

4-13-2026

Algorithmic Ethics in an Era of Agentic AI Advocacy: An Analysis of AI's Impact on the Model Rules of Professional Conduct and the Model Code of Judicial Conduct

Michael D. Murray
University of Kentucky

Follow this and additional works at: <https://commons.stmarytx.edu/lmej>



Part of the [Education Commons](#), and the [Law Commons](#)

Recommended Citation

Michael D. Murray, *Algorithmic Ethics in an Era of Agentic AI Advocacy: An Analysis of AI's Impact on the Model Rules of Professional Conduct and the Model Code of Judicial Conduct*, 16 ST. MARY'S J. ON LEGAL MALPRACTICE & ETHICS 279 (2026).

Available at: <https://commons.stmarytx.edu/lmej/vol16/iss2/4>

This Article is brought to you for free and open access by the St. Mary's Law Journals at Digital Commons at St. Mary's University. It has been accepted for inclusion in St. Mary's Journal on Legal Malpractice & Ethics by an authorized editor of Digital Commons at St. Mary's University. For more information, please contact sfowler@stmarytx.edu, egoode@stmarytx.edu.

ARTICLE

Michael D. Murray

Algorithmic Ethics in an Era of Agentic AI Advocacy: An Analysis of AI's Impact on the Model Rules of Professional Conduct and the Model Code of Judicial Conduct

Author. Michael D. Murray is a University Research Professor and the Spears Gilbert Professor of Law at the University of Kentucky, J. David Rosenberg College of Law. He is the principal investigator of the Artificial Intelligence and the Law Project. Professor Murray thanks Interim Dean James Duff, Acting Associate Dean for Research Joshua Douglas, and the Rosenberg College of Law faculty research fund for their on-going support of his research. Research supporting this article was undertaken pursuant to a 2025-2026 OVPR CURATE Grant and a 2025 University Research Professor Grant of the University of Kentucky.

ARTICLE CONTENTS

I.	Introduction: The New Digital Frontier of Legal Practice.....	284
A.	Defining the Tools: From Generative Assistants to Autonomous Agents.....	289
B.	The Enduring Framework: Applying Rules of Reason to a New Reality.....	290
II.	Navigating the Model Rules of Professional Conduct with AI	291
A.	The Client-Lawyer Relationship: Core Duties Reimagined	292
1.	Rule 1.1 (Competence): Beyond Legal Knowledge to Technological Proficiency.....	292
2.	Rule 1.6 (Confidentiality of Information): Protecting Client Data in the Age of AI Models and Training Protocols	296
3.	Rule 1.4 (Communications): The Duty to Inform and Obtain Consent.....	300
a.	When Is Disclosure Required?.....	300
b.	The Spectrum of Guidance from State Bars	301
4.	Rule 1.5 (Fees): Ethical Billing for Efficiency and Value	302
a.	Universal Consensus over AI Billing.....	302
b.	Billing for AI Costs and Learning Time	306
5.	Rule 1.2 (Scope of Representation): Allocating Authority Between Lawyer, Client, and AI Agent	308

a.	Impact of Agentic AI on the Scope of Representation.....	308
b.	Delegating the “Means” to a Non-Human Non-Lawyer	309
B.	The Lawyer As Advocate: Candor, Fairness, and the AI Co-Counsel	310
1.	Rule 3.1 (Meritorious Claims) & Rule 3.3 (Candor Toward the Tribunal): The Specter of AI “Hallucinations” and the Inviolable Duty of Verification	311
a.	The Epicenter of AI Misconduct	311
b.	The Unanimous Ethical Mandate Regarding Candor and Verification	312
2.	Rule 3.4 (Fairness to Opposing Party and Counsel): AI in Discovery, Negotiations, and Strategic Analysis	313
a.	AI in E-Discovery	314
b.	AI in Negotiation.....	314
C.	Law Firms and Associations: The Imperative of Supervision	316
1.	Rules 5.1 & 5.3 (Supervision): Supervising the “Digital Associate” and Third-Party AI Vendors	316
a.	AI Systems and Agents That Provide “Nonlawyer Assistance”	316
b.	Firm-Level Responsibility	318
c.	Supervising AI Vendors	318
2.	Rule 5.5 (Unauthorized Practice of Law): Identifying the Line an Agentic AI Must Not Cross.....	320

a.	The Unauthorized Practice of Law Frontier.....	320
b.	When Does an AI Agent Cross the Line?	321
D.	Maintaining the Integrity of the Profession...	323
1.	Rule 8.4 (Misconduct): Deceit, Dishonesty, and Unwitting Reliance on Flawed AI.....	323
a.	Unwitting Deceit	324
b.	Bias as Misconduct	324
2.	Rule 7.1 (Communications Concerning a Lawyer's Services): Ensuring Truthfulness in AI-Generated Marketing.	326
a.	AI-Generated Hype	326
b.	Misleading Claims About AI Itself	326
	Table 1: Comparative Analysis of Key State Bar Guidelines on Generative AI Use.....	327
III.	Upholding the Model Code of Judicial Conduct with AI.....	330
A.	Canon 1: Upholding the Independence, Integrity, and Impartiality of the Judiciary.....	331
1.	The Appearance of Impropriety and □Black Box□ Algorithms.....	331
2.	Independence from Undue Influence.....	332
B.	Canon 2: Performing Judicial Duties Impartially, Competently, and Diligently.....	333
1.	Rule 2.2 (Impartiality) & Rule 2.3 (Bias): Confronting Algorithmic Bias in Judicial Decision-Support Tools.....	333
a.	The Problem of Biased Data.....	333
b.	The <i>Loomis</i> Precedent.....	334

2.	Rule 2.5 (Competence and Diligence): The Judge's Duty of Technological Competence.....	335
a.	Judicial Competence with AI.....	335
b.	Diligence in Application.....	336
3.	Rule 2.9 (Ex Parte Communications): Can an AI be an Improper External Influence?	336
a.	AI as an External Influence.....	336
IV.	Conclusion: Charting a Course for Ethical AI Integration in the American Legal System.....	338
A.	Key Recommendations for Practitioners, Firms, and the Judiciary	338
B.	The Future of Legal Ethics, Self-Regulation, and Legal Education	340

I. INTRODUCTION: THE NEW DIGITAL FRONTIER OF LEGAL PRACTICE

The legal profession, historically characterized by its slow pace of change, is now at the inflection point of a technological revolution.¹ This new digital frontier is defined by two distinct but related forms of AI, each with its own set of capabilities and ethical implications: the content-creating multimodal generative AI (GenAI)² and the task-performing agentic AI,³ which both present profound ethical and professional challenges for lawyers and judges.⁴ The integration of these artificial intelligence systems into the American legal system and the American practice of law represents a paradigm shift, introducing efficiencies and capabilities that were unimaginable a decade ago.⁵ The rapid proliferation of sophisticated artificial intelligence tools promises to reshape the practice of law, the administration of justice, and the very nature of legal work.⁶

This Article provides a comprehensive analysis of these challenges, structured through the frameworks of the American Bar Association's (ABA) Model Rules of Professional Conduct (MRPC)⁷ and the Model Code of Judicial Conduct (MCJC).⁸ The core finding of this analysis is that while

1. See Marjorie Richter, *How AI Is Transforming the Legal Profession*, THOMSON REUTERS (Aug. 18, 2025), <https://legal.thomsonreuters.com/blog/how-ai-is-transforming-the-legal-profession/> [https://perma.cc/E4C7-CY79] (reporting a statistical increase in use of AI within the legal field).

2. Shreyas V. Tawalare, Vaishnav A. Gudadhe & Sarthak C. Nimbhorkar, *How Generative AI is Revolutionizing the Future of Content*, 8 INT'L J. MULTIDISCIPLINARY RSCH. SCI., ENGG., & TECH. 1138, 1141 (2025).

3. Shomit Ghose, *The Next Next Big Thing: Agentic AI's Opportunities and Risks*, U.C. BERKELEY SUTARDJA CTR. FOR ENTREPRENEURSHIP & TECH. (Dec. 19, 2024), <https://scet.berkeley.edu/the-next-next-big-thing-agentic-ais-opportunities-and-risks> [https://perma.cc/N5TZ-2CNU] (defining Agentic AI as a group of AI agents designed to work as a team to solve problems).

4. *AI and Law: What Are the Ethical Considerations?*, CLIO, <https://www.clio.com/resources/ai-for-lawyers/ethics-ai-law> [https://perma.cc/5N9S-46AK].

5. John Villasenor, *How AI Will Revolutionize the Practice of Law*, BROOKINGS (Mar. 20, 2023), <https://www.brookings.edu/articles/how-ai-will-revolutionize-the-practice-of-law> [https://perma.cc/8SMQ-DGTQ].

6. A.B.A. Task Force on Law and Artificial Intelligence, *Addressing the Legal Challenges of AI*, A.B.A. (2025), https://www.americanbar.org/groups/centers_commissions/center-for-innovation/artificial-intelligence/impact-of-ai-on-the-legal-profession/ [https://perma.cc/YMU8-RGBK]; Villasenor, *supra* note 5; GARY MARCHANT & JOSEPH R. TIANO, JR., ARTIFICIAL INTELLIGENCE AND LEGAL ETHICS 1 (2023), <https://nacmnet.org/wp-content/uploads/AI-and-Legal-Ethics-Final-White-Paper.pdf> [https://perma.cc/UF4Z-857H].

7. MODEL RULES OF PROFESSIONAL CONDUCT (AM. BAR ASS'N 2026).

8. MODEL CODE OF JUDICIAL CONDUCT (AM. BAR ASS'N 2026).

AI presents novel and complex scenarios, the foundational principles of legal ethics— attorney competence, confidentiality, candor, diligence, and judicial impartiality— remain the primary and sufficient guideposts for professional conduct. The challenge lies not in a wholesale amendment or replacement of these rules, but in their diligent and nuanced application to a new technological reality.

Consider the following scenarios and the MRPC and MCJC rules that are implicated by an attorney’s or judge’s use of AI or by the actions of an agentic AI itself:

AI Scenario	Ethical Duties Implicated	Rules Implicated
An attorney inputs confidential client details into a free AI tool and fails to verify or correct the AI’s “hallucinated” legal theories regarding the claims stated in the AI’s output before filing the document in court.	Competence; Confidentiality; Meritorious Claims; Candor Toward the Tribunal	MRPC 1.1, 1.6, 3.1, 3.3
An attorney deploys an agentic AI to negotiate a client’s settlement with a human opposing counsel without disclosing the use of AI and obtaining informed consent from the client.	Scope of Representation; Communication with a Client	MRPC 1.2, 1.4
An attorney uses AI to draft a brief in 1 hour but bills the client for 5 hours, the usual time it would have taken the attorney to draft the brief without using AI.	Reasonable Attorney Fees	MRPC 1.5

A firm's junior lawyer uses an inadequately supervised AI agent for e-discovery. Due to algorithmic bias, the AI agent systematically hides responsive documents, unfairly prejudicing the opposing party. The firm had no AI governance and oversight policies for both the AI and the junior lawyer.	Fairness to Opposing Party and Counsel; Responsibilities of Partners, Managers, and Supervisory Lawyers; Responsibilities Regarding Nonlawyer Assistance	MRPC 3.4, 5.1, 5.3
A firm's unsupervised agentic AI chatbot offers specific legal advice to a new client that called the firm, performing core lawyer functions without oversight.	Unauthorized Practice of Law; Responsibilities Regarding Nonlawyer Assistance	MRPC 5.3, 5.5
A law firm's AI-generated ad falsely promises "guaranteed wins" using "proprietary, advanced AI," when it is merely using a free public-facing AI tool.	Communications Concerning a Lawyer's Services; Misconduct	MRPC 7.1, 8.4
A judge uploads a case summary to a public-facing AI and asks it for a recommendation on a child custody determination and accepts the AI's suggestion without notice to or input from the parties and without knowing how the "Black Box" AI technology made its recommendation.	Upholding the Independence, Integrity, and Impartiality of the Judiciary; Appearance of Impropriety; Performing Judicial Duties Impartially, Competently, and Diligently; Ex Parte Communications; possibly Impartiality; Bias	MCJC Canon 1 and Rule 1.2, Canon 2 and Rules 2.5, 2.9 Possibly MCJC Canon 2 and Rules 2.2, 2.3 (impartiality and bias)

My research finds that the lawyer's duty of competence has irrevocably expanded to include technological proficiency, including knowledge of the

proper and accurate uses of AI.⁹ The duty of confidentiality faces new vectors of risk from third-party AI platforms that may use uploaded client data for model training or from simpler “human error” scenarios when a less-experienced user of AI in an organization might accidentally or inadvertently obtain confidential information through the model’s retrieval augmented generation capabilities and later disclose it in AI generated outputs.¹⁰ The duties of candor and competence have been brought into sharp relief by high-profile sanctions against lawyers who submitted AI-generated “hallucinations” to courts.¹¹ The duty of supervision now also extends to the digital associates and third-party vendors that populate the AI ecosystem.¹²

9. MODEL RULES OF PROFESSIONAL CONDUCT R. 1.1 cmt 8 (AM. BAR ASS’N 2026); Robert J. Ambrogi, *In First Ethics Ruling on Gen AI, ABA Says Lawyers Must Have Reasonable Understanding of the Technology, But Need Not Become Experts*, LAWSITES (July 30, 2024), <https://www.lawnext.com/2024/07/in-first-ethics-ruling-on-gen-ai-aba-says-lawyers-must-have-reasonable-understanding-of-the-technology-but-need-not-become-experts.html> [https://perma.cc/H968-N7GP]. As the title of this Article states, the ABA does not go so far as to demand that attorneys become experts in the technology but must have a reasonable understanding of the capabilities and limitations of the specific generative AI technology the lawyer might use.

10. Jake Gray, *Artificial Intelligence Poses Threat to Business Data Privacy and Confidential Information*, IFRAH LAW (May 10, 2023), <https://www.ifrahlaw.com/ftc-beat/artificial-intelligence-poses-threat-to-business-data-privacy-and-confidential-information/> [https://perma.cc/EHW6-KZWX]; *Retrieval-Augmented Generation (RAG) in the Legal Industry*, ACROPLANS (Feb. 10, 2024), <https://www.acroplans.com/retrieval-augmented-generation-rag-in-the-legal-industry> [https://perma.cc/N53N-SXLG]. Although not specifically framed as an ethical responsibility under the MRPC, the evidence law attorney-client communication privilege also is directly implicated by generative and agentic AI systems if communications are shared or disclosed (uploaded) to a third-party AI vendor, or if an agentic AI system further disseminates the communication beyond the attorney and the client and their directly related human associates and agents. Ismail Amin, *Client Beware: The Utilization of Artificial Intelligence Platforms and the Potential Waiver of Attorney-Client Privilege*, TALG LAW (Aug. 19, 2025), <https://talglaw.com/artificial-intelligence-platforms-and-the-potential-waiver/> [https://perma.cc/BBG2-3MRQ].

11. See *Mata v. Avianca, Inc.*, 678 F. Supp. 3d 443, 466 (S.D.N.Y. 2023) (sanctioning attorney for including fake, AI-generated legal citations in a filing); *Lacey v. State Farm Gen. Ins. Co.*, No. CV 24-05205, 2025 WL 1363069, at 3 (C.D. Cal. May 5, 2025) (sanctioning a K&L Gates attorney for the same conduct); *Wadsworth v. Walmart, Inc.*, 348 F.R.D. 489, 499 (D. Wyo. 2025) (sanctioning a Morgan & Morgan attorney for the same conduct).

12. Sierra Van Allen, *Do Lawyers Have a Duty to Supervise Legal AI?*, VLEX (Jun. 21, 2025), <https://vlex.com/blog/AI-Supervision> [https://perma.cc/6YNH-PL5W]; Grace Wynn, *Agentic AI Puts a New Twist on Attorney Ethics Obligations*, LAW360 (Sep. 12, 2025), <https://www.law360.com/projectfinance/articles/2385567/agentic-ai-puts-a-new-twist-on-attorney-ethics-obligations> [https://perma.cc/N8V6-QA5B].

For the judiciary, AI introduces the specter of algorithmic bias, which threatens the core tenets of impartiality and fairness.¹³ The use of AI in judicial decision-support roles requires judges to develop a new form of competence to understand and guard against the cognitive and systemic biases these tools can introduce.¹⁴

In this Article, a critical distinction is drawn between generative AI, which primarily impacts the *tasks* of legal practice,¹⁵ and the more advanced agentic AI, which affects professional *autonomy* and *judgment*.¹⁶ Agentic AI, with its capacity for independent action, pushes the boundaries of delegation, the unauthorized practice of law, and the very business model of the legal profession.¹⁷ Ultimately, this Article concludes that ethical integration of AI is possible but requires a proactive and vigilant stance from all legal professionals. It calls for robust firm governance, continuous education, a demand for transparency from technology providers, and an unwavering commitment to the principle that AI is a tool to augment, not replace, human professional judgment. The lawyer and the judge remain the ultimate guarantors of justice, and their ethical obligations are paramount, irrespective of the tools they employ.

13. See Andrej Kr̄tofik, *Bias in AI (Supported) Decision Making: Old Problems, New Technologies*, 16 INT'L J. CT. ADMIN. 1, 3-6 (2025) (providing examples of ways that the use of AI can impact impartiality in a court's decision).

14. See Jerry M. Gewirtz, *Artificial Intelligence May Assist, but Can Never Replace, the Judicial Decision-Making Process of Human Judges*, 98 FLA. BAR J. 8 (2024) (referencing Chief Justice Roberts's urge to approach AI with caution and humility).

15. Kim Martineau, *What Is Generative AI?*, IBM (Apr. 20, 2023), <https://research.ibm.com/blog/what-is-generative-ai> [<https://perma.cc/LE2Z-NZLL>].

16. See *What Is Agentic AI?*, AWS, <https://aws.amazon.com/what-is/agentic-ai> [<https://perma.cc/H8L9-4TF5>] (explaining how agentic AI can act autonomously and with independent judgment); Molly Hayes & Amanda Downie, *AI and the Future of Work*, IBM, <https://www.ibm.com/think/insights/ai-and-the-future-of-work> [<https://perma.cc/Q84P-JYWE>] (explaining advances in AI and its influence on decision-making).

17. See Andrew Fletcher, *The Key to Autonomous Legal Workflows with Agentic AI*, THOMSON REUTERS (Jun. 5, 2025), <https://legal.thomsonreuters.com/blog/the-key-to-autonomous-legal-workflows-with-agentic-ai> [<https://perma.cc/B2LV-SVWN>] (providing examples of tasks AI can conduct for attorneys, as well as how it can help "enhance" decision making, both of which may implicate professional conduct rules).

