. Eight fictitious cases were cited as a result. Neither Morgan nor Goody reviewed the motions before signing them. The defendants discovered the fake citations and brought them to the court's attention. The Court found all three attorneys in violation of Rule 11(b)(2) for citing non-existent cases that were not supported by existing law and for not verifying citations.<sup>49</sup> They also improperly delegated Rule 11 responsibilities by relying on AI or another lawyer.<sup>50</sup> The court emphasized that attorneys have a duty to verify citations personally. Signing a legal document certifies that the attorney has conducted a reasonable inquiry into the law and facts of the case.

In deciding disciplinary actions, the Court considered that the lawyers promptly withdrew their motions after discovering the error, were forthcoming in admitting to the use of AI, and their law firm took remedial action to implement policies and training to prevent future issues. Ayala was fined \$3,000 and his pro hac vice admission was revoked. Morgan and Goody were each fined \$1,000. The firm itself was not sanctioned.

3 *Benjamin v. Costco Wholesale Corp.* 779 F Supp. 3d 341 (C.D. Cal. 2025)<sup>51</sup>

The Plaintiff's attorney, Victoria Hovsepyan, used an AI program to draft a reply in support of a motion to remand. The reply cited seven cases, five of which were non-existent. Hovsepyan admitted to not reviewing or verifying the citations. Parts of the reply argued against the plaintiff's own position, undermining her client. Hovsepyan claimed she used AI because of time constraints. She admitted her mistake and expressed remorse for her actions. She enrolled in a continuing education class focused on the use of AI in legal practice.

The Court found Hovsepyan in violation of Rule 11.<sup>52</sup> Her submission of false information constituted subjective bad faith and compromised the integrity of the judicial system. The Court also noted the use of AI-generated cases caused unnecessary work for both the opposing counsel and the Court. In deciding disciplinary actions against Hovsepyan, the Court considered her cooperation,

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

49 Fed. R. Civ. P. 11(b)(2).

50 Fed. R. Civ. P. 11.

51 *Benjamin v. Costco Wholesale Corp.*, 779 F. Supp. 3d 341 (C.D. Cal. 2025).

52 Fed. R. Civ. P. 11.

remedial action, and remorse as mitigating factors. The court fined Hovsepyan \$1,000 and ordered her to provide a copy of the sanctions to her client.

4 *People v. Crabill No 23PDJ067 (Colo. 2023)*<sup>53</sup>

Attorney Zachariah Crabill used AI to write a motion that he submitted to the court without checking the cited cases. He realized the cases were invalid before appearing in court but did not attempt to rectify his mistake or inform the court. The judge overseeing the case expressed concerns about the caselaw cited in the motion, and the attorney falsely blamed the mistakes on a legal intern. The motion was denied. Crabill submitted an affidavit six days after the hearing, admitting that he had used AI to write the motion.

The Court found Crabill in violation of Colorado Rules of Professional Conduct 1.1, mandating that a lawyer competently represent a client; 1.3, mandating that a lawyer act with reasonable diligence; 3.3(a)(1), mandating that a lawyer not knowingly make false statements, and 8.4(c), mandating that lawyers not engage in conduct involving dishonesty, fraud, deceit or misrepresentation.<sup>54</sup> Crabill was suspended for one year and one day, with 90 days actively served and nine months and one day stayed, subject to probation. In addition, his law firm terminated his employment.

## II Judicial Accountability

Although most documented instances of AI misuse involve attorneys, judges face similar risks as generative AI tools enter judicial chambers. The growing availability of AI tools raises unique challenges for judges that go beyond those faced by practicing attorneys. Judges must maintain procedural fairness, transparency, and independence, and any tool that obscures reasoning or influences outcomes without meaningful oversight threatens these values.<sup>55</sup> As with attorneys, the central issue is not the use of AI, but whether it compromises the integrity of the decision-making process.

Generative AI tools offer practical benefits to courts. However, their use in drafting judicial opinions or evaluating evidence presents serious challenges. Generative AI may introduce fabricated citations, mischaracterize precedent, or omit crucial context. These risks can undermine appellate review and raise questions about whether judicial reasoning is genuinely human.

Judges occupy a particularly sensitive role because they operate within constitutional due process guarantees. If AI influences a judicial decision, parties must be able to evaluate how the decision was reached. Yet AI systems cannot be questioned about omissions or errors, nor can their internal logic be examined.

In 2025, Judge Julien Xavier Neals of the District of New Jersey withdrew a published opinion after a litigant demonstrated multiple significant errors, including misstating the outcomes of three cases where motions to dismiss were

<sup>53</sup> *People v. Crabill*, supra note 10.

<sup>54</sup> Colo. Rules of Pro. Conduct rr. 1.1, 1.3, 3.3(a)(1), 8.4(c) (2023).

