ADOPTING ARTIFICIAL INTELLIGENCE

A Practical Guide for Law Firms

Use Cases, Ethics, Governance, and a Roadmap to Get Started

Colin S. Levy

About This Guide

This guide is written for lawyers who have heard a great deal about artificial intelligence and want a careful, plain-English explanation of what it is, what it can and cannot do in a law firm, what the ethics rules require, how to govern its use, and how to actually begin. It assumes no prior technical background. It is meant to be read cover to cover by a managing partner, scanned by a practice group leader, or used as a reference by a firm's general counsel and innovation team.

The guide is organized in five parts. Part I explains the technology in lay terms. Part II surveys the use cases that have been tested in the practice of law and notes which are mature, which are emerging, and which remain unreliable. Part III addresses the professional responsibility issues, with reference to the American Bar Association Formal Opinion 512[1] and to recent state bar guidance. Part IV sets out a governance framework, including written policies, intake review, vendor diligence, training, and monitoring. Part V is an implementation roadmap with practical milestones for the first one hundred and eighty days.

Citations appear as superscript markers throughout the body and are collected in the Endnotes section at the back of the guide. Where opinions are stated, they are the authors' and not the position of any bar.

Executive Summary

Generative artificial intelligence, the branch of AI that produces fluent text and increasingly capable analysis, has matured to the point that it is being used in production at most of the AmLaw 200 and a rising share of midsize and boutique firms.[2] The technology is genuinely useful for several core legal tasks, including first-pass document review, deposition and transcript summarization, brief and memorandum drafting, contract abstraction, and legal research synthesis. It is also genuinely unreliable in ways that have already produced sanctions, judicial referrals, and embarrassing public errors.[3] The difference between a firm that benefits and a firm that is harmed is almost entirely a function of governance, training, and tool selection.

The headline points are these. First, every American jurisdiction now treats lawyer use of generative AI as falling within existing duties of competence, confidentiality, supervision, candor to the tribunal, and reasonable fees, rather than requiring a new rule.[4] Second, the duty of competence affirmatively requires lawyers to keep up with technology that bears on their practice, which includes a working understanding of generative AI.[5] Third, the duty of confidentiality requires that any AI tool used to handle client information be procured under terms that prohibit training on firm data and that meet the firm's standards for vendor security. Fourth, lawyers remain responsible for every word in every document filed or sent under their name. Fifth, billing for time saved by AI on an hourly basis is not categorically prohibited but is increasingly disfavored, and clients are beginning to demand transparency in writing.

A firm should not wait to act. The cost of a written policy, an enterprise license, two hours of mandatory training, and a quarterly governance committee meeting is small. The cost of a single sanctioned filing, a confidentiality breach, or a regulatory misstep is not.

Part I. Understanding the Technology

1.1 What "AI" Means in 2026

Artificial intelligence is an umbrella term for software that performs tasks historically associated with human intelligence, such as recognizing patterns, classifying documents, predicting outcomes, and producing language. Within that umbrella, the form of AI that has driven the recent wave of legal use cases is the **large language model**, or LLM.[6]

A large language model is a piece of software that has been trained on a very large body of text. During training, the model learns the statistical relationships among words and phrases. Once trained, the model is given a prompt, which is just a piece of text supplied by the user, and it produces an output one token at a time by predicting what is most likely to come next given everything it has seen so far. The result reads like fluent, often very capable, prose. The most widely used examples in 2026 are the GPT family from OpenAI, the Claude family from Anthropic, the Gemini family from Google, and the Llama family from Meta. Specialized legal-domain products, including Harvey, CoCounsel, Hebbia, and Spellbook, are typically built on top of one or more of these underlying models with additional retrieval, search, and workflow features layered on.

Two technical concepts matter for any lawyer evaluating these tools. The first is the **context window**, which is the amount of text the model can take into account at one time. Modern enterprise legal tools commonly accept hundreds of pages, and frontier models now handle the equivalent of a small brief library. The second is **retrieval-augmented generation**, or RAG, in which the model is given source documents at query time and instructed to base its answer on those sources.[7] Retrieval substantially reduces, but does not eliminate, the model's tendency to fabricate.

1.2 What These Models Cannot Do

Large language models do not reason in the way a human associate does. They do not check their work against a reliable database unless one is connected to them. They do not know what is true. They produce text that is plausible given their training, which is a different thing from text that is correct. The most important practical consequence is that any model can **hallucinate**, meaning it can produce a confident, well-formatted, fluently written assertion that is simply false.[8] This includes fabricated case citations, invented statutory section numbers, made-up quotations, and invented procedural histories. The fact that a tool is marketed as "hallucination-free" does not make it so.[9]

Models also have a **training cutoff**, which is the last date on which they ingested data. Anything that happened after that date is unknown to the underlying model unless the product is integrated with a live legal research database or current web search. This means that, absent retrieval, an LLM cannot reliably

tell you whether a case has been overturned, whether a statute has been amended, or whether a settlement has closed.

Finally, models reflect biases present in their training data, can be steered by carefully crafted prompts, and can produce different answers to the same question on different runs. Treating their output as a draft prepared by an enthusiastic but unverified summer associate is a useful frame.

1.3 Vocabulary at a Glance

Term	What it means in practice
Model	The trained software that produces output. GPT, Claude, Gemini, and Llama are models.
Prompt	The instruction or question you give the model. The quality of the prompt strongly affects the quality of the answer.
Token	A piece of a word. Models read and write in tokens. Pricing and length limits are usually expressed in tokens. One thousand tokens is roughly seven hundred and fifty English words.
Context window	How much text the model can consider at once. Bigger is better for long documents.
Hallucination	Fluent output that sounds right but is factually wrong. The defining failure mode of LLMs.
Fine-tuning	Additional training of a base model on narrower data. A firm may fine-tune on its own templates or brand voice.
Retrieval (RAG)	Architecture in which source documents are pulled in at query time and the model answers from them. The standard pattern for legal use.
Agent	A system in which the model takes multi-step actions, such as opening files, running searches, and calling other software, with limited human oversight.
Zero data retention	A vendor commitment that prompts and outputs are not stored after the request is completed and are not used to train any model.
Enterprise tier	A paid product configuration with stronger security, no training on customer data, and contractual protections. The only acceptable tier for client matter work.