A. *Defining the Tools: From Generative Assistants to Autonomous Agents*

Understanding the impact of AI on legal ethics requires a clear distinction between the primary technologies driving this transformation.

Generative AI (GenAI) refers to a class of AI models designed to create new content.¹⁸ These systems, such as OpenAI's ChatGPT, Google's Gemini, and legal-specific platforms like CoCounsel and Lexis+AI, are trained on vast datasets of text and other information.¹⁹ In response to user prompts, they can generate human-like text, summarize documents, draft correspondence, and even produce initial drafts of legal arguments.²⁰ GenAI functions as a powerful assistant, augmenting the lawyer's ability to perform specific tasks more efficiently. However, these tools do not act independently; they are responsive instruments that require direct human input and oversight.²¹ The primary ethical challenges associated with GenAI revolve around the accuracy of its output, the confidentiality of the data input into it, and the proper billing for the efficiencies it creates.²²

Agentic AI represents a significant leap in autonomy and capability. These are more advanced, semi-autonomous systems designed not just to create content, but to *act* in pursuit of a goal.²³ An AI agent can break down

18. Martineau, *supra* note 15.

19. George Lawton, *What Is GenAI? Generative AI Explained*, TECHTARGET (Mar. 13, 2025), <https://www.techtarget.com/searchenterpriseai/definition/generative-AI> [https://perma.cc/X3XW-5B5P].

20. *See generally* Andrea Bucher, *Navigating the Power of Artificial Intelligence in the Legal Field*, 62 HOUS. L. REV. 819 (2025) (listing tasks AI can conduct and how these could be applicable in a law firm); Brent Farese, *How Generative AI For Legal Documents Is Changing Workflows*, ALINE (May 27, 2025), <https://www.aline.co/post/generative-ai-for-legal> [https://perma.cc/S5KQ-V9BQ] (noting GenAI's ability to produce draft and summarize legal documents).

21. *See* Natalie Pierce & Stephanie Goutos, *Why Lawyers Must Responsibly Embrace Generative AI*, 21 BERKELEY BUS. L. J. 469, 473-74, 478-90 (2024) (emphasizing the necessity that lawyers must understand AI through familiarization); Daniel Gall, *AI Is Your Nonstop Legal Assistant*, A.B.A. (Apr. 01, 2025), <https://www.americanbar.org/groups/gpsolo/resources/magazine/2025-mar-apr/ai-your-nonstop-legal-assistant/> [https://perma.cc/43SF-GTZC] (clarifying that AI responds to prompts).

22. Bryce Riddle & Aram Desteian, *Navigating Ethical and Regulatory Challenges of Generative AI in Law Firms*, ATTŶ AT LAW MAG. (Aug. 19, 2024), <https://attorneyatlawmagazine.com/practice-management/legal-ethics/navigating-ethical-and-regulatory-challenges-of-generative-ai-in-law-firms> [https://perma.cc/EK3N-FFLV].

23. Anna Gutowska, *What Are AI agents?*, IBM, <https://www.ibm.com/think/topics/ai-agents> [https://perma.cc/74V4-92CE]; *see generally* Andrew Zhao et al., *ExpeL: LLM Agents Are Experiential*

a complex objective into a sequence of tasks, interact with other software systems, gather new information, reason about its next steps, and execute multi-step workflows with limited human intervention.²⁴ For example, a lawyer could task an agentic AI with “researching precedents in California regarding the duty of care for software developers,” and the agent could formulate a research plan, execute searches across multiple databases, synthesize the findings, and draft a memorandum, all while self-correcting and adapting its approach.²⁵ This evolution from a responsive tool to a proactive agent introduces a more profound set of ethical questions concerning delegation of professional judgment, supervision, accountability for autonomous actions, and the potential for enabling the unauthorized practice of law.²⁶

B. *The Enduring Framework: Applying Rules of Reason to a New Reality*

Despite the novelty of these technologies, the legal profession is not without a map. The ABA’s Model Rules of Professional Conduct and Model Code of Judicial Conduct were designed as “rules of reason.”²⁷ Their purpose is to provide a framework for the ethical practice of law, interpreted with reference to the fundamental purposes of legal representation and the law itself.²⁸ This inherent flexibility allows them to be applied to new and unforeseen circumstances, including those presented by AI.²⁹

Learners, 38 PROCEEDINGS OF THE AAAI CONF. ON A.I. 19632, 19634 (2024) (detailing empirical research regarding the efficacy of learning models).

24. Allie Mistakidis, *Agentic Workflows: How Autonomous AI Executes Complex Tasks*, TRIPLE WHALE (July 7, 2025), <https://www.triplewhale.com/blog/agentic-workflows> [https://perma.cc/7BN6-BX3H].

25. *See id.* (describing the autonomous nature of agentic AI and how it can formulate plans while learning from its mistakes); Anirban Mukherjee & Hannah Hanwen Chang, *Stochastic, Dynamic, Fluid Autonomy in Agentic AI: Implications for Authorship, Inventorship, and Liability*, ARXIV (Apr. 5, 2025), <https://arxiv.org/html/2504.04058v1> [https://perma.cc/Y9R9-EEH6] (describing agentic AI’s ability to handle and solve complex issues).

26. Riddle & Desteian, *supra* note 22; Olga Mack, *How Lawyers Can Embrace the Challenge to Supervise AI*, LEGAL DIVE (July 12, 2023), <https://www.legaldive.com/news/ai-in-legal-practice-ai-supervision-in-house-counsel-legal-ethics/686578/> [https://perma.cc/2NAV-45UH].

27. MODEL RULES OF PROFESSIONAL CONDUCT Preamble ¶ 14 (AM. BAR ASS’N 2026).

28. *Id.* ¶ 9.

29. N.J. ST. BAR, TASK FORCE ON ARTIFICIAL INTELLIGENCE (AI) AND THE LAW: REPORT, REQUESTS, RECOMMENDATIONS, AND FINDINGS (2024); *Do We Need to Amend the Rules of Professional Conduct in Response to the Legal Profession’s Embrace of AI? In My Opinion, No*, ETHICAL GROUNDS (May 16, 2024), <https://vtbarcounsel.wordpress.com/2024/05/16/do-we-need-to-amend-the-rules-of->

The ABA itself has affirmed this approach. In its landmark Formal Opinion 512, *Generative Artificial Intelligence Tools*,³⁰ the Standing Committee on Ethics and Professional Responsibility did not propose new rules. Instead, it systematically applied the existing Model Rules to the challenges posed by AI, demonstrating that the duties of competence, confidentiality, communication with clients, meritorious claims and candor, reasonable fees, and supervision provide a robust framework for governing the use of these new tools.³¹ State bar associations in California,³² Florida,³³ New York,³⁴ Kentucky,³⁵ and elsewhere³⁶ have followed this same interpretive model, issuing guidance that grounds AI ethics in established professional obligations.

Therefore, the central thesis of this Article is that navigating the AI frontier does not require a wholesale reinvention of legal ethics. Rather, it demands a rigorous, thoughtful, and continuous process of interpretation and application. It requires lawyers and judges to understand the capabilities and limitations of these powerful tools and to ensure their use remains consistent with the timeless principles that underpin the integrity of the legal profession and the administration of justice.

II. NAVIGATING THE MODEL RULES OF PROFESSIONAL CONDUCT WITH AI

The introduction of generative and agentic AI into the daily practice of law directly engages the core duties that govern the client-lawyer

professional-conduct-in-response-to-the-legal-professions-embrace-of-ai-in-my-opinion-no [https://perma.cc/CF6F-2RA4].

30. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512 (2024).

31. *Id.*

32. THE STATE BAR OF CALI. STANDING COMM. ON PRO. RESP. & CONDUCT, PRACTICAL GUIDANCE FOR THE USE OF GENERATIVE ARTIFICIAL INTELLIGENCE IN THE PRACTICE OF LAW (2023) [hereinafter CAL. BAR PRACTICAL GUIDANCE].

33. Fla. Bar Ethics Op. 24-1 (2024).

34. N.Y. STATE BAR ASS'N, REPORT AND RECOMMENDATIONS OF THE NEW YORK STATE BAR ASSOCIATION TASK FORCE ON ARTIFICIAL INTELLIGENCE (2024) [hereinafter NYSBA REPORT AND RECOMMENDATION], <https://www.nysba.org/wp-content/uploads/2022/03/2024-April-Report-and-Recommendations-of-the-Task-Force-on-Artificial-Intelligence.pdf> [https://perma.cc/S3NL-JW8V].

35. Ky. Bar Ass'n Ethics Op. KBA E-457 (2024).

36. See *infra* Table 1 (containing a complete list of state bar ethics opinions and guidance pertaining to AI).

relationship, the lawyer's role as an advocate, and the structure of law firm management. This Part examines these impacts through the lens of the Model Rules of Professional Conduct.

A. *The Client-Lawyer Relationship: Core Duties Reimagined*

The foundational duties of competence, confidentiality, communication, and reasonable fees are profoundly tested and reshaped by the integration of AI tools.

1. Rule 1.1 (Competence): Beyond Legal Knowledge to Technological Proficiency

Model Rule 1.1 mandates that "A lawyer shall provide competent representation to a client," which requires the "legal knowledge, skill, thoroughness and preparation reasonably necessary for the representation."³⁷ For over a decade, this duty has been evolving. The Comment to Rule 1.1 explicitly clarifies that maintaining competence requires a lawyer to "keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology."³⁸ With the advent of powerful AI, this technological component of competence has moved from a peripheral consideration to a central professional obligation.³⁹

In the context of generative AI, competence is no longer merely about knowing that these tools exist. It demands a reasonable and practical understanding of their capabilities and, more critically, their inherent limitations.⁴⁰ The most well-documented risk is the phenomenon of hallucinations—the tendency of large language models to generate outputs

37. MODEL RULES OF PROFESSIONAL CONDUCT R. 1.1 (AM. BAR ASS'N 2026).

38. *Id.* cmt. 8.

39. John M. Nolan et al., *What You Need to Know Today About Ethical Use of Generative AI in the Practice of Law*, ASSOC. OF CORP. COUNSEL (Sep. 23, 2024), <https://www.acc.com/resource-library/what-you-need-know-today-about-ethical-use-generative-ai-practice-law> [https://perma.cc/U6RF-M3NF].

40. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 2 (2024); Sergio Brotons, *The Limitations of Generative AI, According to Generative AI*, LINGARO GROUP, <https://lingarogroup.com/blog/the-limitations-of-generative-ai-according-to-generative-ai> [https://perma.cc/39JP-RM36].

that are fluent, plausible, and entirely false.⁴¹ The legal profession has been served a number of stark warnings on this front. The seminal case of *Mata v. Avianca, Inc.*⁴² saw attorneys sanctioned after submitting a legal brief replete with non-existent cases and fabricated citations generated by ChatGPT. The lawyers' failure to verify the AI's output was a clear breach of their duty of competence. Similar sanctions imposed on the law firms K&L Gates⁴³ and Morgan & Morgan⁴⁴ for relying on "bogus AI-generated research" underscore that this is a systemic risk, not an isolated incident.⁴⁵ These cases demonstrate that the duty of competence now includes the non-delegable responsibility to independently verify the factual and legal accuracy of any AI-generated work product.

As the profession moves toward using agentic AI, the demands of competence will escalate further. A lawyer deploying a semi-autonomous agent must understand not only its potential outputs but also its operational logic, the data sources it relies upon, and the predefined boundaries of its delegated authority.⁴⁶ To competently use an AI agent for a task like

41. Daniel E. Ho et. al., *Hallucinating Law: Legal Mistakes with Large Language Models Are Pervasive*, STAN. U. HUM.-CENTERED A.I. (Jan 11, 2024), <https://hai.stanford.edu/news/hallucinating-law-legal-mistakes-large-language-models-are-pervasive> [<https://perma.cc/CP57-4WVK>].

42. See *Mata v. Avianca, Inc.*, 678 F. Supp. 3d 443, 466 (S.D.N.Y. 2023) (sanctioning an attorney for including fake, AI-generated legal citations in a filing).

43. *Lacey v. State Farm Gen. Ins. Co.*, No. 2:24-cv-05205, 2025 WL 1363069, at *1 (C.D. Cal. May 6, 2025); Debra Cassens Weiss, *Sanctions Imposed for "Collective Debacle" Involving AI Hallucinations and 2 Firms, Including K&L Gates*, A.B.A. J. (May 14, 2025, at 12:50 CDT), <https://www.abajournal.com/web/article/judge-imposes-sanctions-for-collective-debacle-involving-ai-hallucinations-and-2-law-firms-including-k> [<https://perma.cc/8VFG-TCW5>].

44. *Wadsworth v. Walmart, Inc.*, 348 F.R.D. 489, 489 (D. Wyo. 2025); Bernie Pazanowski, *Morgan & Morgan Lawyers Fined for "Hallucinated" AI Citations*, BLOOMBERG LAW (Feb. 25, 2025, 8:42 AM CST), <https://news.bloomberglaw.com/litigation/morgan-morgan-lawyers-fined-for-hallucinated-ai-citations> [<https://perma.cc/8UX4-PT38>].

45. See Roland Moore-Colyer, *AI Hallucinates More Frequently as it Gets More Advanced: Is There Any Way to Stop It From Happening, and Should We Even Try?*, LIVE SCIENCE (Jun. 21, 2025), <https://www.livescience.com/technology/artificial-intelligence/ai-hallucinates-more-frequently-as-it-gets-more-advanced-is-there-any-way-to-stop-it-from-happening-and-should-we-even-try> [<https://perma.cc/6YUR-WG94>] (highlighting that AI systems appear to hallucinate more as time progresses).

46. Pamela Langham, *A Lawyer's Primer on AI Agents*, MD. STATE BAR ASSOC. (May 14, 2025), <https://www.msba.org/site/site/content/News-and-Publications/News/General-News/A-Lawyers-Primer-on-AI-Agents.aspx> [<https://perma.cc/94LQ-8R29>]; Daniel J. Siegel, *Joint Formal Ethics Opinion Gives Practical Guidance on Artificial Intelligence*, PHILA. BAR. ORG. (Aug. 8, 2024), <https://philadelphiabar.org/?pg=ThePhiladelphiaLawyerBlog&blAction=showEntry&blogEntry=111283> [<https://perma.cc/Q8QX-IDUQ>].

contract analysis, a lawyer must understand how the agent identifies clauses, what constitutes a "risk" according to its programming, and what might cause it to fail or produce an erroneous result. Deploying a complex tool without this fundamental understanding would be a clear violation of Rule 1.1. This expanded definition of competence is a point of clear consensus among regulators. ABA Formal Opinion 512 states that lawyers must understand AI's capabilities and limitations.⁴⁷ State bar guidelines from California, Florida, Kentucky, and New York echo this mandate.⁴⁸

A competent knowledge of the workings of generative AI systems includes understanding the crucial difference between "non-privacy-protecting" systems, which may use user inputs for training without prior notice or permission, and "locked down" or "enterprise-grade" systems designed for legal practice that prioritizes data security. It also includes a duty to properly vet AI vendors to ensure their technology is reliable and secure.⁴⁹ Unfortunately, the "non-privacy-protecting" models might also be called "normal" models, as it is most common for AI models to use user inputted information rather indiscriminately, which makes the distinction all the more important to be cognizant of.⁵⁰

47. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 2 (2024).

48. CAL. BAR PRACTICAL GUIDANCE, *supra* note 32; Fla. Bar Ethics Op. 24-1 (2024); NYSBA REPORT AND RECOMMENDATION, *supra* note 34; Ky. Bar Ass'n Ethics Op. KBA E-457 (2024).

49. Juliet Luciol, *The AI Vendor Evaluation Checklist Every Leader Needs*, VKTR (Apr. 2, 2025), <https://www.vktr.com/digital-workplace/the-ai-vendor-evaluation-checklist-every-leader-needs/> [<https://perma.cc/2CMV-M2VX>]; *Vetting Your AI Vendors: Key Takeaways for Business Leaders*, BATOI (May 12, 2025), <https://www.batoi.com/blogs/digital-transformation/vetting-ai-vendors-key-takeaways-business-leaders> [<https://perma.cc/F1L2-2F2H>].

50. Lance Eliot, *Generative AI ChatGPT Can Disturbingly Gobble Up Your Private and Confidential Data, Forewarns AI Ethics and AI Law*, FORBES (Jan. 27, 2023, at 08:00 EST), <https://www.forbes.com/sites/lanceeliot/2023/01/27/generative-ai-chatgpt-can-disturbingly-gobble-up-your-private-and-confidential-data-forewarns-ai-ethics-and-ai-law/?sh=3cac963d7fdb> [<https://perma.cc/B28V-2JPV>].

Non-Privacy-Protecting Models (default = uses your data)	Privacy-Protecting or Enterprise Models (default = locked down)
OpenAI ChatGPT (consumer / free / Plus, GPT-5) <input type="checkbox"/> by default your chats <input type="checkbox"/> help improve our models <input type="checkbox"/> ; toggle in <i>Settings</i> → <i>Data Controls</i> to disable. ⁵¹	OpenAI ChatGPT Enterprise/ Team <input type="checkbox"/> <input type="checkbox"/> We never train on your data. <input type="checkbox"/> ⁵²
Google Gemini (consumer apps) <input type="checkbox"/> <input type="checkbox"/> Google uses this data to . . . [improve] machine-learning technologies <input type="checkbox"/> ; <i>Gemini Apps Activity</i> is on by default. ⁵³	Anthropic Claude (free, Pro, API) <input type="checkbox"/> <input type="checkbox"/> We will not use your inputs or outputs to train . . . unless you explicitly opt in. <input type="checkbox"/> ⁵⁴
Microsoft Copilot <input type="checkbox"/> By default it uses your data for training. Opt out is possible, but the option is buried. Click your profile image > your name > Privacy and disable <input type="checkbox"/> Model training on text. <input type="checkbox"/> ⁵⁵	Microsoft Azure OpenAI Service <input type="checkbox"/> prompts and completions <input type="checkbox"/> are NOT used to train . . . foundation models. <input type="checkbox"/> ⁵⁶

51. Jared Newman, *Using ChatGPT or Other AI Tools? Here's Who Can See Your Chat History*, FAST CO. (Aug. 4, 2025), <https://www.fastcompany.com/91379336/using-chatgpt-or-other-ai-tools-heres-who-can-see-your-chat-history> [https://perma.cc/EMS8-3QK9].

52. *Frontier AI Built for Enterprise*, CHATGPT, <https://openai.com/chatgpt/enterprise/> [https://perma.cc/N3KG-76R8].