<sup>55</sup> Code of Conduct for U.S. Judges Canon 1, 2, 3.

granted, not denied, and attributing fake quotes to cases and defendants that do not appear in those opinions.<sup>56</sup> In October 2025, Judge Neals responded to U.S. Senate Judiciary Chairman Charles Grassley's questions, admitting that a law school intern used ChatGPT to perform legal research and the draft opinion should not have been docketed.<sup>57</sup>

Judge Henry T. Wingate of the Southern District of Mississippi encountered similar problems during the same period. In July 2025, Judge Wingate withdrew and replaced a temporary restraining order after litigants identified numerous factual inaccuracies, including incorrect party names, misquoted statutory language, and references to individuals who were not part of the litigation. Subsequent correspondence released by Senator Charles Grassley confirmed that a law clerk had used an AI platform in preparing the draft order and that the order had been docketed before the court's normal internal review process had been completed. Like Judge Neals, Judge Wingate reported adopting additional internal procedures to strengthen the review of AI-assisted judicial work following the incident.<sup>58</sup> Together, the Neals and Wingate incidents demonstrate that AI-related judicial drafting errors have occurred in multiple federal courts, reinforcing concerns about the absence of consistent national guidance for judicial AI use.

To date, no disciplinary action appears to have been taken against either Judge Neals or Judge Wingate. Nevertheless, these incidents underscore the risks associated with the use of generative AI in judicial chambers, particularly when draft opinions or orders are insufficiently reviewed before issuance. Taken together, these two incidents demonstrate that judicial AI misuse is not an isolated occurrence but an emerging institutional concern requiring clearer guidance and more consistent safeguards. Rather than being individual failures, these incidents expose the absence of comprehensive, uniform standards governing the use of generative AI within the judiciary. These incidents reinforce the need for proactive institutional guidance rather than continued reliance on corrective action after errors have already occurred.

### *III Aggravating and Mitigating Factors*

The number of reported AI-related sanctions decisions continues to grow, making it increasingly possible to identify emerging patterns in how courts respond to attorney misuse of generative AI. The George Washington University Ethical Tech Initiative's Database of AI Litigation (DAIL)<sup>59</sup> catalogs ongoing and completed litigation involving artificial intelligence across multiple subject areas, while Damien

<sup>56</sup> ABA Journal, *supra* note 18.

<sup>57</sup> Weiss, '2 Federal Judges Reveal AI Use by Staff Members Led to Error-Riddled Opinions', *ABA Journal* (2025), available at <https://www.abajournal.com/news/article/2-federal-judges-reveal-ai-use-by-staff-members-led-to-error-riddled-opinions>; U.S. Senate Committee on the Judiciary, *Grassley Releases Judges' Responses Owning Up to AI Use, Calls for Continued Oversight and Regulation* (23 Oct. 2025), available at <https://www.judiciary.senate.gov/press/rep/releases/grassley-releases-judges-responses-owning-up-to-ai-use-calls-for-continued-oversight-and-regulation>.

<sup>58</sup> Weiss, *supra* note 57; U.S. Senate Committee on the Judiciary, *supra* note 587.

<sup>59</sup> Ethical Tech Initiative and George Washington University Center for Law and Technology, *DAIL – Database of AI Litigation*, available at <https://blogs.gwu.edu/law-eti/ai-litigation-database/>.

Charlotin's AI Hallucination Cases Database<sup>60</sup> tracks legal decisions involving hallucinated citations and other AI-generated content in court and tribunal proceedings. Although these databases serve different purposes, together they illustrate the rapid growth of AI-related litigation and provide useful context for the emerging body of disciplinary decisions discussed below. The disciplinary decisions examined in this article represent one important subset of that development. Table 1 summarizes the principal AI-related sanctions decisions discussed in this article and illustrates the recurring aggravating and mitigating factors courts have considered when determining appropriate sanctions.

The comparison in Table 1 shows several recurring themes. Across these decisions, many of the same aggravating and mitigating factors influenced the imposed sanctions. Sanctions were generally more severe where attorneys failed to verify AI-generated material, ignored warnings, attempted to shift responsibility, or otherwise engaged in conduct that undermined the judicial process. By contrast, prompt disclosure, cooperation with opposing counsel and the court, remedial training, corrective office procedures, and the absence of a prior disciplinary record frequently weighed in favor of less severe sanctions. Although the number of AI-related disciplinary decisions remains relatively limited, these early cases suggest that courts are applying established principles of professional responsibility consistently while adapting them to misconduct involving generative AI. Importantly, the decisions demonstrate that courts have not treated the mere use of AI as sanctionable conduct. Rather, sanctions have been imposed when attorneys failed to satisfy existing duties of competence, candor, diligence, supervision, and reasonable inquiry. These emerging patterns demonstrate that existing ethical and procedural rules are capable of addressing AI-related misconduct, but they also highlight the limitations of relying primarily on case-by-case discipline rather than prospective guidance.

#### *IV Early Lessons from Case Law*

Early cases show how courts are applying traditional ethical and procedural rules to new forms of technological risk. They demonstrate that lawyers and judges remain fully responsible for the content of their filings, regardless of how it was produced, and that verification cannot be delegated to an AI system. They also highlight the need for stronger guidance and clearer expectations around responsible AI use, both to protect clients and to preserve confidence in the legal system.