53. *Gemini Apps Privacy Hub*, GOOGLE (July 24, 2025), <https://support.google.com/gemini/answer/13594961?hl=en> [https://perma.cc/66K9-JFUT].

54. *Is My Data Used for Model Training?*, ANTHROPIC, <https://privacy.anthropic.com/en/articles/10023580-is-my-data-used-for-model-training> [https://perma.cc/32BY-QJ97].

55. *Data, Privacy, and Security for Azure Direct Models in Microsoft Foundry*, MICROSOFT LEARN, <https://learn.microsoft.com/en-us/azure/ai-foundry/responsible-ai/openai/data-privacy?view=foundation-classic&tabs=azure-portal> [https://perma.cc/9DDC-T4CL].

56. *Data, Privacy, and Security for Azure Direct Models in Microsoft Foundry e*, MICROSOFT LEARN, <https://learn.microsoft.com/en-us/azure/ai-foundry/responsible-ai/openai/data-privacy?tabs=azure-portal> [https://perma.cc/D84Y-ETNU].

<p>xAI Grok (consumer) <input type="checkbox"/> <input type="checkbox"/> By default: uses your data to train AI <input type="checkbox"/> with an opt-out toggle.⁵⁷</p>	<p>Amazon Bedrock <input type="checkbox"/> <input type="checkbox"/> AWS and the third-party model providers will not use any inputs or outputs from Amazon Bedrock to train Amazon Nova, Amazon Titan, or any third-party models.<input type="checkbox"/>⁵⁸</p>
--	---

This evolution signifies that the duty of competence is no longer a static body of knowledge but a dynamic, ongoing obligation. The rapid pace at which AI technology evolves means that a lawyer's understanding must evolve with it. A lawyer who was competent in using basic GenAI tools in 2024 may be deemed incompetent in 2026 if they have not educated themselves on the risks and benefits of the more advanced agentic systems that will have become prevalent. This reality strongly suggests a future need for mandatory continuing legal education (CLE) focused specifically on legal technology to ensure the bar as a whole maintains this dynamic competence.

2. Rule 1.6 (Confidentiality of Information): Protecting Client Data in the Age of AI Models and Training Protocols

Model Rule 1.6 establishes one of the most sacred duties in the legal profession:

- (a) A lawyer shall not reveal information relating to the representation of a client unless the client gives informed consent, the disclosure is impliedly authorized in order to carry out the representation or the disclosure is permitted by paragraph (b).
- (b) A lawyer may reveal information relating to the representation of a client to the extent the lawyer reasonably believes necessary:
 - (1) to prevent reasonably certain death or substantial bodily harm;

57. Lance Eliot, *Generative AI ChatGPT Can Disturbingly Gobble Up Your Private and Confidential Data, Forewarns AI Ethics and AI Law*, FORBES (Jan. 27, 2023, at 08:00 EST), <https://www.forbes.com/sites/lanceeliot/2023/01/27/generative-ai-chatgpt-can-disturbingly-gobble-up-your-private-and-confidential-data-forewarns-ai-ethics-and-ai-law/?sh=3cac963d7fdb> [<https://perma.cc/B28V-2JPV>].

58. *Amazon Bedrock FAQs*, AWS (2025), <https://aws.amazon.com/bedrock/faqs/> [<https://perma.cc/W4ZD-RWZ7>].

(2) to prevent the client from committing a crime or fraud that is reasonably certain to result in substantial injury to the financial interests or property of another and in furtherance of which the client has used or is using the lawyer's services;

(3) to prevent, mitigate or rectify substantial injury to the financial interests or property of another that is reasonably certain to result or has resulted from the client's commission of a crime or fraud in furtherance of which the client has used the lawyer's services;

(4) to secure legal advice about the lawyer's compliance with these Rules;

(5) to establish a claim or defense on behalf of the lawyer in a controversy between the lawyer and the client, to establish a defense to a criminal charge or civil claim against the lawyer based upon conduct in which the client was involved, or to respond to allegations in any proceeding concerning the lawyer's representation of the client;

(6) to comply with other law or a court order; or

(7) to detect and resolve conflicts of interest arising from the lawyer's change of employment or from changes in the composition or ownership of a firm, but only if the revealed information would not compromise the attorney-client privilege or otherwise prejudice the client.

(c) A lawyer shall make reasonable efforts to prevent the inadvertent or unauthorized disclosure of, or unauthorized access to, information relating to the representation of a client.⁵⁹

The use of third-party generative AI platforms creates a direct tension with this duty. The primary threat arises from "non-privacy-protecting" AI systems, such as free, publicly available versions of tools like ChatGPT, which may use data from user prompts to further train their models.⁶⁰ When a lawyer inputs a query containing confidential client information—facts about a case, strategic considerations, or identifying details—they risk that information being absorbed into the AI model and potentially disclosed to other users or used in ways that violate the duty of confidentiality.⁶¹ This

59. MODEL RULES OF PROFESSIONAL CONDUCT R. 1.6 (AM. BAR ASS'N 2026).

60. Eliot, *supra* note 50.

61. Nicole Engler, *Using AI In Legal Work: COPRAC's Tips on Confidentiality and Competence*, BAR ASS'N OF S.F.: JUST. & DIVERSITY CTR. (May 24, 2024), <https://www.sfbbar.org/blog/using-ai-in-legal-work-copracs-tips-on-confidentiality-and-competence/> [<https://perma.cc/SBJ2-T3E9>].

is not a theoretical risk; it is a fundamental aspect of how many large language models learn and improve.⁶²

In response to this clear and present danger, a strong consensus on best practices has emerged from ethics opinions issued by the ABA and leading state bars. To comply with Rule 1.6, lawyers must adopt a multi-layered approach to risk mitigation:

Vet the Vendor: Lawyers must read and understand the terms of service, privacy policies, and data security protocols of any AI provider before using the tool for client matters. This due diligence is a prerequisite for competent and confidential use.⁶³

Prefer Privacy-Protecting Systems: Whenever possible, lawyers should use AI tools designed specifically for the legal profession or enterprise-grade versions that contractually guarantee client data is kept confidential, is not used for model training, and is housed in a secure, closed environment.⁶⁴

Anonymize Data: If using a system where confidentiality cannot be absolutely guaranteed, lawyers have a duty to anonymize client information before inputting it. This means removing all names, locations, specific dates, and any other details that could identify the client or the matter.⁶⁵

Obtain Informed Consent: In situations where the use of a particular AI tool might involve the disclosure of confidential information to a third party, the lawyer must obtain the client's specific, informed consent. As ABA Formal Opinion 512 and other guidance make clear, a generic, boilerplate consent clause in a standard engagement letter is likely insufficient to meet this standard.⁶⁶ The client must be made aware of the specific risks involved.

62. *See id.* (explaining that many AI systems improve through ongoing interaction with user inputs, such that information entered into AI tools may be retained, processed, or reused by the model or vendor, making risks to client confidentiality inherent rather than merely hypothetical).

63. Juliet Luciol, *The AI Vendor Evaluation Checklist Every Leader Needs*, VKTR (Apr. 2, 2025), <https://www.vktr.com/digital-workplace/the-ai-vendor-evaluation-checklist-every-leader-needs/> [<https://perma.cc/2CMV-M2VX>]; *Vetting Your AI Vendors: Key Takeaways for Business Leaders*, BATOI (May 12, 2025), <https://www.batoi.com/blogs/digital-transformation/vetting-ai-vendors-key-takeaways-business-leaders> [<https://perma.cc/JY4B-P5G2>].

64. Eliot, *supra* note 50.

65. Robbie Araiza, *Data Anonymization: Techniques for Protecting Privacy in Data Sets*, FORTRA (Apr. 7, 2025), <https://dataclassification.fortra.com/blog/data-anonymization-techniques-protecting-privacy-data-sets> [<https://perma.cc/LSZ3-SB32>].

66. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 708 (2024).

The advent of agentic AI introduces a new and more complex vector for the inadvertent disclosure of confidential information.⁶⁷ While the risk from generative AI is tied to a lawyer's direct action of *inputting* data, the risk from agentic AI stems from its capacity for *autonomous action*.⁶⁸ Consider an agent tasked with a high-level goal, such as "conduct due diligence on Target Company X" or "negotiate a settlement with opposing counsel's agent." In pursuing this goal, the agent might autonomously access the law firm's document management system, retrieve sensitive client files, and share them with an external system or another AI agent without a specific command from the lawyer.⁶⁹ This potential for unforeseen disclosure shifts the ethical burden. It is no longer enough to be careful about what one types into a prompt box.⁷⁰ Lawyers and firms must now engage in the technical and strategic work of programming precise "guardrails," prohibitions, and oversight mechanisms into the agent's operational parameters.⁷¹ This, in turn, creates a feedback loop to the duty of competence under Rule 1.1, as

67. David McInerney, *Privacy Unplugged: Tackling the Challenges of Agentic AI in Business*, SEC. INFO. WATCH (June 27, 2025), <https://www.securityinfowatch.com/ai/article/55300108/privacy-unplugged-tackling-the-challenges-of-agentic-ai-in-business> [<https://perma.cc/8VSG-Z5PH>].

68. See Phaedra Boinodiris & Jon Parker, *The Evolving Ethics and Governance Landscape of Agentic AI*, IBM (Mar. 21, 2025), <https://www.ibm.com/think/insights/ethics-governance-agentic-ai> [<https://perma.cc/WE4Q-9GA7>] (discussing the ethics surrounding agentic AI); Mack, *supra* note 26 (noting the risks of agentic AI and lawyers' role in overseeing it).

69. Deven R. Desai & Mark O. Riedl, *Responsible AI Agents*, ARXIV, 4:5 (Feb. 20, 2025), <https://arxiv.org/pdf/2502.18359> [<https://perma.cc/7ZGC-C3F6>]; Mark O. Reidl & Deven R. Desai, *AI Agents and the Law*, ARXIV, 1:2 (Aug. 12, 2025), <https://arxiv.org/pdf/2508.08544> [<https://perma.cc/XU2W-2J6M>].

70. N.C. State Bar Comm. on Ethics & Pro. Resp., Formal Op. 2024-1 (2024) (emphasizing that lawyers remain fully responsible for client data when using AI tools "including ensuring confidentiality and system security" not merely mindful prompts); see generally CAL. BAR PRACTICAL GUIDANCE, *supra* note 32 (stressing that a lawyer must critically review both input and AI output, anonymize sensitive data, and prevent inadvertent disclosures "not just rely on cautious prompting).

71. See A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 2:5 (2024) (advising that lawyers must understand GAI's capabilities and risks, secure confidentiality, and establish clear supervisory policies governing its use "akin to building 'guardrails' around the technology"); see generally *ABA Ethics Rules and Generative AI*, THOMSON REUTERS (Mar. 27, 2025), <https://legal.thomsonreuters.com/blog/generative-ai-and-aba-ethics-rules/> [<https://perma.cc/6H95-XMR4>] (reinforcing that lawyers must supervise AI assistants like human nonlawyer staff under Rules 5.1 and 5.3 and embed oversight mechanisms).

the ability to set these guardrails effectively is itself a form of technological skill.⁷²

3. Rule 1.4 (Communications): The Duty to Inform and Obtain Consent

Model Rule 1.4 requires a lawyer to (1) promptly inform the client of any decision or circumstance with respect to which the client's informed consent . . . is required . . . ; (2) reasonably consult with the client about the means by which the client's objectives are to be accomplished; (3) keep the client reasonably informed about the status of the matter.⁷³

a. When Is Disclosure Required?

The application of this rule to AI use is a subject of evolving debate, but a consensus on the core principles is forming. It is generally agreed that a lawyer does not need to inform a client about every minor or routine use of technology, such as employing a spell-checker or a grammar-assist tool embedded in a word processor.⁷⁴ However, ethics bodies like the ABA and various state bars suggest that disclosure, and in some cases explicit consent, becomes necessary under several conditions⁷⁵:

Confidentiality Risks: If a lawyer intends to use an AI tool that processes client data externally or otherwise poses a potential risk to confidentiality under Rule 1.6, communication with the client is essential.⁷⁶

Substantive Delegation: When AI is used for substantive legal tasks that a client would reasonably expect to be performed by a lawyer—such as legal research, drafting significant portions of a brief, or analyzing key evidence—

72. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 2 (2024) (citing Model Rule 1.1's comment that lawyers must understand the benefits and risks of relevant technology, thus including AI governance as part of technological competence); Andrew N. Choi & Chiara D. Kalogjera-Sackellares, *Generative AI and the Duty of Competence Conundrum*, HOLLAND & KNIGHT (Feb. 13, 2025), <https://www.hklaw.com/en/insights/publications/2025/02/generative-ai-and-the-duty-of-competence-conundrum> [<https://perma.cc/KY6B-3BU9>] (noting that under Rule 1.1 and Comment 8, lawyers must not only know how to use generative AI but also supervise its deployment with technological understanding).

73. MODEL RULES OF PROFESSIONAL CONDUCT R 1.4 (AM. BAR ASS'N 2026).

74. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 8–9 n.43 (2024).

75. *Id.*

76. *Id.* at 9.

disclosure is likely required. The client has a right to know about the means being used to achieve their objectives.⁷⁷

Cost Implications: If the use of a specific AI tool will result in a direct cost being passed on to the client, Rule 1.4, in conjunction with Rule 1.5 (Fees) mandates that this be communicated and agreed upon.⁷⁸

b. The Spectrum of Guidance from State Bars

While the principles are similar, the specific guidance from state bars varies in its prescriptive force.⁷⁹

Florida's Advisory Opinion 24-1 takes a firm stance, advising lawyers to obtain the client's informed consent before using any third-party generative AI program that involves the disclosure of confidential information.⁸⁰

California's Practical Guidance adopts a more flexible, context-dependent approach, recommending that lawyers *consider* disclosing their intent to use AI based on factors such as the technology's novelty, the associated risks, and the client's sophistication.⁸¹

New York's Task Force report suggests that lawyers should consider including a general disclosure in their engagement letters stating that AI technologies may be used during the representation.⁸²

77. *Id.*; Michael Cole, *From Code to Conduct: Ethical Considerations for AI in Legal Practice*, REUTERS (Aug. 13, 2024), <https://www.reuters.com/legal/legalindustry/code-conduct-ethical-considerations-ai-legal-practice-2024-08-13> [https://perma.cc/RS2M-89RB].

78. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 11 (2024); Mich. Bar Ass'n, *Artificial Intelligence for Attorneys' FAQs* (Nov. 18, 2024), <https://www.michbar.org/opinions/ethics/AIFAQs> [https://perma.cc/SH6V-LU9H].

79. *See, e.g.*, N.Y. City Bar Ass'n Comm. on Pro. Ethics, Formal Op. 2024-5 (2024) (acknowledging that, unlike more prescriptive bar guidance, New York's opinion intentionally provides guardrails and not hard-and-fast restrictions).

80. *See* Fla. Bar Ethics Op. 24-1 (2024) (mandating that a lawyer obtain the affected client's informed consent prior to utilizing a third-party generative AI program if the utilization would involve the disclosure of any confidential information).

81. *See* CAL. BAR PRACTICAL GUIDANCE, *supra* note 32, at 4 (endorsing a context-dependent approach, urging lawyers to consider disclosure based on novelty, risks, and client sophistication).

82. *See* NYSBA REPORT AND RECOMMENDATION, *supra* note 34, at 57 (recommending lawyers consider including wording in engagement letters to advise clients that AI tools may be used in the representation); Sara Merken, *New York Lawyers Urged to Use AI with Care in New State Bar Guidelines*, REUTERS (Apr. 8, 2024, at 19:33 CDT), <https://www.reuters.com/legal/transactional/new-york-lawyers-urged-use-ai-with-care-new-state-bar-guidelines-2024-04-08/> [https://perma.cc/AW4U-7G5B] (The New York task force also suggested that lawyers consider telling clients that AI tools could be used in their work, and the report provides a sample statement that could be included in a client engagement letter).

This duty to communicate, however, should be viewed as more than a defensive, risk-mitigation exercise. Clients hire lawyers for their expertise, diligence, and professional judgment. A client who discovers that significant portions of their case were handled by an AI tool without their knowledge may feel that the lawyer has abdicated their fundamental role, eroding trust. Therefore, proactive communication about AI can be a powerful tool for building and maintaining the client relationship. By framing the use of AI as a means to enhance efficiency, reduce costs, and deliver a more thorough work product—augmenting, rather than replacing, the lawyer's core judgment—practitioners can turn a potential point of concern into a demonstration of value and technological savvy.⁸³ This reframes the duty under Rule 1.4 from a mere obligation to a strategic opportunity.

4. Rule 1.5 (Fees): Ethical Billing for Efficiency and Value

Model Rule 1.5 states, "A lawyer shall not make an agreement for, charge, or collect an unreasonable fee or an unreasonable amount for expenses."⁸⁴ The introduction of AI, a technology whose primary benefit is a massive increase in efficiency, creates a direct and unavoidable tension with the legal profession's traditional hourly billing model.⁸⁵

a. Universal Consensus over AI Billing

The central ethical question under Rule 1.5 is how a lawyer should bill for work completed in a fraction of the time it would have taken to do manually. On this point, there is universal and unambiguous consensus from every ethics body that has addressed the issue, including the ABA in Formal Opinion 512 and state bars in California, Florida, Kentucky, Pennsylvania

83. See N.Y. City Bar Ass'n Comm. on Pro. Ethics, Formal Op. 2024-5 (2024) (emphasizing that lawyers should "take into account . . . the duty to consult with clients . . ." and maintain independent judgment when using AI); see also Marcin M. Krieger & David R. Cohen, *Navigating the Seven Cs of Ethical Use of AI by Lawyers*, REUTERS (Dec. 20, 2024), <https://www.reuters.com/legal/legalindustry/navigating-seven-cs-ethical-use-ai-by-lawyers-2024-12-20/> [<https://perma.cc/UZG2-5AG3>] (articulating client communication as part of the "Consent" and "Candor" in the "Seven Cs" framework).

84. MODEL RULES OF PROFESSIONAL CONDUCT R 1.5 (AM. BAR ASS'N 2026).

85. See generally Jonah E. Perlin, *How the Billable Hour Can Survive Generative AI*, 5 STETSON BUS. L. REV. 41 (2025) (discussing the effects on generative AI on the billable hour model).

(and Philadelphia), and Texas:⁸⁶ A lawyer cannot bill the client for the time saved by using AI.