### **E Gaps in the U.S. Regulatory Framework**

Although the FRCP and the ABA Model Rules provide mechanisms for addressing AI-related misconduct, neither system was designed with artificial intelligence in mind. As a result, U.S. regulation of AI use in legal practice is largely reactive and fragmented. Courts and disciplinary bodies often address problems only after they

60 Damien Charlotin, *AI Hallucination Cases Database*, available at <https://www.damiencharlotin.com/hallucinations/>.

Table 1. Comparative Analysis of Selected AI-Related Sanctions Decisions

| Case                                                       | Nature of AI Misuse                                                                                                                | Rule(s) Breached                                          | Aggravating Factors                                                                                                                                                                            | Mitigating Factors                                                                                                     | Sanction                                                                                                                             |
|------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------|
| <i>Mata v. Avianca, Inc.</i> (S.D.N.Y. 2023)               | Attorney submitted AI-generated fabricated judicial opinions, and fictitious citations, and false quotations without verification. | FRCP 11(b)(2);<br>FRCP 11(c)(1) (law firm responsibility) | Failed to verify AI-generated authorities; continued relying on fabricated cases after opposing counsel questioned them; signed filings without independent review; wasted judicial resources. | Expressed remorse; significant negative publicity; law firm implemented mandatory AI training for attorneys and staff. | \$5,000 monetary sanction; ordered to notify client and judges falsely cited; required to file proof of compliance.                  |
| <i>Wadsworth v. Walmart Inc.</i> (D. Wyo. 2025)            | Attorneys submitted motions containing eight fictitious AI-generated case citations.                                               | FRCP 11(b)(2);<br>Rule 11 certification responsibilities  | Failed to verify AI-generated citations; attorneys signed filings without adequate review; improperly delegated Rule 11 responsibilities.                                                      | Promptly withdrew motions; cooperated with the court; admitted AI use; law firm adopted AI policies and training.      | Ayala fined \$3,000 and prohibited vice admission revoked; Morgan and Goody each fined \$1,000; no sanction imposed on the law firm. |
| <i>Benjamin v. Costco Wholesale Corp.</i> (C.D. Cal. 2025) | Attorney filed reply brief containing five nonexistent AI-generated cases and arguments contrary to the client's position.         | FRCP 11                                                   | Failed to review AI-generated brief; submitted fabricated authorities; caused unnecessary work for opposing counsel and the court.                                                             | Promptly admitted error; expressed remorse; completed continuing legal education on AI use.                            | \$1,000 monetary sanction and ordered to provide the sanctions order to the client.                                                  |

|                                          |                                                                                                                                                          |                                                                         |                                                                                                                                                       |                                                  |                                                                                                                                                  |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>People v. Crabill</i><br>(Colo. 2023) | Attorney filed AI-generated motion containing fabricated authorities, later falsely blamed a legal intern, and failed to immediately correct the record. | Colorado Rules of Professional Conduct 1.1, 1.3, 3.3 (a)(1), and 8.4(c) | Failed to verify authorities; knowingly misrepresented the source of the errors; failed to promptly correct false information presented to the court. | Eventually admitted responsibility by affidavit. | Suspended for one year and one day, with 90 days actively served and the balance stayed subject to probation; employment terminated by law firm. |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|

**Source:** Compiled by the author from *Mata v. Avianca, Inc.*, *Wadsworth v. Walmart Inc.*, *Benjamin v. Costco Wholesale Corp.*, and *People v. Crabill*.

arise, leaving lawyers and judges with no clear guidance on how to use AI tools responsibly. The absence of uniform standards creates uncertainty for lawyers and judges who increasingly rely on AI tools in their daily work.

One significant gap is the absence of any national ethical standard specifically addressing the use of AI by attorneys and judges. The ABA has not fully amended the Model Rules to include AI-specific commentary, and no Model Rule directly requires disclosure of AI use, verification of AI-generated content, or a defined level of technological competence beyond the general language of Rule 1.1. While Comment 8 encourages lawyers to keep abreast of changes in technology, it offers no concrete guidance on how attorneys should assess or supervise AI tools. This leaves practitioners unsure when client consent is required, whether certain AI tools are inappropriate for confidential work, or what level of review is ethically mandated. As a result, courts may apply different expectations, creating inconsistencies in how sanctions are imposed for similar conduct.

A second major gap arises from the patchwork of state-level initiatives addressing AI use.<sup>61, 62</sup> Several jurisdictions have begun issuing policies or standing orders, but they vary widely in scope and language.<sup>63</sup> For example, the Illinois Supreme Court adopted a 2024 policy allowing the use of AI in court proceedings so long as lawyers comply with existing confidentiality and diligence requirements.<sup>64</sup> Some federal courts, including the Northern District of Texas, have adopted standing orders requiring lawyers to certify that AI-generated content was reviewed for accuracy or to disclose the extent of AI involvement.<sup>65</sup> Other states have issued only advisory guidance but no binding rules. Meanwhile, many states provide no guidance at all. This uneven landscape makes it difficult for attorneys who practice in multiple jurisdictions to know what is required of them, particularly when court policies may conflict with or exceed statewide ethical rules.

Another significant gap concerns the absence of regulation directed toward AI vendors whose tools are widely used in legal practice. While e-discovery platforms and research databases face some industry oversight, widely available generative AI tools like ChatGPT, Gemini, and Claude are not regulated specifically for legal use. These tools may store user data, incorporate confidential information into training sets, or generate fabricated legal content without warning. Yet no federal or state entity sets standards for data accuracy, privacy, or verification for AI systems marketed to attorneys. Lawyers, judges, and law firms remain solely

61 Illinois Supreme Court, *Policy on Artificial Intelligence* (2024), available at <https://ilcourtsaudio.blob.core.windows.net/antilles-resources/resources/e43964ab-8874-4b7a-be4e-63af019cb6f7/Illinois%20Supreme%20Court%20AI%20Policy.pdf>.

62 Standing Order, N.D. Tex. (2023), available at <https://www.texenrls.org/wp-content/uploads/2024/07/A-sampling-of-AI-Court-Orders.pdf>.

63 University of Chicago Library, *Generative AI in Legal Research, Education, and Practice* (2024), available at [https://guides.lib.uchicago.edu/AI/Practice#:~:text=Judge%20Brantley%20Starr%20\(N.D.,AI%20on%20May%2030%2C%202023](https://guides.lib.uchicago.edu/AI/Practice#:~:text=Judge%20Brantley%20Starr%20(N.D.,AI%20on%20May%2030%2C%202023).

64 Illinois Supreme Court, *supra* note 61.

65 University of Chicago Library, *supra* note 63.

responsible for identifying risks they may not be technically equipped to assess, increasing the likelihood of misuse.

The fragmented approach also affects judges. Although the position changed in July 2025 when the Administrative Office of the United States Courts distributed interim guidance on the use of generative AI within the federal judiciary, the guidance remains limited in scope. It cautions judges and court personnel against delegating core judicial functions to AI, emphasizes that users remain responsible for verifying AI-generated content, advises particular caution when AI is used in connection with novel legal questions, and leaves decisions regarding disclosure to the discretion of individual courts.<sup>66</sup> Importantly, the guidance is expressly interim and applies only within the federal judiciary. Below the federal level the position remains uneven: some courts have adopted standing orders or disclosure requirements, others strongly discourage reliance on AI, and many have no policy at all. Without widely accepted standards, judges must determine for themselves when AI tools may be used, what level of verification is required, and whether disclosure to litigants is necessary. This inconsistency creates uncertainty for litigants and may raise due process concerns when parties cannot determine whether AI influenced a judicial ruling.

The Administrative Office's interim guidance represents an important development because it reflects many of the same principles advocated in this article.<sup>67</sup> Its emphasis on human accountability, independent verification of AI-assisted work, and caution in the use of AI for substantive judicial functions demonstrates that the federal judiciary is already moving toward a more structured approach to AI governance. At the same time, because the guidance is interim, non-binding outside of the federal judiciary, and leaves key issues such as disclosure to local discretion, it does not eliminate the fragmented and inconsistent approach that continues to characterize AI regulation across the United States.

Despite these encouraging developments, significant regulatory gaps remain. These regulatory gaps matter because they contribute to the reactive nature of AI-related discipline. As the cases in Section 4 illustrate, sanctions typically occur only after something goes wrong, after fabricated cases appear in a brief, after a judge discovers falsified information, or after opposing counsel uncovers inaccuracies. Without clear preventative guidance, lawyers and judges may not realize their responsibilities or the risks associated with AI tools until misconduct has already occurred.

A more unified and proactive regulatory framework would reduce these risks by establishing clearer expectations for verification, supervision, confidentiality, and disclosure. The CI Arb Guideline provides one example of a proactive model, but the United States currently lacks an equivalent nationwide standard. As AI becomes more integrated into legal practice, filling these gaps will be critical to

66 Administrative Office of the United States Courts, *Director's Annual Report 2025: Court Operations* (2025), available at <https://www.uscourts.gov/data-news/reports/annual-reports/directors-annual-report/annual-report-2025/court-operations-annual-report-2025>.

67 Administrative Office of the United States Courts, *supra* note 66.

protecting clients, maintaining public confidence, and ensuring the fair administration of justice.

## **F International Models and Comparative Approaches**

While the United States continues to rely substantially on a decentralized combination of existing ethical rules, procedural requirements, and case-based enforcement, recent judicial and professional guidance reflects movement toward a more proactive approach. Several international frameworks, including those implemented by the European Union, the United Kingdom, Canada, and Australia, provide insight into how different legal systems balance innovation, transparency, and professional responsibility. These models, along with developments in international arbitration, illustrate how rapidly the global regulatory landscape is evolving and highlight key differences in how jurisdictions respond to technological risk. Although these jurisdictions regulate AI differently, each provides an example of how legal systems are beginning to address AI use by lawyers and judges.

### *I The CIArb Guideline on the Use of AI in Arbitration (2025)*

The Chartered Institute of Arbitrators (CIArb) Guideline on the Use of AI in Arbitration (2025) offers one of the clearest early examples of a structured, proactive approach to regulating AI in adjudicative settings.<sup>68</sup> Unlike the more decentralized U.S. framework, which continues to rely substantially on existing professional rules and case-by-case enforcement, the Guideline establishes AI-specific expectations in advance. Its core elements include:

- Tribunal directed disclosure requirements concerning certain uses of AI in arbitral proceedings;
- Explicit retention of human responsibility, making clear that AI cannot replace judgment;
- Verification requirements for AI-assisted work; and
- Model procedural orders addressing AI use in arbitral proceedings.