Ethics Opinion	Description
<p>ABA Formal Op. 512 (2024)⁸⁷</p>	<p><input type="checkbox"/>GAI tools may provide lawyers with a faster and more efficient way to render legal services to their clients, but lawyers who bill clients an hourly rate for time spent on a matter must bill for their actual time. . . . <input type="checkbox"/>It nonetheless will not be permissible to charge the client for more hours than were actually expended on the matter.<input type="checkbox"/></p> <p><input type="checkbox"/>If using a GAI tool enables a lawyer to complete tasks much more quickly than without the tool, it may be unreasonable under Rule 1.5 for the lawyer to charge the same flat fee when using the GAI tool as when not using it.<input type="checkbox"/></p> <p><input type="checkbox"/>Absent an agreement, the firm may charge the client no more than the direct cost associated with the tool (if any) plus a reasonable allocation of expenses directly associated with providing the GAI tool. . . .<input type="checkbox"/></p>
<p>California: Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law (2023)⁸⁸</p>	<p><input type="checkbox"/>A lawyer may use generative AI to more efficiently create work product and may charge for actual time spent (e.g., crafting or refining generative AI inputs and prompts, or reviewing and editing generative AI outputs). A lawyer must not charge hourly fees for the time saved by using generative AI.<input type="checkbox"/></p>

86. See A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 8–13 n.43 (2024) (making clear that while lawyers may use generative AI to increase efficiency, they may not bill clients for time not actually expended, expressly rejecting any attempt to charge for “time saved” through the use of AI tools); Cole, *supra* note 77 (discussing transparency in billing); Mich. Bar Ass’n, *supra* note 78 (noting Rule 1.5 requires lawyers to charge clients for actual time expended).

87. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 12, 14 (2024).

88. CAL. BAR PRACTICAL GUIDANCE, *supra* note 32, at 4.

<p>Florida: Pro. Ethics Comm. Op. 24-1 (2024)⁸⁹</p>	<p>□ Use of generative AI does not permit a lawyer to engage in improper billing practices such as double-billing. . . . Lawyers are ultimately responsible for the work product that they create regardless of whether that work product was originally drafted or researched by a nonlawyer or generative AI. □</p>
<p>Kentucky: KBA Ethics Op. E-457 (2024)⁹⁰</p>	<p>No Billing for Saved Time: You cannot bill a client on an hourly basis for the time you save by using AI. If a task that previously took five hours now takes one hour with AI assistance, you may only bill for the one hour of actual work.</p> <p>Reduced Fees May Be Necessary: Because AI can create significant efficiencies, a lawyer should consider reducing the overall fee to ensure it remains reasonable under Kentucky Supreme Court Rule 3.130(1.5). An expeditious result from AI may warrant a lower fee.</p> <p>Client Consent for Direct Costs: You can only charge a client for the direct expense of using a specific AI tool for their matter if the client agrees to the charge in writing beforehand.</p> <p>Overhead vs. Billable Costs: General AI subscription costs (like a firm-wide license for a legal research tool with AI features) are considered overhead and cannot be passed on to the client.</p> <p>However, a case-specific AI service might be billable with the client's consent.</p> <p>No Billing for Learning: Time spent learning how to use a new AI tool is considered professional development and is part of your overhead; it cannot be billed to a client.</p>

89. Fla. Bar Ethics Op. 24-1 (2024).

90. Ky. Bar Ass'n Ethics Op. KBA E-457 (2024).

<p>Pennsylvania: Pa. Bar Ass'n & Phila. Bar Ass'n, Joint Formal Op. 2024-200 (2024)⁹¹</p>	<p>□AI has tremendous time-saving capabilities. Lawyers must, therefore, ensure that AI-related expenses are reasonable and appropriately disclosed to clients.□</p>
<p>Texas: Tex. Comm. on Pro. Ethics, Op. 705 (2025)⁹²</p>	<p>□[A] lawyer will likely be able to charge the client for the actual time the lawyer spends using a generative AI program for purposes of the representation, including to refine the program's outputs and check the work. A lawyer may not, however, charge hourly fees for the time that was □saved□ by using the generative AI program.□</p>

The official sources from these various bodies all indicate a general and unified consensus that a lawyer is strictly prohibited from billing their client for the time saved by using artificial intelligence. Repeatedly, the guidance and opinions use the example that if a complex research task would have previously required a junior associate to spend ten hours□ but it now can be completed by a lawyer in only two hours through the use of a sophisticated AI assistant□ the client can only be billed for the two hours of the lawyer's actual, verifiable work (a period which must include the time spent crafting prompts for the AI, and, most crucially, the time spent meticulously reviewing and independently verifying its output).⁹³ The act of billing the client for eight □phantom□ hours that were not actually worked would unquestionably be classified as an unreasonable fee and would therefore constitute a clear and sanctionable violation of the principles set forth in Rule 1.5.

91. Pa. Bar Ass'n Comm. on Legal Ethics and Pro. Resp. & Phila. Bar Ass'n Pro. Guidance Comm., Joint Formal Op. 2024-200 (2024).

92. Tex. Comm. on Pro. Ethics, Op. 705 (2025).

93. *See, e.g.*, A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 12 (2024) (illustrating that when generative AI reduces a task that previously required ten hours of junior-associate research to two hours of attorney work, the lawyer may bill only for the actual time reasonably expended, including time spent formulating prompts and independently reviewing and verifying the AI's output).

b. Billing for AI Costs and Learning Time

The guidance and opinions also provide much-needed clarity on several related and important billing issues. Regarding charging for the AI tools themselves, it may be permissible for lawyers to charge their clients for the direct costs of using a specific, sophisticated AI tool for a particular client matter.⁹⁴ However, it is important to note that this is not an automatic entitlement. Any such charge must be treated as a distinct expense, fully and transparently disclosed to the client and explicitly agreed to in advance, preferably within the written fee agreement.⁹⁵

More generalized AI subscriptions or tools that are utilized across the entire firm are typically considered part of the firm's standard operational overhead and therefore cannot be billed directly to any single client.⁹⁶ Regarding the practice of charging for education and training time, the time a lawyer invests in learning to use a general artificial intelligence tool effectively is considered a fundamental component of their professional duty of competence, as established under Rule 1.1.⁹⁷ As such, this time is properly classified as an investment in their own continuing professional development and, consequently, cannot be billed as a separate line item to a client.

The continuing rise of agentic AI technology is poised to accelerate these existing pressures to a breaking point and might, in fact, compel a fundamental, revolutionary paradigm shift in the legal profession's

94. *See id.* at 14 (costs associated with generative AI may be charged to the client, but a fee agreement should explain the basis for all fees and costs, including those associated with the use of generative AI).

95. *Id.*

96. *See* CAL. BAR PRACTICAL GUIDANCE, *supra* note 32, at 4 (Costs associated with generative AI may be charged to the clients in compliance with applicable law. A fee agreement should explain the basis for all fees and costs, including those associated with the use of generative AI.); *ABA Guidance on Ethical Use of AI for Lawyers*, SMOKEBALL (Nov. 20, 2024), <https://www.smokeball.com/blog/aba-guidance-ethical-ai-for-lawyers> [<https://perma.cc/4KPM-RTQD>] ([A] lawyer should refrain from charging a client for these costs and should consider the expenses associated with legal AI tools as overhead.).

97. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 14 (2024); *see* Candace Whitman, *ABA Issues First Guidance on the Use of Artificial Intelligence for Lawyers*, ADVOC. CAP. (Aug. 1, 2024), <https://www.advocatecapital.com/blog/aba-issues-first-guidance-on-the-use-of-artificial-intelligence-for-lawyers/> [<https://perma.cc/Y3XR-7WED>] (It is unethical to charge clients for the time spent learning how to use the AI tool.).

underlying business model.⁹⁸ The current ethical framework that governs billing practices is still largely anchored to the traditional concept of direct human effort, which is typically measured and quantified in terms of hours worked.⁹⁹ While generative AI already places significant strain on this established model by drastically reducing the hours required to complete certain tasks, agentic AI can perform complex, multi-step workflows entirely autonomously over extended periods.¹⁰⁰ A lawyer might, for example, spend just thirty minutes defining a clear objective and setting the operational parameters for an AI agent, which could then proceed to work for the subsequent twenty hours completely independently to complete a comprehensive due diligence review. In a scenario such as this, the very concept of the “billable hour” as a meaningful measure of value becomes almost entirely nonsensical. The actual value that is delivered to the client—in this case, the completed and detailed report—is now completely disconnected from the lawyer’s direct and personal time input. This technological reality will undoubtedly create immense and perhaps ultimately unavoidable pressure on law firms to abandon the traditional hourly billing model for many different types of legal work and to transition toward alternative fee arrangements, such as fixed fees, subscription-based models, or other innovative forms of value-based billing that are designed to compensate the firm for the ultimate *outcome* and the delivered *value* of the legal service, rather than for the specific amount of time a human being spent providing it.¹⁰¹

98. Sam Skolnik, *Big Law’s Increasing AI Use Could Speed Billable Hours’ Demise*, BLOOMBERG L. (May 30, 2025, at 14:20 CST), <https://news.bloomberglaw.com/legal-ops-and-tech/big-laws-increasing-ai-use-could-speed-billable-hours-demise> [https://perma.cc/D16Q-YF6K]; Robert J. Couture, *The Impact of Artificial Intelligence on Law Firms’ Business Models*, HARV. L. SCH. CTR. ON THE LEGAL PRO.: INSIGHT (Feb. 24, 2025), <https://clp.law.harvard.edu/knowledge-hub/insights/the-impact-of-artificial-intelligence-on-law-law-firms-business-models/> [https://perma.cc/3ZGP-7P46].

99. Mark C. Palmer, *The AI Revolution in Legal and the Billable Hour: Is the End Near?*, 2CIVILITY (Jul. 9, 2024), <https://www.2civility.org/the-ai-revolution-in-legal-and-the-billable-hour-is-the-end-near/> [https://perma.cc/PQ8A-ZYSM]; Couture, *supra* note 98.

100. Perlin, *supra* note 85, at 45; *Agentic Workflows for Legal Professionals: A Smarter Way to Work with AI*, THOMSON REUTERS (Aug. 14, 2025), <https://legal.thomsonreuters.com/blog/agentic-workflows-for-legal-professionals-a-smarter-way-to-work-with-ai> [https://perma.cc/GJJ5-U49A] (explaining that an agentic system “plans, executes, and adapts” across multi-step processes); Riedl & Desai, *supra* note 69.

101. *See generally* Palmer, *supra* note 99 (discussing how integrating AI into legal practice will lead to a reform of traditional billing practices); Couture, *supra* note 98 (delving into billing challenges AI

5. Rule 1.2 (Scope of Representation): Allocating Authority Between Lawyer, Client, and AI Agent

Model Rule 1.2(a) establishes the fundamental allocation of authority in the client-lawyer relationship: [a lawyer shall abide by a client's decisions concerning the objectives of representation and . . . shall consult with the client as to the means by which they are to be pursued.]¹⁰² Means and ends have often been divided in a straightforward way—ends are the objectives that belong to the client, while means require technical expertise that belongs to the lawyer. But as with the other topics above, employing agentic AI throws a wrench into the normal suppositions of [means] to carry out a client matter.

a. Impact of Agentic AI on the Scope of Representation

While generative AI acts as a tool to help a lawyer pursue the [means], agentic AI, with its autonomous decision-making capabilities, profoundly challenges this traditional allocation of authority.¹⁰³ When a lawyer deploys an autonomous agent with a high-level objective, such as [negotiate the best possible settlement for the client, not to exceed \$500,000], it blurs the lines of who is making the critical strategic decisions during the process.¹⁰⁴ The argument would be that it was the client who set the ultimate objective of settlement, but it is less conclusive to assert that the lawyer who programmed the agent's parameters and constraints determined the means because the AI agent could potentially make thousands of micro-decisions and tactical choices in real-time during a negotiation with another AI, none of which will be under the direct control of the attorney.¹⁰⁵ At this point,

tools bring with them); Perlín, *supra* note 85, at 62 (discussing studies show how generative AI decreases lawyer's time for tasks but not their effectiveness).

102. MODEL RULES OF PROFESSIONAL CONDUCT R. 1.2(a) (AM. BAR ASS'N 2026).

103. See N.Y. City Bar Ass'n Comm. on Pro. Ethics, Formal Op. 2024-5 (2024) (emphasizing lawyers' duty to maintain independent professional judgment and to consult with clients); see also Desai & Riedl, *supra* note 69 (addressing the various issues with agentic AI); Andrew Fletcher, *The Key to Autonomous Legal Workflows with Agentic AI*, THOMSON REUTERS (June 5, 2025), <https://legal.thomsonreuters.com/blog/the-key-to-autonomous-legal-workflows-with-agentic-ai/> [<https://perma.cc/N46E-P2B6>] (describing agentic systems that plan, act, and adapt toward predefined objectives beyond assistive, prompt-level tools).

104. MODEL RULES OF PROFESSIONAL CONDUCT R. 1.5.

105. See generally Jake Jones, *Stop Calling Workflows [Agents]—A Guide to Real Agentic AI*, ARTIFICIAL LAW. (Aug. 26, 2025), <https://www.artificiallawyer.com/2025/08/26/stop-calling-workflows-agents-a-guide-to-real-agentic-ai/> [<https://perma.cc/U9N7-NQEX>] (defining true agentic

under Rule 1.2(a), the client should be informed of the arrangement and how it will be performed as the “means” to carry out the client’s objective so that the client can consent to the arrangement.¹⁰⁶

b. Delegating the “Means” to a Non-Human Non-Lawyer

A lawyer generally has discretion over the tactical “means” of a representation.¹⁰⁷ However, as noted above, delegating a substantive function like negotiation or key decision-making to an autonomous agent (a non-human non-lawyer) is a “means” so significant and novel that it almost certainly requires consultation and informed consent from the client under Rule 1.2.¹⁰⁸ The client must understand not just the objective, but also the nature of the powerful and autonomous tool being used to pursue it.¹⁰⁹ Deploying such an agent without this consultation would risk usurping the client’s ultimate authority over their matter.¹¹⁰

autonomy as goal-holding, tool-choosing, acting across external systems, and replanning without step-by-step human control); Qingyun Wu et al., *AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation*, ARXIV (Oct. 3, 2023), <https://arxiv.org/abs/2308.08155> [<https://perma.cc/QYE7-9BM6>] (documenting multi-agent LLM frameworks where agents iteratively plan and act via tool-use and inter-agent dialogue).

106. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 8–9 (2024) (directing lawyers to communicate with clients regarding AI use consistent with Rules 1.2 and 1.4, and to maintain independent professional judgment); MODEL RULES OF PROF'L CONDUCT R. 1.2(a).

107. See MODEL RULES OF PROF'L CONDUCT R. 1.2(a) (“[A] lawyer shall abide by a client’s decisions concerning the objectives of representation and . . . shall consult with the client as to the means by which they are to be pursued.”)

108. See Ill. State Bar Ass'n, Pro. Conduct Advisory Op. No. 19-04 (2019) (discussing requirements for client disclosure and informed consent when delegating “complete or substantial responsibility” for a matter to a nonlawyer or outside lawyer); N.C. State Bar, Formal Ethics Op. 12 (2008) (requiring disclosure and informed written consent before outsourcing to nonlawyers/foreign assistants); see also A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 9 (2024) (explaining decisions about whether and how to employ generative AI in performing substantive legal functions may constitute a significant “means” of representation, such that delegating core tasks to autonomous AI tools can trigger the lawyer’s duty under Rule 1.2 to consult with the client and obtain informed consent).

109. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 8–9 (2024); CAL. BAR PRACTICAL GUIDANCE, *supra* note 32, at 4; see Or. State Bar, Formal Op. No. 2025-205: Artificial Intelligence Tools (2025) (advising that novel tech use may require client disclosure/consultation under Rules 1.2 & 1.4).

110. N.Y. City Bar Ass'n Comm. on Pro. Ethics, Formal Op. 2024-5 (2024).

This challenge necessitates a re-evaluation of how the scope of representation is defined and documented.¹¹¹ A traditional engagement letter that sets out broad objectives may no longer be sufficient in an age of agentic AI.¹¹² To comply with Rule 1.2 and properly manage risk, law firms will likely need to develop more granular and dynamic agreements.¹¹³ This could take the form of "AI Use Protocols" included as an addendum to the engagement letter.¹¹⁴ Such protocols would explicitly detail which types of tasks may be delegated to an AI agent, the specific parameters of its autonomy, the client's role in approving those parameters, and, crucially, the predefined points at which the agent must stop and allow human beings to take the wheel.¹¹⁵ This approach transforms the scope of representation from a static, one-time agreement into a living document that is revisited and adjusted as the role of autonomous technology in the matter evolves.¹¹⁶

B. *The Lawyer As Advocate: Candor, Fairness, and the AI Co-Counsel*

When a lawyer acts as an advocate before a tribunal, the duties of candor and fairness are paramount. The use of AI as a "co-counsel" in research and drafting puts these duties under intense scrutiny.

111. NYSBA REPORT AND RECOMMENDATION, *supra* note 34, at 57.

112. *Id.*; Or. State Bar, *supra* note 109.

113. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 5 (2024); *see* NAT'L INST. OF STANDARDS AND TECH. U.S. DEPT OF COM., ARTIFICIAL INTELLIGENCE RISK MANAGEMENT FRAMEWORK (AIRMF 1.0) 2 (2023) (discussing a "framework" for governing AI risk with documented roles, processes, and controls).

114. NYSBA REPORT AND RECOMMENDATION, *supra* note 34, at 34-36; *see also* TEX. BAR PRAC., SAMPLE AI DISCLOSURES FOR ATTORNEY-CLIENT CONTRACTS (2025), <https://www.texasbarpractice.com/wp-content/uploads/2025/05/Sample-AI-Disclosures-for-Attorney-Client-Contracts-final.pdf> [<https://perma.cc/2FBH-E5AG>] (providing sample clauses to disclose AI use, supervision, and confidentiality safeguards).

115. *See generally* NAT'L INST. OF STANDARDS AND TECH., *supra* note 113, at 12, 21-24 (explaining responsible AI practices to manage risks associated with its use); CAL. BAR PRACTICAL GUIDANCE, *supra* note 32, at 44 (offering examples of parameters for attorney-client agreements).

116. *See* NAT'L INST. OF STANDARDS AND TECH., *supra* note 113, at 48; Or. State Bar, *supra* note 109 (detailing obligations of attorneys to evolve agreements with clients parallel to evolution of artificial intelligence).

1. Rule 3.1 (Meritorious Claims) & Rule 3.3 (Candor Toward the Tribunal): The Specter of AI □Hallucinations□ and the Inviolable Duty of Verification

Model Rule 3.1 prohibits a lawyer from bringing a claim or asserting a contention unless there is a basis in law and fact for doing so that is not frivolous.¹¹⁷ Model Rule 3.3, the duty of candor toward the tribunal, forbids a lawyer from knowingly making a false statement of fact or law to a court or failing to correct a false statement of material fact or law previously made.¹¹⁸

a. The Epicenter of AI Misconduct

The duties of candor and responsibility for bringing meritorious claims and valid arguments under Model Rules 3.1 and 3.3 have become the epicenter of AI-related professional misconduct, generating the most high-profile disciplinary actions and cautionary tales. The canonical example is the 2023 case of *Mata v. Avianca, Inc.*,¹¹⁹ in which two New York lawyers filed a brief opposing a motion to dismiss that cited more than half a dozen entirely fictitious judicial opinions, complete with fake quotes and bogus citations, all generated by ChatGPT.¹²⁰ When opposing counsel could not locate the cases, the lawyers, instead of retracting, doubled down, even producing fabricated text of the □opinions□ provided by the AI.¹²¹ The lawyers eventually pled ignorance and □confusion□ about the AI model and how it worked, hoping that this would exonerate them under the mental state requirement (knowingness),¹²² but the court rejected this defense and ultimately imposed a \$5,000 sanction, finding that the attorneys acted with □subjective bad faith□ by consciously avoiding the truth and misleading the court.¹²³

117. MODEL RULES OF PROFESSIONAL CONDUCT R. 3.1 (AM. BAR ASS'N 2026).

118. MODEL RULES OF PROFESSIONAL CONDUCT R. 3.3(a).

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

120. *See id.* at 464-65 (sanctioning attorney for including fake, AI-generated legal citations in a filing). The court applied FED. R. CIV. P. 11 to the conduct rather than an ethical rule.

121. *Id.* at 449, 452-56.

122. *Id.* at 463. Rule 11 requires an affirmation upon □knowledge, information, and belief□ FED. R. CIV. P. 11.

123. *Mata*, 678 F. Supp. 3d at 463, 464-66.

This was not an isolated event. In May 2025, a special master sanctioned two law firms, including the global firm K&L Gates, for a "collective debacle" in which a brief containing "bogus AI research" was filed.¹²⁴ The error originated with a lawyer at one firm using AI to create an outline, which was then incorporated into the final brief by lawyers at the other firm without independent verification.¹²⁵ Similarly, in February 2025, three lawyers from Morgan & Morgan were sanctioned for filing motions that cited eight non-existent cases generated by the firm's own in-house AI tool.¹²⁶

b. The Unanimous Ethical Mandate Regarding Candor and Verification

The above cases solidified a clear and non-negotiable ethical mandate: the lawyer, not the AI, is absolutely and unequivocally responsible for the accuracy and veracity of every statement made in a court filing.¹²⁷ There is no viable defense of "the AI did it" or "I didn't know the AI could be wrong." Every ethics opinion on the subject—from ABA Formal Opinion 512 to the guidelines from California, Florida, and others—mandates that lawyers must critically review, independently verify, and take full responsibility for all AI-generated content before it is submitted to any tribunal.¹²⁸

124. *Lacey v. State Farm Gen. Ins.*, No. CV 24-5205 FMO, 2025 WL 1363069, at *1 (C.D. Cal. May 5, 2025); see Weiss, *supra* note 43 (highlighting the sanction against K&L Gates).

125. See *Lacey*, 2025 WL 1363069, at *1-2 (noting the AI-generated outline contained the research at issue); see also Weiss, *supra* note 43 (discussing sanctions imposed).

126. *Wadsworth v. Walmart, Inc.*, 348 F.R.D. 489, 493 (D. Wyo. 2025); Pazanowski, *supra* note 44.

127. See *Mata*, 678 F. Supp. 3d at 466 (sanctioning attorney for including fake, AI-generated legal citations in a filing); *Lacey*, 2025 WL 1363069, at *5 (striking Plaintiff's supplemental briefs drafted with AI, as well as monetary sanctions the plaintiff's law firms had to pay to the defendant); Weiss, *supra* note 43 (highlighting the sanctions in *Lacey*); *Wadsworth*, 348 F.R.D. at 493 (imposing sanctions for citing nonexistent cases); Pazanowski, *supra* note 44 (discussing the sanctions imposed in *Wadsworth*).

128. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 10 (2024); CAL. BAR PRACTICAL GUIDANCE, *supra* note 32, at 4; Fla. Bar Ethics Op. 24-1 (2024); N.J. COMM. ON AI AND THE CTS., PRELIMINARY GUIDELINES ON THE USE OF AI BY NEW JERSEY LAWYERS 3-4 (2024); Pa. Bar Ass'n Comm. on Legal Ethics and Pro. Resp. & Phila. Bar Ass'n Pro. Guidance Comm., Joint Formal Op. 2024-200 (2024); Tex. Comm. on Pro. Ethics, Op. 705 (2025).

This duty of verification is an indispensable component of the duties of both competence (Rule 1.1) and candor (Rule 3.3).¹²⁹ The emergence of agentic AI creates the potential for a new and insidious form of this violation: a candor violation by omission. Current disciplinary cases have stemmed from a lawyer's affirmative act of filing a document containing false information.¹³⁰ However, agentic AI is being developed to autonomously execute multi-step workflows, including the "drafting" and "filing [of] motions."¹³¹ Imagine a scenario where a busy lawyer deploys an AI agent to handle routine filings, such as motions for an extension of time. If the agent, in an effort to bolster the motion, independently generates a supporting "fact" or a citation that is a hallucination and then files the document with the court without the lawyer's final, line-by-line review, the lawyer has violated Rule 3.3.¹³² The violation in this future scenario would occur not through an act of flawed review, but through a complete *omission* of review resulting from an improper delegation of a core professional function. This risk underscores the critical importance of implementing firm-wide policies that require a "human in the loop" for any AI-assisted action that involves a representation to a court.¹³³

2. Rule 3.4 (Fairness to Opposing Party and Counsel): AI in Discovery, Negotiations, and Strategic Analysis

Model Rule 3.4 requires a lawyer to be fair to the opposing party and their counsel.¹³⁴ This includes prohibitions on unlawfully obstructing another

129. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 10 (2024).

130. Sara Merken, *AI "Hallucinations" in Court Papers Spell Trouble for Lawyers*, REUTERS (Feb. 18, 2025, at 14:55 CST), <https://www.reuters.com/technology/artificial-intelligence/ai-hallucinations-court-papers-spell-trouble-lawyers-2025-02-18/> [<https://perma.cc/Y5RN-MUNP>] (reporting that generative AI has led to filings with fictitious case citations, generating disciplinary risk" but so far only for affirmative filings).

131. Brandon Fierro, *Short Circuit Court: AI Hallucinations in Legal Filings and How to Avoid Making Headlines*, REUTERS (Aug. 4, 2025, at 08:27 CDT), <https://www.reuters.com/legal/legalindustry/short-circuit-court-ai-hallucinations-legal-filings-how-avoid-making-headlines-2025-08-04/> [<https://perma.cc/P7GU-84KR>] (underscoring the need for "human oversight" a "human in the loop" to prevent AI-generated hallucinations from being submitted).

132. See Merken, *supra* note 130 (reporting that generative AI has led to filings with fictitious case citations, generating disciplinary risk" but so far only for affirmative filings); Fierro, *supra* note 131 (underscoring the need to prevent AI-generated hallucinations from being submitted).

133. Merken, *supra* note 130; Fierro, *supra* note 131.

134. MODEL RULES OF PROFESSIONAL CONDUCT R. 3.4 (AM. BAR ASS'N 2026).

party's access to evidence, falsifying evidence, or engaging in other conduct that prejudices the integrity of a proceeding.¹³⁵

a. AI in E-Discovery

AI has long been used in e-discovery to help lawyers analyze vast volumes of documents to identify those that are relevant and privileged.¹³⁶ The ethical obligations of fairness and candor apply directly to this process. A lawyer would violate Rule 3.4 if they intentionally used an AI tool to hide or obscure responsive documents.¹³⁷ This could be done, for example, by manipulating search parameters in a technology-assisted review process to systematically exclude certain types of documents or by using a biased algorithm known to under-identify responsive materials from certain custodians.¹³⁸ The duty of fairness may require transparency and cooperation with opposing counsel regarding the AI protocols used in discovery, especially if the process is challenged.¹³⁹

b. AI in Negotiation

The prospect of using agentic AI in negotiations raises novel fairness issues. An AI agent could be deployed to negotiate a settlement with another AI agent or a human counterpart.¹⁴⁰ A fairness issue under Rule 3.4 (and a truthfulness issue under Rule 4.1) would arise if the AI were programmed to use deceptive tactics, knowingly misrepresent material facts,

135. *Id.*

136. *Compass, Inc. v. Zillow, Inc.*, No. 1:25-CV-05201-JAV, 2025 WL 2330701, at *2 (S.D.N.Y. Aug. 13, 2025); *Everlast Roofing, Inc. v. Wilson*, No. 1:23-CV-828, 2025 WL 1959345, at *4 (M.D. Pa. July 16, 2025); *Jennings v. USAA Cas. Ins.*, No. 3:23-CV-06171-DGE, 2025 WL 1167850, at *2 (W.D. Wash. Apr. 22, 2025); *Da Silva Moore v. Publics Groupe*, 868 F. Supp. 2d 137, 142 (S.D.N.Y. 2012).

137. *E.g.*, *Qualcomm Inc. v. Broadcom Corp.*, No. 05CV1958-B, 2008 WL 66932, at *9 (S.D. Cal. Jan. 7, 2008), *vacated in part*, 2008 WL 638108 (S.D. Cal. Mar. 5, 2008).

138. *See* Nikhil Guha et. al., *Vulnerabilities in Discovery Tech*, 35 HARV. J.L. & TECH. 582, 615-17 (2022) (discussing claims of mis-programmed technology-assisted review and court review of the claims).

139. *E.g.*, Sedona Conference, *The Sedona Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production*, 19 SEDONA CONF. J. 28, 30-31, 76-77 (2017) (recognizing modern search/analytics for ESI).

140. *See* Bob Ambrogi, *A Compendium of Legal Ethics Opinions on Gen AI (As Compiled by You Guessed It—Gen AI)*, LAW SITES (Feb. 7, 2025), <https://www.lawnext.com/2025/02/a-compendium-of-legal-ethics-opinions-on-gen-ai-as-compiled-by-you-guessed-it-gen-ai.html> [<https://perma.cc/N26Y-WS67>] (highlighting potential uses of AI in the legal field).

or exploit vulnerabilities in the opposing system.¹⁴¹ Even if not explicitly programmed to be deceptive, an agent driven solely by the objective of maximizing a client's outcome could develop emergent behaviors that cross ethical lines. Furthermore, an AI agent could be used to scrape publicly available data (e.g., social media data) about the opposing counsel or their client to build a psychological profile and exploit perceived biases or weaknesses in the negotiation, raising serious questions of fairness and respect for the rights of third persons (Rule 4.4).¹⁴²

The use of AI in litigation also has broader, systemic implications for fairness. The Preamble to the Model Rules expresses a core value of the profession: the responsibility to seek "improvement of the law, access to the legal system, [and] the administration of justice."¹⁴³ The rise of sophisticated predictive AI could create a technological "arms race" that undermines this goal.¹⁴⁴ AI tools can be used for predictive analytics, forecasting the likely outcomes of cases, predicting the rulings of specific judges, or modeling jury behavior.¹⁴⁵ Large, well-resourced law firms will invariably have access to more powerful and accurate predictive and agentic AI systems than solo practitioners, small firms, or legal aid organizations.¹⁴⁶ This creates a significant asymmetry of power.¹⁴⁷ One side can model and "war-game" the litigation to a degree the other cannot, giving them an unfair advantage

141. Sara Merken, *Lawyers Using AI Must Heed Ethics Rules, ABA Says in First Formal Guidance*, REUTERS (Jul. 29, 2024, at 16:13 CDT), <https://www.reuters.com/legal/legalindustry/lawyers-using-ai-must-heed-ethics-rules-aba-says-first-formal-guidance-2024-07-29> [https://perma.cc/LK8W-BGJY].

142. See N.Y.C. Bar Ass'n, Formal Op. 2012-2 (2012) (discussing the ethical constraints on juror research via social media, including respect for privacy, avoiding misrepresentation, and noting that research must not involve deceptive means); see also D.C. Bar, Ethics Op. 371 (2016) (addressing accessing public social media of potential jurors/adversaries, privacy, public data, and duties under Rules 1.1, 3.1, etc.).

143. MODEL RULES OF PROFESSIONAL CONDUCT Preamble ¶ 6 (AM. BAR ASS'N 2026).

144. Matthew Dylag, *Impact of Artificial Intelligence on Access to Justice: Predictive Analytics and the Legal Services Market*, 48 DAL. L.J. 1, 22 (2025).

145. *Id.* at 8-10, 12.

146. Dylag, *supra* note 144, at 22, 23, 28; Ashwin Telang, *The Promise and Peril of AI Legal Services to Equalize Justice*, HARV. J.L. & TECH. ONLINE (Mar. 14, 2023), <https://jolt.law.harvard.edu/digest/the-promise-and-peril-of-ai-legal-services-to-equalize-justice> [https://perma.cc/4X78-628D]; Drew Simshaw, *Access to A.I. Justice: Avoiding an Inequitable Two-Tiered System of Legal Services*, 24 YALE J.L. & TECH. 150, 205 (2022) (highlighting how rising reliance on legal AI could widen, rather than shrink, justice gaps unless carefully regulated).

147. Telang, *supra* note 146; Simshaw, *supra* note 146, at 158-60.

in setting case strategy, valuing settlements, and making decisions about whether to proceed to trial.¹⁴⁸ This technological disparity threatens to exacerbate the existing justice gap, making it more difficult for the poor and under-resourced to secure adequate legal counsel and a fair outcome, a direct contradiction of the profession's aspirational goals.¹⁴⁹

C. *Law Firms and Associations: The Imperative of Supervision*

The integration of AI tools, particularly those that function with some degree of autonomy, places significant new responsibilities on law firm partners and supervising lawyers. The existing rules on supervision provide a clear, if challenging, framework for managing these new risks.

1. Rules 5.1 & 5.3 (Supervision): Supervising the "Digital Associate" and Third-Party AI Vendors

The Model Rules create a clear hierarchy of supervisory responsibility. Rule 5.1 requires that partners and lawyers with "comparable managerial authority" make reasonable efforts to ensure their firm has policies and procedures in place to give "reasonable assurance that all lawyers in the firm conform to the Rules of Professional Conduct."¹⁵⁰ It also places a duty on lawyers with direct supervisory authority over another lawyer to ensure that lawyer's compliance.¹⁵¹ Rule 5.3 extends a parallel duty to the supervision of "nonlawyer assistants," requiring lawyers to provide appropriate instruction and supervision and to be responsible for their work product.¹⁵²

a. AI Systems and Agents That Provide "Nonlawyer Assistance"

There is a broad and firm consensus among ethics authorities that an AI tool or an AI agent, for the purposes of professional responsibility, should be treated as an entity that provides "nonlawyer assistance" under

148. Telang, *supra* note 146; Simshaw, *supra* note 146, at 158-60.

149. Telang, *supra* note 146; Simshaw, *supra* note 146, at 158-60; Dylag, *supra* note 144, at 5, 6, 28.

150. MODEL RULES OF PROF'L CONDUCT R. 5.1(a) (AM. BAR ASS'N 2026).

151. *Id.* R 5.1(b).

152. *Id.* R 5.3.

Rule 5.3.¹⁵³ This characterization has critical implications.¹⁵⁴ It means that a lawyer's professional judgment can never be delegated to the AI.¹⁵⁵ The lawyer must give the "assistant" appropriate instruction (e.g., through carefully crafted prompts and custom instructions), supervise its work (e.g., by diligently reviewing its performance and its outputs), and ultimately be responsible for the final work product.¹⁵⁶

In the supervision of AI agents, two concepts are paramount in importance: observability and traceability.¹⁵⁷ Observability refers to the ability of a supervisor to see the agent's work and its thinking and reasoning while it is performing a multistep task.¹⁵⁸ One can see what steps the agent took to reach outputs, understand its tool usage and data retrieval logic, and identify where the agent's reasoning paths diverged or failed.¹⁵⁹ With traceability, sometimes called agent tracing or simply debugging, you can supervise and review how the outputs were created by the agent step-by-step through action logs, inter-agent communication maps, state transition histories, and error localization.¹⁶⁰ Without observability and traceability the agent providing nonlawyer assistance becomes a black box and an attorney supervisor has no means to second guess the thinking, reasoning, and steps

153. See A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 10 (2024) (explaining that lawyers must treat AI tools as they would nonlawyer assistants under Rule 5.3 and remain responsible for their supervision); N.C. State Bar Comm. on Ethics & Pro. Resp., Formal Op. 2024-1 (2024) (requiring that generative AI be treated as a nonlawyer assistant under Rule 5.3, with duties of instruction, supervision, and responsibility resting on the lawyer).

154. CAL. BAR PRACTICAL GUIDANCE, *supra* note 32, at 1 (emphasizing that AI use implicates professional obligations including competence, confidentiality, and supervision).

155. *Id.* at 3 ("[A] lawyer's professional judgment may never be delegated to generative AI").

156. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 11 (2024) (stating that lawyers must instruct, supervise, and review the work of AI tools as they would nonlawyer assistants, and remain ultimately responsible for the output).

157. Gregg Lindemulder & Annie Badman, *Why Observability Is Essential for AI Agents*, IBM (Aug. 14, 2025), <https://www.ibm.com/think/insights/ai-agent-observability> [https://perma.cc/R6F7-QTJX]; Kuldeep Paul, *Agent Tracing for Debugging Multi-Agent AI Systems*, MAXIM (Aug. 14, 2025), <https://www.getmaxim.ai/articles/agent-tracing-for-debugging-multi-agent-ai-systems/> [https://perma.cc/PQ75-RRHM].