These provisions are designed to reduce problems before they occur, promote transparency, and preserve the integrity of arbitral proceedings. They also bring AI-specific expectations together within a single framework, in contrast to the more fragmented approach that continues to characterize the United States.

### *II The European Union's Artificial Intelligence Act (2024)*

One of the most significant developments in AI regulation is the European Union's Artificial Intelligence Act (2024), which establishes a binding, risk-based regulatory framework for AI systems.<sup>69</sup> Although the Act does not directly regulate lawyers or judges simply because they use AI, certain AI systems intended for use in the administration of justice are classified as "high-risk". This includes certain sys-

<sup>68</sup> CIArb Guideline, *supra* note 11, Arts. 4.4, 7.1-7.7.

<sup>69</sup> Regulation (EU) 2024/1689, *supra* note 12.

tems intended to assist judicial authorities in researching and interpreting facts and law or applying the law to a concrete set of facts, but does not extend to purely ancillary administrative activities.<sup>70</sup> High-risk systems are subject to requirements, including:<sup>71</sup>

- Risk management and data governance;
- Technical documentation and record keeping;
- Transparency and the provision of information to deployers;
- Human oversight; and
- Accuracy, robustness, and cybersecurity.

The significance of the Act for the legal system therefore depends on the purpose for which an AI system is intended and deployed, rather than on the fact that AI is used in legal work. For judicial use in particular, the Act's treatment of certain AI systems used in the administration of justice as high-risk reflects concerns that automated tools may affect fundamental rights, due process, and the independence of judicial decision-making.

This risk-based approach differs from the United States' more decentralized approach. The EU establishes regulatory obligations in advance for AI systems falling within defined risk categories, whereas the United States continues to rely primarily on existing professional duties, procedural rules, court specific policies, and developing institutional guidance. The EU model therefore provides an important example of proactive regulation: rather than waiting for AI-related misconduct to occur, it imposes safeguards based on the risks associated with particular uses of AI. As a result, the EU approach aligns more closely with CI Arb's proactive model than with the more decentralized and largely reactive U.S. framework.

### *III United Kingdom: Judicial Guidance and Professional Caution*

The United Kingdom's (UK) Judicial Office has taken a more cautious but still forward-looking approach to the use of generative AI. In December 2023, it issued guidance advising judicial office holders not to use generative AI to analyze confidential material or draft substantive portions of judicial decisions.<sup>72</sup> The guidance emphasizes that judicial office holders remain personally responsible for the accuracy, confidentiality, and integrity of their work, regardless of whether AI tools are used to assist with administrative or research tasks. Among other things, the guidance emphasizes:

- Limited AI use for administrative tasks.
- Strict avoidance of AI for legal reasoning or evidence evaluation.
- Heightened vigilance regarding confidentiality risks.
- Transparency when AI assistance influences court work.

70 Regulation (EU) 2024/1689, Annex III, point 8(a); Recital 61.

71 Regulation (EU) 2024/1689, Arts. 9-15.

72 Judicial Office, *Artificial Intelligence (AI): Guidance for Judicial Office Holders* (31 October 2025), available at <https://www.judiciary.uk/wp-content/uploads/2025/10/Artificial-Intelligence-AI-Guidance-for-Judicial-Office-Holders-2.pdf>.

Similarly, the Bar Council of England and Wales has warned barristers that AI cannot be relied on as a legal research tool and that all outputs must be independently verified. Together, these judicial and professional guidance documents illustrate a coordinated approach to AI governance that places primary emphasis on human accountability rather than prohibiting the use of technology. These warnings emphasize that lawyers, not software, remain responsible for accuracy and ethical compliance.

The UK's approach resembles the CI Arb Guideline in its emphasis on professional accountability and transparency. It differs from the United States in that the judiciary itself has taken the lead in providing ethical expectations, whereas U.S. courts generally act only after problems arise.

#### *IV Canada: Early Efforts Toward Technological Competence*

Canadian regulators have begun incorporating AI-specific concerns into existing standards of professional and judicial competence. The Law Society of Ontario has published guidance addressing the professional obligations of lawyers using generative AI, including technological competence, confidentiality, and verification of AI-generated material.<sup>73</sup> The Canadian Judicial Council has similarly issued guidance addressing risks associated with judicial use of AI, including bias, hallucinations, and data privacy. Although Canada has not adopted a comprehensive AI-specific framework for the legal profession comparable to the EU's approach, these developments reflect a movement toward addressing AI through existing principles of technological competence, professional responsibility, and judicial ethics.<sup>74</sup>

#### *V Australia: Principles-Based AI Governance*

Australia has taken a principles-based approach to AI use in the legal system. The Australian Institute of Judicial Administration and the Law Council of Australia have emphasized responsible use, technological competence, human oversight, and the continuing responsibility of legal professionals for work produced with AI assistance. These principles reinforce the need for lawyers and judges to understand the limitations of AI tools and ensure that technology does not replace pro-

<sup>73</sup> Law Society of Ontario, *Licensee Use of Generative Artificial Intelligence* (2024), available at <https://lawsocietyontario-dwd0dscmayfwh7bj.a01.azurefd.net/media/lso/media/lawyers/practice-supports-resources/white-paper-on-licensee-use-of-generative-artificial-intelligence-en.pdf>.

<sup>74</sup> Felsky and Eltis, *Guidelines for the Use of Artificial Intelligence in Canadian Courts* (Canadian Judicial Council, 2024), available at <https://cjc-ccm.ca/sites/default/files/documents/2024/AI%20Guidelines%20-%20FINAL%20-%202024-09%20-%20EN.pdf>.

fessional legal judgment.<sup>75,76,77</sup> This principles-based approach is also reflected in Australia's debate over AI regulation. In September 2025, the Law Council of Australia responded to an Interim Report published by the Productivity Commission, supporting expansion of existing regulation where necessary while opposing the adoption of a comprehensive AI act.<sup>78</sup> Although Australia has not adopted a comprehensive AI specific framework for the legal profession, its developing guidance shows how AI-related concerns are being incorporated into existing standards of professional responsibility and judicial governance.

## VI *International Arbitration as a Model for AI Governance*

International arbitration has emerged as an important setting for the development of guidance on the responsible use of AI. The CIArb Guideline establishes common expectations for AI use by parties, counsel, and arbitrators from different legal systems. Because international arbitration frequently involves participants from jurisdictions with differing approaches to AI regulation, guidance that sets common expectations regarding verification, disclosure, confidentiality, and human responsibility can promote greater consistency across jurisdictions.

Rather than creating entirely new obligations, the CIArb Guideline brings existing principles of disclosure, verification, confidentiality, and human responsibility together in one framework. In an international setting, those common expectations can reduce uncertainty where participants are otherwise subject to different domestic approaches to AI governance. The CIArb model provides a useful comparator for the United States, where AI-related obligations remain dispersed among procedural rules, professional conduct standards, standing orders, ethics guidance, and the interim guidance issued for the federal judiciary.

## VII *Implications for U.S. Legal Practice*

The growing divergence between U.S. and international approaches raises important questions:

- How might the EU AI Act affect U.S. lawyers engaged in cross-border practice involving AI systems subject to the Act?
- Should U.S. courts consider adopting CIArb-style disclosure rules or procedural orders?

75 Australian Institute of Judicial Administration, *AI Decision-Making and the Courts: A Guide for Judges, Tribunal Members and Court Administrators* (2022), available at [https://aija.org.au/wp-content/uploads/woocommerce\\_uploads/2022/06/AI-DECISION-MAKING-AND-THE-COURTS\\_Report\\_V5-2022-06-20-11zkl.pdf](https://aija.org.au/wp-content/uploads/woocommerce_uploads/2022/06/AI-DECISION-MAKING-AND-THE-COURTS_Report_V5-2022-06-20-11zkl.pdf).

76 Law Council of Australia, *Artificial Intelligence in the Legal Profession* (2025), available at <https://lawcouncil.au/policy-agenda/advancing-the-profession/artificial-intelligence-and-the-legal-profession>.

77 Law Council of Australia, *Use of AI in the Courts* (Submission, 16 June 2025).

78 Law Council of Australia, *Harnessing Data and Digital Technology: Interim Report* (2025), available at <https://lawcouncil.au/resources/submissions/harnessing-data-and-digital-technology-interim-report>.

- Will inconsistent state guidance disadvantage lawyers practicing in multiple jurisdictions?
- Could international arbitration norms eventually influence U.S. judicial expectations?

These questions highlight an important difference in regulatory approach. The United States continues to rely substantially on existing professional duties, procedural rules, and case-by-case enforcement, although recent judicial and professional guidance reflects movement toward more proactive governance. By comparison, the international approaches examined here generally place greater emphasis on setting expectations in advance through binding regulation, judicial guidance, professional standards, or structured procedural frameworks. If the U.S. framework remains fragmented, American lawyers engaged in cross-border practice may increasingly encounter expectations concerning AI use that are more specific than those imposed by domestic professional rules.

## G Convergence and Divergence in U.S. and International Approaches

The international models discussed above reveal both common themes and important distinctions in how legal systems are responding to the rise of AI in professional practice. Across the approaches examined here, several recurring principles emerge, including human responsibility, technological competence, accuracy, transparency, and appropriate oversight of AI-assisted work.<sup>79</sup> These shared themes reflect a growing recognition that AI must be supervised, verified, and used in ways that preserve the integrity of legal processes.

Where the approaches diverge is primarily in the level of formality and centralization. The United States continues to rely heavily on a decentralized regulatory structure in which ethical obligations arise from the ABA Model Rules, state bar authorities, procedural rules, court specific policies, and case-based discipline. Recent developments, including the Administrative Office's interim guidance for the federal judiciary, demonstrate movement toward more proactive governance. Nevertheless, AI-specific expectations remain dispersed across institutions and jurisdictions, creating uncertainty about the standards applicable to lawyers and judges.

By contrast, several of the international models examined in this article establish more explicit expectations before misconduct occurs. The EU AI Act uses a binding, risk-based framework, while the CIArb Guideline establishes AI specific expectations for arbitral proceedings. Although these frameworks differ in scope and legal effect, both illustrate approaches that establish safeguards before AI-related problems arise rather than relying primarily on discipline after misconduct occurs. Their preventive orientation provides an important contrast with the more fragmented U.S. framework.