158. Lindemulder & Badman, *supra* note 157; Paul, *supra* note 157.

159. Paul, *supra* note 157.

160. *Id.*; *Top Tools and Techniques for Debugging Agentic AI Systems*, AMLEWORK (Jul. 22, 2025), <https://www.amplework.com/blog/debugging-agentic-ai-tools-techniques/> [https://perma.cc/B4VD-D9X6].

taken of the agent as it is working and no means to trace and debug the errors found when the outputs are reviewed.¹⁶¹

b. Firm-Level Responsibility

The duty of supervision is not merely an individual one. Rule 5.1(a) makes it a firm-wide, managerial responsibility to establish effective policies and procedures.¹⁶² In the age of AI, this requires firms to move beyond ad-hoc usage and implement a system of governance.¹⁶³ This includes creating clear, written AI usage policies that dictate which tools are approved for use, establish protocols for handling confidential client data, mandate verification procedures for all AI-generated work, and provide for ongoing training.¹⁶⁴ The sanctioning of the *firm* K&L Gates, not just the individual lawyers, serves as a powerful reminder that the failure to implement such measures is a collective, organizational failure for which the firm itself can be held responsible.¹⁶⁵

c. Supervising AI Vendors

The principles of Rule 5.3 also apply when a lawyer uses nonlawyers *outside* the firm, such as retaining an investigative service or a document management company.¹⁶⁶ This duty extends directly to the use of third-party AI vendors.¹⁶⁷ When a law firm retains the services of an AI provider, it has a supervisory duty to make reasonable efforts to ensure that the vendor's platform and practices are compatible with the lawyer's own ethical

161. Ofir Pearl, *First Look, Then Leap: Why Observability Is the First Step in Securing Your AI Agents*, ZENITY (Jul. 24, 2025), <https://zenity.io/blog/security/observability-the-first-step-in-securing-your-ai-agents> [<https://perma.cc/NPL6-A9HB>]; *Top Tools and Techniques for Debugging Agentic AI Systems*, *supra* note 160.

162. MODEL RULES OF PROFESSIONAL CONDUCT R. 5.1 cmt. 1-3, 5 (AM. BAR ASS'N 2026).

163. See A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 10 (2024) (recommending law firms establish firm-wide AI governance, policies, and training).

164. See *id.* at 11 (recommending law firms establish firm-wide AI governance, policies, and training); NAT'L INST. OF STANDARDS AND TECH., *supra* note 113, at 19 (calling for organizational AI governance policies, verification procedures, and continuous training).

165. *Lacey v. State Farm Gen. Ins. Co.*, No. 2:24-cv-05205, 2025 WL 1363069, at *4 (C.D. Cal. May 6, 2025); Weiss, *supra* note 43.

166. MODEL RULES OF PROFESSIONAL CONDUCT R. 5.3 cmt. 3 (explaining that Rule 5.3 duties apply equally to services performed by nonlawyers outside the firm, including investigators and document management companies).

167. *E.g.*, N.C. State Bar Comm. on Ethics & Pro. Resp., Formal Op. 2024-1 (2024).

obligations, particularly regarding data security, confidentiality, and reliability.¹⁶⁸ This requires thorough due diligence before engaging a vendor and ongoing monitoring of their services.¹⁶⁹

Agentic AI creates a complex, layered supervisory challenge by merging the duties of Rule 5.1 and 5.3. Imagine a junior associate, a subordinate lawyer under Rule 5.1, who uses an agentic AI tool, a nonlawyer assistant under Rule 5.3, to perform a complex task.¹⁷⁰ Under Rule 5.1, they are responsible for the associate's conduct, including whether the associate competently deployed the agent and diligently reviewed its output.¹⁷¹ Under Rule 5.3, they are responsible for the AI tool's work product.¹⁷² If the AI agent makes a critical error that the junior associate fails to catch, the supervising partner's potential liability is compounded: Did they adequately supervise the associate's review process? And, more fundamentally, did the firm have adequate policies, training, and vetting procedures in place for the use of that specific agentic AI tool?¹⁷³ This dual-layered risk makes robust, firm-wide AI governance, not just case-by-case individual supervision, an

168. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 498 (2021).

169. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 477R (2017); Pa. Bar Ass'n Comm. on Legal Ethics and Pro. Resp. & Phila. Bar Ass'n Pro. Guidance Comm., Joint Formal Op. 2024-200 (2024).

170. *See generally* MODEL RULES OF PROFESSIONAL CONDUCT R. 5.1, 5.3 (imposing affirmative duties on lawyers with managerial or supervisory authority to establish and enforce measures ensuring that both subordinate lawyers and nonlawyer assistants comply with ethical obligations, making supervising attorneys responsible for reasonable oversight and for preventing or remedying professional misconduct within their practice).

171. *See* MODEL RULES OF PROFESSIONAL CONDUCT R. 5.1 (clarifying that a lawyer may be subject to a jurisdiction's professional-responsibility rules on practicing law even when providing services outside that jurisdiction if those services involve legal advice or assistance, meaning lawyers must not engage in the unauthorized practice of law and must comply with relevant local regulations when their work (including AI-facilitated services) has a substantial connection to the forum state).

172. *See generally* MODEL RULES OF PROFESSIONAL CONDUCT R. 5.3 (requiring that lawyers ensure nonlawyer assistants, including technology vendors and staff, act in a manner compatible with the lawyer's professional obligations by implementing appropriate supervision, training, and oversight to prevent ethical breaches meaning attorneys remain responsible for the conduct of nonlawyers (including AI tools) that they engage or direct in their practice).

173. *See* A.B.A. Standing Comm. on Ethics & Pro. Resp., Formal Op. 467, at 10 (2014) (explaining that under Rules 5.1 and 5.3 prosecutors have managerial and supervisory obligations to ensure that both lawyers and nonlawyer assistants under their direction such as staff or technology act in compliance with professional conduct rules, emphasizing that supervisory lawyers may be responsible for preventing, mitigating, or addressing ethical violations by those they supervise).

absolute necessity for mitigating liability and fulfilling the duties of Rules 5.1 and 5.3.¹⁷⁴

2. Rule 5.5 (Unauthorized Practice of Law): Identifying the Line an Agentic AI Must Not Cross

Model Rule 5.5 prohibits a lawyer from practicing law in a jurisdiction in violation of that jurisdiction's regulations, and, critically, from "assisting another person" in the unauthorized practice of law (UPL).¹⁷⁵ UPL rules are designed to protect the public by ensuring that legal services requiring professional judgment are provided only by individuals who are licensed, competent, and subject to professional discipline.¹⁷⁶

a. The Unauthorized Practice of Law Frontier

For standard generative AI, the UPL risk is relatively low. When used as a drafting or research tool under the direct control and supervision of a licensed attorney, the AI is clearly an assistant, and the legal service is being provided by the lawyer.¹⁷⁷ However, agentic AI, with its capacity for autonomous reasoning and action, pushes directly against the UPL frontier.¹⁷⁸ The risk of a lawyer "assisting" the AI in UPL arises when the agent's autonomous functions cross the line from performing a task to providing a legal service that requires a lawyer's independent professional judgment.¹⁷⁹

174. See NAT'L INST. OF STANDARDS & TECH., *supra* note 113, at 21 (providing a voluntary, principles-based structure for identifying, assessing, and managing risks throughout the AI lifecycle—emphasizing functions such as governance, mapping, measuring, and managing risks to promote trustworthy and responsible AI systems while enabling organizations to tailor risk practices to their specific contexts); A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 10 (2024).

175. MODEL RULES OF PROF'L CONDUCT R. 5.5 cmt 1.

176. *Id.* cmt 2.

177. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 11 (2024).

178. *Agentic AI & Legal Ethics: The UPL Question*, NEXLAW (July 4, 2025) [hereinafter *Agentic AI & Legal Ethics*], <https://www.nexlaw.ai/blog/ai-ethics-autonomous-brief> [https://perma.cc/NWD6-PZRT]; *Ethical Uses of Generative AI in the Practice of Law*, THOMSON REUTERS (Aug. 8, 2025), <https://legal.thomsonreuters.com/blog/ethical-uses-of-generative-ai-in-the-practice-of-law> [https://perma.cc/N59B-A3FU].

179. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 4 (2024); *Agentic AI & Legal Ethics*, *supra* note 178; *Ethical Uses of Generative AI in the Practice of Law*, *supra* note 178.

b. When Does an AI Agent Cross the Line?

The UPL line is crossed when the AI system begins operating without meaningful human oversight in a way that replicates the core functions of a lawyer.¹⁸⁰ Several potential scenarios illustrate this risk:

AI-Powered Intake: A law firm chatbot for client intake that goes beyond simple fact-gathering and begins to analyze a prospective client's legal problem, assess the merits of their claim, and provide specific legal advice or recommend a course of action would likely be engaging in UPL.¹⁸¹

Autonomous Legal Opinions: An agentic system that is fed a contract and then, without a lawyer's substantive review and analysis, provides the client with a definitive legal opinion on the contract's validity, risks, and enforceability would be practicing law.¹⁸²

Automated Negotiation and Agreement: An AI agent authorized to negotiate and enter into a legally binding settlement or contract on behalf of a client, making substantive decisions along the way, is performing a function at the very heart of legal practice.¹⁸³

180. See *Agentic AI & Legal Ethics*, *supra* note 178 (discussing how emerging agentic AI systems that autonomously plan and execute legal tasks raises critical ethical concerns for U.S. lawyers by potentially crossing into the unauthorized practice of law unless lawyers maintain transparent "glass box" oversight, understand and approve each step of the process, and ensure compliance with core duties such as competence, supervision, and avoiding UPL despite the technology's increased autonomy); *Ethical Uses of Generative AI in the Practice of Law*, *supra* note 178 (asserting that AI should act as a legal assistant to enhance, rather than replace, professional judgment); N.C. State Bar Comm. on Ethics & Pro. Resp., Formal Op. 2024-1 (2024) (holding lawyers ultimately responsible for reviewing work produced by someone (or something) other than the lawyer).

181. See *AI for Legal Services: Is it Unauthorized Practice of Law (UPL)?*, BY DESIGN L. (July 1, 2024), <https://www.bydesignlaw.com/ai-for-legal-services-is-it-unauthorized-practice-of-law-upl> [<https://perma.cc/25F3-TMUC>] (describing key features of UPL in the State of Washington, including permissible and prohibited activities within the practice of law); Jayne R Reardon, *Advancing Technology Poses Challenges to Enforcing UPL*, A.B.A. L. PRAC. TODAY (Mar. 28, 2023), https://www.americanbar.org/groups/law_practice/resources/law-practice-today/2023-march/advancing-technology-poses-challenges-to-enforcing-upl/ [<https://perma.cc/QV78-F7DL>] (summarizing proscribed activities in the practice of law under UPL statutes).

182. Reardon, *supra* note 181; *Can AI Give Legal Advice? Exploring the Capabilities and Limitations of AI in Law*, CASE STATUS (June 24, 2025), <https://www.casestatus.com/blog/can-ai-give-legal-advice> [<https://perma.cc/VGW3-CMLU>].

183. Reardon, *supra* note 181; *Can AI Give Legal Advice? Exploring the Capabilities and Limitations of AI in Law*, *supra* note 181; see Vivan Marwaha, *Legal AI Agents: Accelerating Contract Workflows with Intelligent Automation*, LEGALON TECH. (Feb. 13, 2025), <https://www.legalontech.com/post/legal-ai-agents> [<https://perma.cc/GG6Z-UPMQ>] (depicting how AI agents perform tasks); see also Memorandum from the U.S. Dep't of Just. & the F.T.C. to David Beach, [Clerk of Court for the

The risk of UPL in the age of AI is not solely about the output of the tool; it is also about the public's perception and reliance. UPL statutes are fundamentally consumer protection laws.¹⁸⁴ If a technology company markets a product to the public as an "AI Lawyer" or a tool that can "solve your legal problems," it is inviting consumers to rely on that tool for legal judgment, a core element of the practice of law.¹⁸⁵ Even with fine-print disclaimers, the functionality, marketing, and user interface of such a product could create a situation where the public is receiving what it reasonably perceives to be legal services from an unlicensed entity—the AI.¹⁸⁶ This could lead state bar associations and regulators to shift their enforcement focus.¹⁸⁷ In addition to disciplining individual lawyers who misuse AI, they may begin to target the *providers* of these sophisticated agentic systems for engaging in the unauthorized practice of law on a mass scale. The proposed definition and regulation of Automated Decisionmaking Technology (ADMT) in California is an early signal of this potential regulatory direction, focusing on replacing human decision-making.¹⁸⁸

Supreme Court of Virginia 2 (Jan. 3, 1997) [hereinafter Memorandum to David Beach] (on file with author) (proposing the prohibition of settlement (AI) services for real estate-related transactions).

184. Memorandum to David Beach, *supra* note 183.

185. The boast of DoNotPay that invited persons to allow its AI lawyer to represent them in a court proceeding achieved a regulatory shut down order from the FTC and a class action lawsuit alleging UPL in California. See *DoNotPay*, FTC (Sep. 25, 2024), <https://www.ftc.gov/legal-library/browse/cases-proceedings/donotpay> [<https://perma.cc/LY3Y-8GNU>] (providing proposed order); Press Release, FTC, FTC Finalizes Order with DoNotPay That Prohibits Deceptive "AI Lawyer" Claims, Imposes Monetary Relief, and Requires Notice to Past Subscribers (Feb. 11, 2025), <https://www.ftc.gov/news-events/news/press-releases/2025/02/ftc-finalizes-order-donotpay-prohibits-deceptive-ai-lawyer-claims-imposes-monetary-relief-requires> [<https://perma.cc/3FSN-2YMV>] (providing press release on disposition); Sara Merken, *Lawsuit Pits Class Action Firm Against "Robot Lawyer" DoNotPay*, REUTERS (Mar. 9, 2023), <https://www.reuters.com/legal/lawsuit-pits-class-action-firm-against-robot-lawyer-donotpay-2023-03-09/> [<https://perma.cc/4LR2-XDR4>] (discussing class action alleging that DoNotPay engaged in UPL under California law).

186. LegalZoom was confronted by ethics charges and lawsuits in multiple states alleging that it was engaging in UPL when it offered "do it yourself" interactive legal documents to the public, leading one commentator to suggest that UPL rules were outdated even in 2019, well before the current GenAI and agentic AI boom. Caroline Shipman, *Unauthorized Practice of Law Claims Against LegalZoom: Who Do These Lawsuits Protect, and is the Rule Outdated?*, 32 GEORG. J. LEG. ETHICS 939, 940–43, 952–54 (2019).

187. *Id.* at 944–45; *DoNotPay*, *supra* note 185.

188. See CALIFORNIA PRIVACY PROTECTION AGENCY, FACT SHEET: DRAFT AUTOMATED DECISIONMAKING TECHNOLOGY (ADMT) REGULATIONS 1 (2024–2025), <https://cpa.ca.gov>.

D. *Maintaining the Integrity of the Profession*

The final set of rules governs the lawyer's overarching duty to uphold the integrity of the profession, prohibiting misconduct and ensuring truthful communications.¹⁸⁹ AI introduces new ways these duties can be breached, often unintentionally.¹⁹⁰

1. Rule 8.4 (Misconduct): Deceit, Dishonesty, and Unwitting Reliance on Flawed AI

Model Rule 8.4 defines professional misconduct in broad terms.¹⁹¹ Most relevant to AI are Rule 8.4(c), which makes it misconduct for a lawyer to engage in conduct involving dishonesty, fraud, deceit or misrepresentation;¹⁹² and Rule 8.4(g), which prohibits engag[ing] in conduct that the lawyer knows or reasonably should know is harassment or discrimination.¹⁹²

ca.gov/meetings/materials/adt_regulations.pdf [https://perma.cc/UH6K-BCC8] (defining "automated decisionmaking technology" (ADMT) as technology that replaces or substantially replaces human decision-making and as technology such as AI and profiling systems that processes personal information and makes or substantially replaces human decision-making, and explains that businesses subject to the California Consumer Privacy Act would be required under the draft rules to provide pre-use notices, opt-out mechanisms, and access disclosures to consumers about how ADMT is used and how decisions are made, although the draft regulations were not yet in formal rulemaking or effect at the time of publication); *see also* CALIFORNIA PRIVACY PROTECTION AGENCY, DRAFT AUTOMATED DECISIONMAKING TECHNOLOGY REGULATIONS 3 (2023), https://coppa.ca.gov/meetings/materials/20231208_item2_draft.pdf [https://perma.cc/U9JY-ZS25] (proposing definitions and business obligations under the then-draft rules defining "automated decisionmaking technology" to include systems that process personal information to make or assist decisions, require clear pre-use notices to consumers about how ADMT is used and their opt-out and access rights, and outline detailed opt-out procedures and disclosures for decisions with legal or significant effects while emphasizing that the draft was for board discussion and public participation and not yet in formal rulemaking or effect).

189. MODEL RULES OF PROFESSIONAL CONDUCT R. 8.4 (AM. BAR ASS'N 2026).

190. *E.g.*, Jane Susskind, *AI-Generated Deficiencies in Filings: Sanctions and How to Avoid Them*, CLARK CNTY. BAR ASS'N (July 3, 2025), <https://clarkcountybar.org/ai-generated-deficiencies-in-filings-sanctions-and-how-to-avoid-> [https://perma.cc/5RJH-9YQD] (describing risks of unintentional misconduct through AI).

191. MODEL RULES OF PROFESSIONAL CONDUCT R. 8.4.

192. *Id.* R. 8.4(c), (g).

a. Unwitting Deceit

A lawyer can violate Rule 8.4(c) without possessing a specific intent to deceive.¹⁹³ The act of submitting a court filing with hallucinated case law, as occurred in the *Mata* case, is a misrepresentation to the tribunal, plain and simple.¹⁹⁴ The lawyers' claims of ignorance about ChatGPT's capacity for error were unavailing because their failure to verify the information constituted a reckless disregard for the truth.¹⁹⁵ This recklessness, this failure to perform basic due diligence before making a representation to a court, is itself a form of conduct that involves deceit and misrepresentation and thus constitutes misconduct.¹⁹⁶

b. Bias as Misconduct

The use of AI can also implicate Rule 8.4(g)'s prohibition on discrimination.¹⁹⁷ It is well-documented that AI models trained on historical data can inherit and perpetuate the biases present in that data.¹⁹⁸ If a law firm uses an AI tool for tasks like screening resumes for hiring or evaluating potential clients, and that tool is known or reasonably should be known to have a discriminatory bias against individuals based on race, gender, or

193. See Susskind, *supra* note 190 (analyzing how courts are imposing monetary and non-monetary sanctions— including fines, pro hac vice revocations, apologies, and professional discipline— on attorneys who file briefs with AI-generated errors like fabricated citations, and offers practical advice that lawyers must verify AI outputs, understand AI's limitations, and take remedial steps, such as correcting filings and being transparent, to ethically incorporate AI into legal practice).

194. See *Mata v. Avianca, Inc.*, 678 F. Supp. 3d 443, 448–49, 460 (S.D.N.Y. 2023) (sanctioning Plaintiff's attorneys under FED. R. CIV. P. 11— imposing a \$5,000 fine and requiring remedial measures— after they filed briefs citing numerous non-existent judicial opinions generated by ChatGPT and then failed to verify or correct the fabricated citations when their authenticity was challenged, underscoring that lawyers retain responsibility to ensure the accuracy of legal authority even when assisted by generative AI).