These differences have important implications for the legal profession. As cross-border practice becomes more common and as AI tools become more deeply

<sup>79</sup> OECD, *AI Principles* (2021).

embedded in legal work, the United States may face pressure to adopt clearer and more consistent standards. Aligning domestic regulation with emerging global norms may also be necessary to preserve public confidence in the justice system and maintain the credibility of U.S. legal institutions in international contexts.

## H Reforms: Toward a Unified Regulatory Approach

The cases and comparative frameworks discussed above demonstrate that existing ethical and procedural rules can address many instances of AI misuse, but expectations remain fragmented and enforcement frequently occurs only after misconduct has arisen. Because generative AI is evolving faster than existing ethical frameworks, incremental or reactive updates will not be sufficient.<sup>80</sup> Waiting for disciplinary cases to clarify expectations places courts, lawyers, and litigants at unnecessary risk. A more effective system would set expectations in advance, guiding lawyers and judges before problems occur. A coherent U.S. framework for regulating AI in legal practice would combine the strengths of both domestic and international systems. The following reforms aim to create a predictable, preventive framework that supports responsible AI use across the legal system.

### I *Disclosure Requirements*

Drawing on the disclosure framework developed by the Chartered Institute of Arbitrators, U.S. courts and bar associations should consider rules requiring lawyers and judges to disclose when AI has played a substantive role in drafting, research, or analysis that influences a filing or decision.<sup>81</sup> Such a requirement could promote transparency without discouraging responsible use. Several courts have begun requiring parties to certify or disclose certain uses of AI in filings, but these requirements remain inconsistent across jurisdictions and generally focus on lawyers rather than judicial use of AI. Clear, uniform disclosure rules would reassure litigants that any AI-generated content was reviewed and verified by a human decision-maker.

Disclosure would not undermine advocacy or judicial independence; rather, it would reinforce public trust by demonstrating transparency. A brief notation that AI materially assisted with legal research, drafting, or analysis, accompanied by confirmation that the resulting content was independently reviewed and verified, could promote accountability without imposing significant burdens on lawyers or courts.

### II *Model Procedural Orders*

Courts and bar associations should consider adopting standardized procedural orders governing AI use, drawing on the model procedural orders provided by the CI Arb for arbitration proceedings.<sup>82</sup> These orders could:

80 National Institute of Standards and Technology, *AI Risk Management Framework* (2023), available at <https://www.nist.gov/itl/ai-risk-management-framework>.

81 CI Arb Guideline, *supra* note 11, Arts. 4.4, 7.1-7.7.

82 CI Arb Guideline, *supra* note 11, Appendix B.

- Identify acceptable and prohibited uses of AI in litigation;
- Specify verification obligations for lawyers;
- Clarify whether and how AI use must be disclosed; and
- Establish limits on the use of AI in judicial decision-making.

Standardized procedural orders could reduce inconsistencies among courts and provide clearer expectations for lawyers practicing across jurisdictions. They could also provide a framework for integrating AI into litigation in a controlled and transparent manner.

### *III Competence and Education Standards*

As AI tools become increasingly embedded in legal practice, both lawyers and judges must develop sufficient technological competence to use these tools responsibly. While Rule 1.1's duty of competence already requires attorneys to understand the benefits and risks associated with relevant technology, the rapid evolution of generative AI demands more targeted and systematic training. Without formal guidance, professionals may inadvertently compromise confidentiality, accuracy, or fairness. Strengthening educational standards and institutional support is essential to ensuring ethical and effective AI use across the legal system.

For lawyers, mandatory continuing legal education (CLE) focused specifically on AI risks should become a standard requirement. *Formal Opinion 512* reinforces the need for AI-related competence by explaining that lawyers using generative AI must understand the technology's capabilities and limitations and remain mindful of their existing duties concerning confidentiality, competence, candor, and supervision.<sup>83</sup> Such training should address data privacy concerns, confidentiality obligations under Rule 1.6, verification practices for AI-generated research and citations, bias detection, and the limits of automation in legal reasoning. CLE programs should also teach attorneys how to evaluate AI vendors and understand data retention, model training practices, and reliability measures. Because AI tools increasingly perform tasks once delegated to junior attorneys or paralegals, lawyers must establish appropriate policies and review procedures to ensure AI-assisted work complies with their professional obligations.

Judges require similar, if not heightened, training. The Administrative Office's interim guidance represents an important step in this direction by emphasizing human accountability, independent verification, and caution when AI is used in connection with substantive judicial functions.<sup>84</sup> Judicial reliance on AI has broader implications because it affects not only individual litigants but also public confidence in the courts. Judicial education should therefore include instruction on due process considerations, risks associated with AI-generated summaries or evidence evaluations, and the importance of transparency when AI informs judicial work. Training programs should help judges understand when AI may be used appropriately, such as for administrative tasks, and when its use could undermine independence or appellate review. Courts may also benefit from developing internal

<sup>83</sup> ABA Formal Opinion 512, *supra* note 31.

<sup>84</sup> Administrative Office of the United States Courts, *supra* note 66.

review procedures or committees to promote consistent implementation of AI policies and best practices.