195. *Id.* at 451 (noting confusion over the contents of submissions still can constitute subjective bad faith); *Ethics of Using Artificial Intelligence in Practice*, N.H. BAR ASS'N (July 2023), <https://www.nhbar.org/using-artificial-intelligence-in-practice/> [<https://perma.cc/L4NU-J635>] (summarizing *Mata* and noting lawyers' unavailing claims of ignorance about ChatGPT).

196. Susskind, *supra* note 190.

197. See Keith Robert Fisher, *ABA Ethics Opinion Generative AI Offers Useful Framework*, A.B.A. BUS. L. SEC. (Oct. 3, 2024), https://www.americanbar.org/groups/business_law/resources/business-law-today/2024-october/aba-ethics-opinion-generative-ai-offers-useful-framework [<https://perma.cc/DEW6-RFJW>] (reviewing the applications of the Model Rules with AI).

198. See Jon Kleinberg, et al., *Discrimination in the Age of Algorithms*, ARXIV, 23, 30 (Feb. 5, 2019), <https://arxiv.org/abs/1902.03731> [<https://perma.cc/984T-9D88>] (A major problem that occurs repeatedly, and is often overlooked, is that we can only predict outcomes using training observations about those for whom we observe the outcome.).

another protected characteristic, the continued use of that tool could constitute professional misconduct under Rule 8.4(g).¹⁹⁹

The very nature of some advanced AI systems creates a novel and troubling ethical dilemma: the risk of unknowable misconduct caused by the “black box” nature of AI systems, where even the model designers and creators cannot fully explain why and how the model does things.²⁰⁰ Rule 8.4 generally implies some level of knowledge, or at least recklessness, on the part of the lawyer.²⁰¹ However, as stated, some complex AI models are referred to as having a black box nature, meaning that their internal decision-making processes are so intricate that even their own creators cannot fully explain how they arrived at a specific output.²⁰² If a lawyer relies on such an opaque tool for a critical legal judgment— for instance, a settlement valuation or a risk assessment— and the tool’s reasoning is secretly flawed or biased in a way that is practically undetectable, has the lawyer committed misconduct?²⁰³ This possibility pushes the ethical burden upstream, back to the duty of competence under Rule 1.1. It suggests that lawyers have an affirmative duty to favor AI tools that are transparent and explainable (white box systems) over those that are opaque.²⁰⁴ Relying on a

199. Fisher, *supra* note 197.

200. See Matthew Kosinski, *What Is Black Box Artificial Intelligence (AI)?*, IBM (Oct. 29, 2024), <https://www.ibm.com/think/topics/black-box-ai> [<https://perma.cc/AQ34-32VM>] (“A black box AI is an AI system whose internal workings are a mystery to its users. Users can see the system’s inputs and outputs, but they can’t see what happens within the AI tool to produce those outputs.”).

201. See MODEL RULES OF PROFESSIONAL CONDUCT R. 8.4(a) (AM. BAR ASS’N 2026) (defining the “knowing” standard). *But see* CAL. BAR PRACTICAL GUIDANCE, *supra* note 32, at 4 (failing to define an explicit “knowing” standard).

202. Kosinski, *supra* note 200; Miss. Bar, Formal Op. 267 (2014), <https://www.msbar.org/ethics-discipline/ethics-opinions/formal-opinions/267/> [<https://perma.cc/E2J3-PSCF>].

203. See CAL. BAR PRACTICAL GUIDANCE, *supra* note 32, at 4 (underscoring that generative AI must be used in a manner consistent with existing professional responsibility obligations— such as duties of confidentiality, competence, diligence, supervision, communication, candor, compliance with the law, and nondiscrimination— requiring lawyers to understand AI risks and limitations, carefully review outputs, safeguard client information, and apply their own judgment even as the technology and regulatory landscape evolve); Miss. Bar, Formal Op. 267 (2014) (discussing the lawyer’s responsible use of generative AI).

204. Lindemulder & Badman, *supra* note 157; Paul, *supra* note 157; Pearl, *supra* note 161; *Top Tools and Techniques for Debugging Agentic AI Systems*, *supra* note 160.

process that is fundamentally unknowable is an inherently risky, and therefore potentially reckless, act.²⁰⁵

2. Rule 7.1 (Communications Concerning a Lawyer's Services):
Ensuring Truthfulness in AI-Generated Marketing

Model Rule 7.1 is straightforward: "A lawyer shall not make a false or misleading communication about the lawyer or the lawyer's services."²⁰⁶ "A communication is false or misleading if it contains a material misrepresentation of fact or law or omits a fact necessary to make the statement as a whole not materially misleading."²⁰⁷

a. AI-Generated Hype

Law firms are increasingly using generative AI to create marketing content, such as website copy, blog posts, social media updates, and client alerts.²⁰⁸ While this can be an efficient way to generate content, the ultimate responsibility for its truthfulness remains with the lawyer.²⁰⁹ An AI tool, tasked with creating a flattering biography, might generate text that describes a lawyer as an "expert" or "specialist" in a practice area where they do not meet the specific jurisdictional requirements for such a designation. It might create content that implies a guaranteed result (such as winning all of a lawyer's cases), which is a clear violation. The lawyer must review and edit all AI-generated marketing materials to ensure they are accurate and comply with Rule 7.1.

b. Misleading Claims About AI Itself

A firm's communications about its own use of technology must also be truthful. It would be misleading for a firm to claim it uses cutting-edge

205. See Vinky G., *The Risks of Agentic AI: Unintended Consequences of Autonomous Decision-Making*, RPA TECH (Dec. 19, 2024), <https://www.rpatech.ai/risks-of-agentic-ai/> [<https://perma.cc/V5T5-RYFW>] ("In scenarios where AI systems are tasked with making ethical decisions, such as healthcare or criminal justice, Agentic AI may make decisions that conflict with human ethical standards.")

206. MODEL RULES OF PROFESSIONAL CONDUCT R. 7.1 (AM. BAR ASS'N 2026).

207. *Id.*

208. Fla. Bar Ethics Op. 24-1 (2024) (noting law firms are using generative AI and its use implicates ethical rules including those regulating advertising).

209. Micah Buchdahl, *LP Magazine* "Ethical Issues When Incorporating AI Into Law Firm Marketing," MKTG. ATTY. BLOG (Oct. 8, 2024), <https://www.marketingattorney.com/lp-magazine-ethical-issues-when-incorporating-ai-into-law-firm-marketing> [<https://perma.cc/8QLF-KEHJ>].

proprietary AI to deliver superior results if, in reality, its lawyers are merely using the free, public version of ChatGPT.²¹⁰ Furthermore, as AI-powered chatbots become more common for website intake, the duty to avoid misleading communication is critical. Florida's ethics opinion specifically highlights this, stating that such chatbots must clearly and conspicuously disclose that they are an AI program and not a lawyer or employee of the firm, to ensure that prospective clients are not misled about who or what they are communicating with.²¹¹

TABLE 1: COMPARATIVE ANALYSIS OF KEY STATE BAR GUIDELINES ON GENERATIVE AI USE

The following table summarizes the guidance issued by the ABA and several leading state bar associations, highlighting areas of consensus and divergence in the emerging ethical landscape for AI in legal practice.

<p>ABA Formal Opinion 512²¹² July 2024 Duties: Competence, Confidentiality, Supervision, Candor, Fees AI Disclosure required? Case-dependent; required when confidentiality is at risk or AI use is substantive. AI Billing Guidance: AI efficiency should reduce hourly fees; cannot charge for learning time.</p>
<p>Alaska Ethics Opinion 2025-1²¹³ Apr. 2025 Duties: Competence, Confidentiality, Communication, Supervision, Fees AI Disclosure required? Recommended based on client matter and nature and scope of GAI use. AI Billing Guidance: Must disclose the basis on which the charges will be computed. May not bill clients for time saved that the lawyer did not work.</p>

210. See Fla. Bar Ethics Op. 24-1 (2024) ("Lawyers may advertise their use of generative AI but cannot claim their generative AI is superior to those used by other lawyers or law firms unless the lawyer's claims are objectively verifiable").

211. *Id.*

212. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 512, at 10 (2024).

213. Alaska Bar Ass'n Ethics Op. 2025-1 (2025).

California Practical Guidance²¹⁴ □ Nov. 2023

Duties: Confidentiality, Competence, AI Disclosure, Supervision, Candor, Anti-Discrimination

AI Disclosure required? Not mandatory but recommended based on client sophistication and risk.

AI Billing Guidance: Efficiency gains must be passed on to clients; cannot charge for time saved.

D.C. Bar Ethics Opinion 388²¹⁵ □ Sep. 2024

Duties: Competence, AI Verification, Supervision, Client Files

AI Disclosure required? Required in specific situations.

AI Billing Guidance: No excess fees for AI use.

Florida Advisory Opinion 24-1²¹⁶ □ Jan. 2024

Duties: Confidentiality, Supervision, Fees, AI Transparency in Advertising

AI Disclosure required? Informed consent required before disclosing confidential information to a third-party AI system.

AI Billing Guidance: Must not charge for AI time savings; must inform client in writing to charge for AI tool costs.

Kentucky Ethics Opinion E-457²¹⁷ □ Mar. 2024

Duties: Competence, AI Oversight, Client Communication, Confidentiality, Billing

AI Disclosure required? No ethical duty to disclose AI use unless the work is being outsourced to a third party.

AI Billing Guidance: Must not charge for AI time savings; must inform client in writing to charge for AI tool costs.

New York City Bar Formal Opinion 2024-5²¹⁸ □ Aug. 2024

Duties: Confidentiality, Competence, Supervision, Conflicts, Advertising

214. CAL. BAR PRACTICAL GUIDANCE, *supra* note 32, at 2□4.

215. D.C. Bar Ethics Op. 388 (2024).

216. Fla. Bar Ethics Op. 24-1 (2024).

217. Ky. Bar Ass'n Ethics Op. KBA E-457 (2024).

218. N.Y.C. Bar Ass'n Comm. on Pro. Ethics, Formal Op. 2024-5 (2024).

AI Disclosure required? Required for external AI use involving client information; not for routine internal tools.

AI Billing Guidance: No overcharging for AI use; fees must be reasonable.

New York State Bar Task Force Report²¹⁹ □ Apr. 2024

Duties: Competence, Bias, AI Oversight, Client Commun., Confidentiality

AI Disclosure required? Recommended to disclose potential AI use in engagement letters.

AI Billing Guidance: Clients should benefit from AI-driven efficiency.

New Jersey Supreme Court Prelim. Guidelines²²⁰ □ Jan. 2024

Duties: Competence, Candor, Client Commun., AI Hallucinations, Confidentiality

AI Disclosure required? Not required, unless necessary to permit the client to make informed decisions.

AI Billing Guidance: Use of AI likely will affect lawyer billing practices.

North Carolina Formal Ethics Opinion 2024-1²²¹ □ Nov. 2024

Duties: Competence, Supervision, Confidentiality, AI Oversight

AI Disclosure required? Yes, in some instances, to allow client autonomy and consent; must protect confidentiality.

AI Billing Guidance: AI-based efficiency should lower fees and costs.

Pa. & Phila. Bars Joint Opinion 2024-200²²² □ June 2024

Duties: Competence, Candor, AI Hallucinations, Confidentiality, Billing

AI Disclosure required? Yes, in some AI-use cases, to ensure transparency.

AI Billing Guidance: Cannot bill for AI time as if it were manual work; expenses must be reasonable.

219. NYSBA REPORT AND RECOMMENDATION, *supra* note 34.

220. N.J. COMM. ON AI AND THE CTS., *supra* note 128.

221. N.C. State Bar Comm. on Ethics & Pro. Resp., Formal Op. 2024-1 (2024).

222. Pa. Bar Ass'n Comm. on Legal Ethics and Pro. Resp. & Phila. Bar Ass'n Pro. Guidance Comm., Joint Formal Op. 2024-200 (2024).

Texas Ethics Committee Opinion 705²²³ Feb. 2025

Duties: Competence, Candor, AI Hallucinations, Confidentiality, Billing

AI Disclosure required? Lawyer should consider informing clients about the associated risks and may need to secure client consent.

AI Billing Guidance: Lawyer may charge for the actual time spent using a generative AI program. Lawyer may not charge for the time that was saved.

USPTO Practice Guidance (2023-2024) Apr. 2024

Duties: Candor, Confidentiality, AI Use in Legal Submissions

AI Disclosure required? Court compliance required.

AI Billing Guidance: Legal AI use cannot inflate costs.

Virginia AI Guidance Update²²⁴ Aug. 2024

Duties: Confidentiality, Billing, Supervision, AI Court Compliance

AI Disclosure required? Not mandated but recommended.

AI Billing Guidance: AI costs must align with ethical billing.

III. UPHOLDING THE MODEL CODE OF JUDICIAL CONDUCT WITH AI

The integration of AI into the judicial branch presents a distinct set of ethical challenges that strike at the heart of public trust in the justice system.²²⁵ The Model Code of Judicial Conduct, which governs the behavior of judges, provides the essential framework for navigating this new terrain.²²⁶

223. Tex. Comm. on Pro. Ethics, Op. 705 (2025).

224. Nicole Black, *Practical and Adaptable AI Guidance Arrives from the Virginia State Bar*, N.Y. L. BLOG (Sep. 3, 2024), <https://nylawblog.typepad.com/suigeneris/2024/09/practical-and-adaptable-ai-guidance-arrives-from-the-virginia-state-bar.html> [<https://perma.cc/W2DE-52WE>].

225. *See AI & the Courts: Judicial and Legal Ethics Issues*, NAT'L CENTER FOR STATE CTS., <https://www.ncsc.org/resources-courts/ai-courts-judicial-and-legal-ethics-issues> [<https://perma.cc/YH34-BC5C>] (explaining the challenges of AI integration and how this process interacts with existing legal responsibilities and principles).

226. *See generally* MODEL CODE OF JUD. CONDUCT (AM. BAR. ASS'N 2026) (providing canons that guide judges as to their baseline conduct requirements in the court of law).

A. *Canon 1: Upholding the Independence, Integrity, and Impartiality of the Judiciary*

Canon 1 is the bedrock of judicial ethics.²²⁷ It commands that “[a] judge shall uphold and promote the independence, integrity, and impartiality of the judiciary, and shall avoid impropriety and the appearance of impropriety.”²²⁸ This canon is concerned not only with actual misconduct but also with public perception, because, as the commentary notes, “[p]ublic confidence in the judiciary is eroded by improper conduct and conduct that creates the appearance of impropriety. This principle applies to both the professional and personal conduct of a judge.”²²⁹

1. The Appearance of Impropriety and “Black Box” Algorithms

The use of certain AI tools by judges can create a significant appearance of impropriety.²³⁰ This is particularly true of so-called “black box” algorithms, where the logic and data used to arrive at a recommendation are proprietary, opaque, and not available for inspection by the parties or the public.²³¹ Justice must not only be done, but it must “manifestly and undoubtedly be seen to be done.”²³² When a judge relies on a secret, inscrutable algorithm—perhaps one provided by a private, for-profit company—to inform a decision on matters as critical as bail, sentencing, or child custody, the public cannot see how justice is being done.²³³ This opacity undermines public confidence and creates the appearance that the judge is outsourcing their judicial function to an unaccountable

227. *See id.* (emphasizing a judge’s duty to promote integrity in the judiciary and avoidance of impropriety).

228. *Id.* Canon 1.

229. *Id.* R. 1.2 cmt. 1

230. *See id.* Canon 1 (stating that judges should avoid the appearance of impropriety).

231. *See* Jenna Burrell, *How the Machine “Thinks”: Understanding Opacity in Machine Learning*, 3 *BIG DATA & SOCIETY* 1, 1 (2016) (introducing different reasons that machine learning algorithms may be opaque).

232. A phrase that draws from the wisdom of *Rex v. Sussex Justices, ex parte McCarthy*, [1924] 1 K.B. 256, 259 (Hewart C.J.), https://www.iclr.co.uk/wp-content/uploads/media/vote/1915-1945/McCarthy_kb1924-1-256.pdf [<https://perma.cc/48DC-RF27>]. *Speech by Mr. Justice Cobb: “Justice Must Be Seen to Be Done”*, CTS. AND TRIBUNALS JUDICIARY (Oct. 16, 2024), <https://www.judiciary.uk/speech-by-mr-justice-cobb-justice-must-be-seen-to-be-done/> [<https://perma.cc/6CAM-U2QN>].

233. *See, e.g.*, *State v. Loomis*, 881 N.W.2d 749, 757–76 (Wis. 2016) (discussing due process concerns with proprietary COMPAS risk assessment).

technological process, thereby compromising the integrity of the court.²³⁴

2. Independence from Undue Influence

A judge's independence— their ability to act without fear or favor— can also be compromised by an over-reliance on AI tools.²³⁵ Two well-documented cognitive biases are particularly relevant here:

- **Automation Bias:** This is the human tendency to over-trust and uncritically accept decisions made by automated systems, even when contradictory information is available.²³⁶ A judge, faced with a heavy caseload, may be tempted to defer to an AI's recommendation rather than engage in a more demanding independent analysis.²³⁷
- **Anchoring Effect:** This is the tendency to over-rely on the first piece of information received.²³⁸ If a judge is presented with an AI-generated "recidivism risk score" for a defendant before hearing arguments from the prosecution and defense, that initial number can act as a powerful cognitive anchor, unconsciously influencing their subsequent judgment and making it difficult to deviate from the AI's initial assessment.²³⁹

234. See Recent Case, *Criminal Law* — *Sentencing Guidelines* — *Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessments in Sentencing*, 881 N.W.2d 749 (*Wis.* 2016), 130 HARV. L. REV. 1530, 1531–36 (2017) [hereinafter Recent Case, *Criminal Law*] (summarizing transparency and accountability concerns when methodology is not disclosed).

235. See CONF. OF STATE CT. ADM'RS, GENERATIVE AI AND THE FUTURE OF THE COURTS: RESPONSIBILITIES AND POSSIBILITIES 5–7 (2024) (explaining that it is imperative for judges to adapt to modern technology and be familiar with its application to maintain independence and integrity in the justice system).

236. See Linda J. Skitka, Kathleen L. Mosier & Mark Burdick, *Does Automation Bias Decision-making?*, 51 INT'L J. HUMAN-COMPUT. STUD. 991, 992 (1999) (introducing the idea that, as a result of automation, human error may actually increase in certain areas).

237. See MODEL CODE OF JUD. CONDUCT (AM. BAR. ASS'N 2026) (emphasizing the notion that judges have a duty to analyze risk when using AI in their judicial role).