Institutional policies will be critical to supporting this increased emphasis on technological competence. Law firms should adopt written AI-use policies specifying acceptable uses, verification requirements, and confidentiality safeguards. Courts and judicial systems should consider similar guidelines to establish consistent expectations for judges and court staff. Institutions may also consider developing certification programs or internal training modules that provide structured pathways for lawyers and judges to demonstrate competence with AI tools.

Strengthening technological competence across the legal profession is not simply a matter of keeping pace with innovation. It is a foundational requirement for ensuring that AI supports, rather than undermines, the traditional values of competence, diligence, candor, and independent judgment. By adopting clear training standards and institutional policies, the legal system can better align AI use with the profession's ethical commitments and protect both the integrity of legal processes and public trust.

#### *IV Clarifying Ethical Duties for AI Use*

The ABA Model Rules should be amended or supplemented with commentary that explicitly addresses AI. Although existing duties on competence, candor, confidentiality, and supervision provide a basis for addressing AI-related conduct, they do not directly address issues such as hallucinations, transparency, embedded bias, or unauthorized data retention.<sup>85</sup> Clear commentary could specify that:

- Rule 1.1 includes an obligation to understand the capabilities and limits of AI tools used in legal practice;
- Rule 1.6 requires lawyers using AI tools to take reasonable measures to protect confidential client information;
- Rule 3.3's duty of candor applies to the submission of AI-generated legal authority;
- Rule 5.3's supervisory principles apply when AI tools perform functions analogous to those traditionally performed by nonlawyer assistants; and
- Rule 8.4 applies when misuse involves dishonesty, deceit, or misrepresentation or is prejudicial to the administration of justice.

These clarifications could remove ambiguity and promote greater consistency in the application of existing ethical duties to AI-related conduct.

#### *V Lessons from Arbitration Standards*

The CIArb Guideline provides a useful model for regulating AI through clear expectations and structured safeguards that may also inform judicial and professional ethics. Part IV emphasizes that arbitrators must retain responsibility for their

<sup>85</sup> ABA Model Rules of Professional Conduct rr. 1.1, 1.6, 3.3, 5.3, 8.4 (2025); ABA Formal Opinion 512, *supra* note 31.

decision-making and independently verify AI-generated information on which they intend to rely.<sup>86</sup> The Guideline also provides a framework under which disclosure of certain AI uses may be required.<sup>87</sup> Although the Guideline was written for arbitration, the principles behind it translate well to judicial and legal practice. These principles reinforce the importance of appropriate disclosure, independent verification, and human accountability in maintaining confidence in adjudicative processes.

## VI *Judicial-Specific Reforms*

Judges may require separate guidance because of the heightened due process concerns associated with judicial decision-making. The Administrative Office's interim guidance represents an important first step at the federal level, but more specific and consistent policies are needed to govern judicial use of generative AI.<sup>88</sup> Courts should adopt policies specifying:

- Acceptable and prohibited uses of AI in drafting opinions;
- Verification obligations for any AI-assisted analysis;
- Circumstances requiring disclosure to litigants; and
- Internal review procedures to promote compliance with applicable AI policies and verification requirements.

Because judicial opinions shape legal standards and public trust, judges must exercise particular caution when using generative AI tools. Clear standards, judicial transparency, and human accountability are essential to maintaining the legitimacy and integrity of judicial decision-making.

## I **Conclusion**

The growing use of artificial intelligence in legal practice offers clear benefits, but it also brings risks that the profession is only beginning to confront. Early disciplinary cases show that when lawyers rely on AI without verification or hide their use of it, courts and disciplinary bodies must rely on duties of competence, candor, and honesty. Existing rules have proven adaptable, but reliance on case-by-case enforcement has contributed to uneven expectations and uncertainty about acceptable AI use. Recent judicial and professional guidance represents movement toward a more proactive approach, although significant gaps remain.

International models, particularly the Chartered Institute of Arbitrators' Guideline on the Use of AI in Arbitration (2025), demonstrate a proactive approach by establishing expectations concerning appropriate disclosure, verification, and human responsibility before problems arise. These principles provide a useful model for how domestic regulation could evolve. Clearer professional guidance, more consistent procedural standards, appropriate disclosure requirements, and

86 CI Arb Guideline, *supra* note 11, pt. IV.

87 CI Arb Guideline, *supra* note 11, Arts. 4.4, 7.1-7.7.

88 Administrative Office of the United States Courts, *supra* note 66.

expanded education on AI risks would help shift the profession from after-the-fact correction to prevention.

Whatever tools lawyers and judges adopt, the core obligations of the profession remain the same. AI may assist with research, drafting, and analysis, but it cannot replace independent judgment or relieve professionals of responsibility for the accuracy of their work. Strengthening the ethical and procedural framework around AI will help ensure that technology supports, rather than undermines, the fairness and integrity of the justice system. The challenge moving forward is not whether AI will be used, but whether it will be used in a way that preserves the values on which the legal profession depends. Greater consistency across jurisdictions would provide clearer expectations for lawyers and judges while allowing the legal profession to adopt useful technologies without sacrificing accountability, fairness, or public trust.