238. See Birte English, Thomas Mussweiler & Fritz Strack, *Playing Dice with Criminal Sentences: The Influence of Irrelevant Anchors on Experts' Judicial Decision Making*, 32 PERSONALITY SOC. PSYCH. BULL. 188, 188 (2006) (explaining anchoring theory and how seemingly arbitrary factors included in a decision can influence human behavior).

239. See Dan Pilat, *Why Do We Compare Everything to the First Piece of Information We Received? The Anchoring Bias, Explained*, THE DECISION LAB, <https://thedecisionlab.com/biases/anchoring-bias> [<https://perma.cc/9DB7-84VW>] (expanding on how the first piece of information we receive can influence later decisions on the matter by acting as an unconscious reference point in the context of

This potential for AI to exert an undue influence on judicial decision-making poses a direct threat to the independence guaranteed by Canon 1.

B. *Canon 2: Performing Judicial Duties Impartially, Competently, and Diligently*

Canon 2 governs the judge's performance of their official duties, requiring them to act impartially, competently, and diligently.²⁴⁰ AI impacts each of these components.

1. Rule 2.2 (Impartiality) & Rule 2.3 (Bias): Confronting Algorithmic Bias in Judicial Decision-Support Tools

Rule 2.2 requires a judge to "uphold and apply the law, and [the judge] shall perform all duties of judicial office fairly and impartially."²⁴¹ Rule 2.3 explicitly commands that "[a] judge shall not, in the performance of judicial duties, by words or conduct manifest bias or prejudice. . . ."²⁴²

a. The Problem of Biased Data

The single greatest threat AI poses to judicial impartiality is algorithmic bias.²⁴³ AI systems are not inherently objective; they learn from the data they are trained on and can inherit its biases.²⁴⁴ If an AI tool used for sentencing recommendations or bail assessments is trained on historical judicial data, it will inevitably learn and replicate the systemic societal biases present in that data.²⁴⁵ This has been demonstrated empirically.²⁴⁶ The well-known COMPAS recidivism-risk assessment tool, used in courtrooms across the country, was shown in a landmark ProPublica investigation to be significantly more likely to falsely flag black defendants as being at high risk

AI); CONF. OF STATE CT. ADM'RS, *supra* note 238, at 506 (emphasizing that judges must understand the technical limits and risks of generative AI before relying on the information).

240. MODEL CODE OF JUD. CONDUCT Canon 2.

241. *Id.* R. 2.2.

242. *Id.* R. 2.3(B).

243. *See generally* SOLON BAROCAS, MORITZ HARDT & ARVIND NARAYANAN, FAIRNESS AND MACHINE LEARNING: LIMITATIONS AND OPPORTUNITIES 76-101 (2023) (highlighting the different types of unfairness and analyzing them under various scenarios such as criminal justice).

244. Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 4-5 (2014).

245. *See id.* (stating that humans are the ones who provide initial data at the beginning of the AI learning process, which could create bias as it continues making decisions).

246. *See id.* at 14 (citing a study concluding that credit scoring under a presumably objective predictive algorithm was found to discriminate against poor and minority groups).

of reoffending than white defendants.²⁴⁷ Using such a tool can lead to discriminatory outcomes under the guise of objective, data-driven analysis.²⁴⁸

b. The *Loomis* Precedent

The Wisconsin Supreme Court addressed this issue in the 2016 case of *State v. Loomis*,²⁴⁹ where the defendant argued that the use of the COMPAS tool at his sentencing violated his due process rights.²⁵⁰ The court held that the use of such a tool was permissible, provided that the presentence report given to the judge included written warnings about the tool's limitations.²⁵¹ However, critics have forcefully argued that these warnings are insufficient.²⁵² They contend that judges often lack the statistical expertise to truly evaluate the tools, face institutional pressure to use them for the sake of efficiency, and are susceptible to the cognitive biases that lead to over-reliance on their outputs.²⁵³

The use of AI in judicial functions fundamentally shifts the nature of a legal challenge to bias.²⁵⁴ A traditional motion for recusal argues that a *specific human judge* has a personal bias that prevents them from being impartial.²⁵⁵ When an AI decision-support tool is involved, the potential bias is no longer personal and idiosyncratic; it is systemic, statistical, and embedded in the

247. Julia Angwin et. al., *Machine Bias*, PRO PUBLICA (May 23, 2016), <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> [<https://perma.cc/QFK3-UEVA>]; *State v. Loomis*, 881 N.W.2d 749, 763 (Wis. 2016) (Abrahamson, J., concurring); Recent Case, *Criminal Law*, *supra* note 237, at 1533-35.

248. Angwin et. al., *supra* note 250.

249. *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016).

250. *Id.* at 757.

251. *Id.* at 770.

252. See Bernard E. Harcourt, *Risk As a Proxy for Race: The Dangers of Risk Assessment*, 27 FED. SENT'G. REP. 237, 237 (2015) (advocating that the current predictive risk tools utilized are still not up to equitable standards); Recent Case, *Criminal Law*, *supra* note 237, at 1530, 1533 (noting the ineffectiveness of the warnings and the widespread endorsement of the reliability of the tools).

253. Recent Case, *Criminal Law*, *supra* note 237, at 1533-36; see *Loomis*, 881 N.W.2d at 774 (Abrahamson, J., concurring) (describing how the court system is still in the unknown regarding the COMPAS technology).

254. See Citron & Pasquale, *supra* note 247, at 16-20 (explaining how to integrate new technology in scoring ethically).

255. MODEL CODE OF JUD. CONDUCT R. 2.11 (AM. BAR. ASS'N 2026).

tool's very code.²⁵⁶ A litigant's argument would no longer be, "Your Honor, you are biased," but rather, "Your Honor, the tool you are relying on is demonstrably biased against a protected class, and its use in my case violates my constitutional rights to due process and equal protection."²⁵⁷ This transforms the evidentiary basis for such a challenge.²⁵⁸ It moves the courtroom debate from questions of judicial temperament to questions of statistical validity, algorithmic fairness, and data provenance, requiring complex expert testimony from data scientists, statisticians, and computer scientists to resolve.²⁵⁹

2. Rule 2.5 (Competence and Diligence): The Judge's Duty of Technological Competence

Canon 2, Rule 2.5 requires that "[a] judge shall perform judicial and administrative duties, competently and diligently."²⁶⁰

a. Judicial Competence with AI

Just as the duty of competence for lawyers has expanded to include technological proficiency, so too has it for judges.²⁶¹ To competently use an AI tool, a judge must have a reasonable understanding of what the tool does, how it works, what its known limitations are, and for what purpose it is appropriate to use.²⁶² A judge who uses a risk assessment tool without understanding that it is probabilistic and not deterministic, or without being aware of its documented racial biases, is not performing their duties

256. See Citron & Pasquale, *supra* note 247, at 16–20 (describing the issues and bias within technology in scoring).

257. See Angwin et al., *supra* note 250 (describing the racial bias within new technology).

258. See Citron & Pasquale, *supra* note 247, at 22–26 (pushing for legislation that calls for the allowance of validity testing for predictive algorithms).

259. *Id.*; Hannah Webster, *Are Risk Assessment Tools Setting the Stage for AI Judges?*, THE BAIL PROJECT (June 28, 2024), <https://bailproject.org/learn/are-risk-assessment-tools-setting-the-stage-for-ai-judges/> [<https://perma.cc/R2L2-R5VY>]; Alix Faulkner, *From Case Law to Code: Evaluating AI's Role in the Justice System*, MONTREAL AI ETHICS INST. (May 26, 2025), <https://montrealthics.ai/from-case-law-to-code-evaluating-ais-role-in-the-justice-system/> [<https://perma.cc/3TJ9-CMHG>].

260. MODEL CODE OF JUD. CONDUCT R. 2.5 (AM. BAR. ASS'N 2026).

261. *Id.*; CONF. OF STATE CT. ADM'RS, *supra* note 238, at 5.

262. MODEL CODE OF JUD. CONDUCT R. 2.5; CONF. OF STATE CT. ADM'RS, *supra* note 238, at 5.

competently.²⁶³ Judicial education in this area is therefore essential.²⁶⁴

b. Diligence in Application

The duty of diligence of judges is also critical.²⁶⁵ A diligent judge will not simply accept an AI's output as dispositive.²⁶⁶ They will treat the AI's recommendation as a single data point among the many other pieces of evidence and argument presented in a case.²⁶⁷ Diligence requires the judge to exercise their own independent legal reasoning and moral judgment, critically assessing the AI's output in the context of the specific facts and human factors of the case before them.²⁶⁸

3. Rule 2.9 (Ex Parte Communications): Can an AI be an Improper External Influence?

Rule 2.9 strictly prohibits a judge from initiating, permitting, or considering ex parte communications—that is, communications concerning a pending case that are made to the judge outside the presence of the parties and their lawyers.²⁶⁹ The rule is designed to ensure that decisions are based only on the evidence and arguments presented in the adversarial process.²⁷⁰

a. AI as an External Influence

Canon 2, Rule 2.9 raises an intriguing question in the context of AI: If a judge, struggling with a complex legal issue in a pending case, were to consult a general-purpose AI like ChatGPT for analysis or research into the

263. See Angwin et al., *supra* note 250; Harcourt, *supra* note 255; Recent Case, *Criminal Law*, *supra* note 237, at 1536 (providing an example of a risk judges face when using predictive sentencing tools).

264. CONF. OF STATE CT. ADM'RS, *supra* note 238, at 5-6; *AI & the Courts: Judicial and Legal Ethics Issues*, *supra* note 228.

265. See MODEL CODE OF JUD. CONDUCT R. 2.5 (enumerating a judge's duty to act diligently in their role).

266. *AI & the Courts: Judicial and Legal Ethics Issues*, *supra* note 228.

267. See Herbert B. Dixon et al., *Navigating AI in the Judiciary: New Guidelines for Judges and Their Chambers*, 26 SEDONA CONF. J. 1, 5-8 (2025) (emphasizing that judges are ultimately still responsible for all orders with their name on it, and that the use of AI does not change this responsibility); see also Martin Sposato, *A Call for Caution and Evidence-Based Research on the Impact of Artificial Intelligence in Education*, 2 INT'L J. ORG. THEORY & BEHAVIOR 158, 160 (2025) (asserting that AI tools should enhance rather than replace human practices).

268. See Dixon et al., *supra* note 270 (emphasizing that AI output cannot be trusted blindly).

269. MODEL CODE OF JUD. CONDUCT R. 2.9.

270. A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 478 (2017).

factual history and context of the case before her, could this be considered an improper external communication? If the action is characterized as performing independent research into the facts²⁷¹ of a case, the answer is “yes.”²⁷² The AI is not a party to the case, nor is it a law clerk or staff attorney (i.e., a licensed attorney) working under the judge’s direction.²⁷³ It is an external source of information and analysis, and its input is being received and considered by the judge outside the adversarial process, without giving the parties an opportunity to respond to or rebut the AI’s reasoning.²⁷⁴

This concern becomes even more acute when considering proprietary, “black box” AI tools used in judicial decision-making.²⁷⁵ Such a tool can be analogized to an improper ex parte communication with the hidden “experts” who designed it.²⁷⁶ The fundamental problem with an ex parte communication is that one party does not have the opportunity to hear, challenge, and rebut the information being provided to the decision-maker.²⁷⁷ When a judge uses a proprietary AI tool, they are receiving an “opinion” or “recommendation” from the algorithm.²⁷⁸ The parties to the

271. It is important to note that independent research into the *law* by judges always is permitted. CHARLES G. GEYH ET AL., JUDICIAL CONDUCT AND ETHICS §5.04 (5th ed. 2013) (noting that “independent investigation of the law has always been permitted” and that judges are “experts on matters of law who are charged with the duty of declaring what the law is”).

272. See MODEL CODE OF JUD. CONDUCT R. 2.9(C) [(stating independent research by a judge is prohibited); see also *AI & the Courts: Judicial and Legal Ethics Issues*, *supra* note 228 (explaining AI use by judges may create issues with Rule 2.9); A.B.A. Comm. on Ethics & Pro. Resp., Formal Op. 478 (2017) (noting a judge may not conduct research into the facts via the internet)].

273. These labels refer to exceptions to the ex parte communications rule. MODEL CODE OF JUD. CONDUCT R. 2.9(A)(1), (3).

274. See *AI & the Courts: Judicial and Legal Ethics Issues*, *supra* note 228 (explaining how AI could be in violation of judicial conduct regarding ex parte communications).

275. See Burrell, *supra* note 234, at 1 (explaining black box code and why AI companies can maintain privacy in their proprietary methods); Matthew T. Christ, *The Hidden Dark Side of AI in the Legal System*, AI J. (Mar. 6, 2025), <https://aijourn.com/the-hidden-dark-side-of-ai-in-the-legal-system/> [<https://perma.cc/7NUD-PWQC>] (addressing the “black box” issue arises when judges and attorneys cannot access risk assessments).

276. See Christ, *supra* note 278 (analogizing how black box code could be considered ex parte due to the confidential and inaccessible nature of the code itself).

277. *Id.*

278. See Jeff Ward, *10 Things Judges Should Know About AI*, JUDICATURE (2019), <https://judicature.duke.edu/articles/10-things-judges-should-know-about-ai> [<https://perma.cc/5DL7-7QSJ>] (contrasting human decisions based on behavioral experiences and opinions with AI

case, however, do not know the vendor's methodology, the data the tool was trained on, the factors it considers, or the weight it assigns to them.²⁷⁹ They cannot cross-examine the algorithm's logic or challenge its hidden assumptions.²⁸⁰ Therefore, a judge's reliance on this opaque, third-party analysis is functionally equivalent to having a private, un-rebuttable conversation with an undisclosed team of experts. This raises profound concerns under Rule 2.9 and implicates the core due process rights of the litigants.²⁸¹

IV. CONCLUSION: CHARTING A COURSE FOR ETHICAL AI INTEGRATION IN THE AMERICAN LEGAL SYSTEM

The integration of generative and agentic artificial intelligence into the fabric of the American legal system is not a distant prospect; it is a present and rapidly accelerating reality. As this Article has detailed, this transformation challenges the legal profession to apply its most foundational ethical principles to a new and complex set of technological circumstances. The analysis, structured through the Model Rules of Professional Conduct and the Model Code of Judicial Conduct, reveals that while the tools are new, the core duties of competence, confidentiality, candor, supervision, and judicial impartiality remain the essential lodestars. The path forward requires not a revolution in ethics, but a renewed commitment to their diligent application.

A. *Key Recommendations for Practitioners, Firms, and the Judiciary*

Based on the analysis of the rules and the emerging consensus from ethics bodies, the following recommendations provide a course for responsible AI

decisions that are based on opaque inputted data); Faulkner, *supra* note 262 (discussing the exploration of AI sentencing tools to assist judges).

279. Faulkner, *supra* note 262; Burrell, *supra* note 234; Christ, *supra* note 278.

280. John M. McNichols, *How Do You Cross-Examine Siri If You Think She's Lying?*, 47 A.B.A. LITIG. NEWS 2, 2-3 (2022); see Faulkner, *supra* note 262 (noting concerns surrounding transparency and accountability); Christ, *supra* note 278 (discussing the lack of transparency and accountability for "black box" AI).

281. See Chris Chambers Goodman, *AI, Can You Hear Me? Promoting Procedural Due Process in Government Use of Artificial Intelligence Technologies*, 28 RICH. J. L. & TECH. 700, 702-07 (2022) (describing how a lack of transparency in the methods AI uses to come to conclusions may create due process violations).

integration:

For Individual Practitioners:

- **Embrace Continuous Education:** The duty of competence is now dynamic. Lawyers must commit to ongoing learning about the benefits, risks, and limitations of the AI tools relevant to their practice. This includes understanding the difference between open and closed systems, the danger of hallucinations, and the capabilities of agentic AI.
- **Adopt a “Trust but Verify” Mindset:** Never delegate professional judgment to an AI. Treat all AI-generated output—whether legal research, factual summaries, or draft arguments—as a starting point provided by an unsupervised junior assistant. Every fact and every legal citation must be independently verified before it is used.
- **Prioritize Confidentiality:** Default to using “closed” AI systems designed for legal practice that guarantee data privacy. When using any other tool, meticulously anonymize all client information and obtain specific, informed client consent if any risk to confidentiality exists.

For Law Firms and Legal Organizations:

- **Establish Systemic Governance:** Move from ad-hoc AI use to a formal governance structure. This requires drafting and implementing clear, written AI usage policies that are communicated to all personnel. These policies should specify approved vendors, data handling protocols, verification requirements, and billing practices.
- **Conduct Rigorous Vendor Due Diligence:** Before adopting any third-party AI platform, the firm must conduct thorough due diligence on the vendor’s data security, confidentiality policies, and the reliability of their technology. This is a core supervisory duty under Rules 5.1 and 5.3.
- **Invest in Mandatory Training:** Firms must provide regular, mandatory training for all lawyers and staff on the firm’s AI policies and the ethical obligations associated with using these tools. The sanctions against entire firms highlight that ignorance is not a defense.

For the Judiciary:

- **Demand Transparency and Explainability:** Judges and court systems should refuse to use any "black box" AI decision-support tool. Any algorithm used to inform judicial decisions, particularly in criminal matters, must be transparent, with its source code, training data, and logic available for inspection and challenge by the parties.
- **Foster Judicial Education:** The judiciary must invest in comprehensive education for judges on the nature of AI, the risks of cognitive and algorithmic bias, and the proper, limited role of these tools in supporting—not making—judicial decisions.
- **Exercise Cautious Diligence:** Judges must consciously guard against automation bias and the anchoring effect. AI-generated outputs should be treated as a single, fallible piece of information to be weighed against all other evidence and argument, with the final decision resting solely on the judge's independent and impartial judgment.

B. *The Future of Legal Ethics, Self-Regulation, and Legal Education*

The challenges posed by AI will only intensify as the technology advances from generative assistance to agentic autonomy. Agentic AI will continue to push the boundaries of what it means to practice law, forcing the profession to constantly re-evaluate the lines between a tool and an agent, between delegation and abdication, and between efficiency and the core exercise of professional judgment.

The legal profession's tradition of self-regulation is being tested. To maintain its relative autonomy, the bar must demonstrate that it can effectively govern the use of these powerful new technologies in the public interest. This will require a proactive, collaborative effort among the ABA, state bar associations, law schools, and the courts to develop and refine ethical guidance, create new educational curricula, and ensure that the promise of AI is harnessed to enhance justice, not to undermine it. The ultimate challenge is to leverage the immense power of artificial intelligence to improve the delivery of legal services and the administration of justice, without sacrificing the human values of fairness, empathy, integrity, and wisdom that lie at the very heart of the rule of law.