Part II. Use Cases in the Practice of Law

The honest answer to "what can AI do for a law firm" is that it depends on whether the task is one of language, where models are now strong, or one of reasoning over privileged facts and authority, where models still need close supervision. The matrix below organizes the most common use cases by maturity. "Mature" means the technology, when used through a reputable enterprise tool with proper inputs, performs at or near the level of a competent first-year associate, with human review. "Emerging" means it can be useful but requires careful evaluation. "Risky" means it is not yet reliable enough for unsupervised work product.[10]

2.1 Use Case Maturity Matrix

Use case	Maturity	Notes
Document summarization (depositions, transcripts, expert reports)	Mature	Strong with retrieval. Always sample-check against source for material facts.
First-pass contract review and abstraction	Mature	Best when configured against a defined playbook of clauses and positions.
Drafting routine correspondence and internal memos	Mature	A senior associate can finalize in a fraction of the time. Confidentiality controls required.
eDiscovery first-pass relevance and privilege screening	Mature	Continuation of predictive coding. Defensibility requires documented protocol and statistical sampling.
Legal research synthesis through specialized tools	Emerging	Specialized legal AI products are far better than general chat tools, but still hallucinate. Verify every cite.
Brief and motion drafting	Emerging	Useful for outlines, headings, and statement of facts from record. Cite checking remains a human task.
Due diligence reports	Emerging	Effective on contract abstraction; less reliable on judgment calls about materiality.
Client-facing chat assistants	Emerging	Possible for general information; unauthorized practice and confidentiality concerns require

Use case	Maturity	Notes
		careful design.
Predicting case outcomes or settlement values	Risky	Statistical tools exist, but generative models should not be used to give numerical predictions to clients.
Unsupervised legal advice to clients	Risky	Implicates competence, confidentiality, and unauthorized practice. Not appropriate today.
Autonomous "agentic" workflows acting on client systems	Risky	Promising but immature. Pilot only with strong scoping, logging, and rollback.

2.2 Practical Examples

A litigation team uses an enterprise tool to summarize a one-hundred-page deposition into a five-page outline keyed to an issue list. A senior associate, who already has the issue list, can verify the outline in roughly twenty minutes, where reading the underlying transcript would take several hours. The same tool can produce a witness-by-witness chronology across a multi-deponent case. This is a paradigmatic mature use.

A corporate team feeds a stack of two hundred vendor contracts into a contract abstraction tool configured against the firm's renewal-and-termination playbook. The tool returns a spreadsheet of governing law, term length, auto-renewal triggers, change-of-control provisions, and indemnification caps. A paralegal verifies the spreadsheet against the source contracts in a structured sampling protocol. The work product would have taken weeks of associate time to produce manually.

A regulatory team asks a general-purpose chat tool to "find recent SEC enforcement actions involving misuse of customer data." The tool returns five plausible-sounding case names, none of which exist. This is the failure mode that produced the sanctions in Mata v. Avianca[11] and the referral in Park v. Kim.[12] The lesson is that general-purpose chat tools are not legal research tools and should not be used as such.

Part III. Ethics and Professional Responsibility

3.1 The Operating Principle

The American Bar Association's Standing Committee on Ethics and Professional Responsibility issued Formal Opinion 512 in July 2024.[13] The opinion is the most authoritative ethics guidance to date and frames the analysis under the existing Model Rules. It is consistent with the views of the California, Florida, New York, New Jersey, and District of Columbia bars, among others.[14] The operating principle, in three sentences, is this. Lawyer use of generative AI is permissible. It is governed by the existing duties of competence, confidentiality, communication, supervision, candor, and reasonable fees. The lawyer remains professionally responsible for the work product the AI helps produce.

3.2 Duty of Competence (Rule 1.1)

Comment 8 to Model Rule 1.1 has, since 2012, instructed lawyers to keep abreast of the benefits and risks associated with relevant technology.[15] Generative AI is now relevant to virtually every practice. Competence in 2026 includes a working understanding of what an LLM does, what it cannot do, what hallucination is, and how to verify output. It does not require the partner to be able to fine-tune a model. It does require the partner to know enough not to file a brief drafted by ChatGPT without checking the citations.

Competence Checklist

- Confirm the lawyer using a tool has been trained on the tool's capabilities, limits, and approved use cases.
- Verify any factual or legal assertion produced by the model against an authoritative source before it leaves the firm.
- Document the verification, particularly for citations to authority. A pre-filing cite check by a human is now table stakes.
- Match the tool to the task. Do not use a general-purpose chat product for legal research.

3.3 Duty of Confidentiality (Rule 1.6)

Rule 1.6 prohibits a lawyer from disclosing information relating to the representation of a client without informed consent and requires the lawyer to make reasonable efforts to prevent inadvertent or unauthorized disclosure.[16] Pasting client documents into a free, consumer-grade chatbot can be a disclosure to a third party in violation of Rule 1.6. It can also expose the data to use as training material for the underlying model, which has implications that are difficult to undo.

The path to compliance is straightforward. Use only enterprise-tier products for matter work. Require, in writing, that the vendor will not use firm data to train any model and will retain data only for the period

necessary to deliver the service.[17] Procure under a master services agreement with a data processing addendum,[18] a security exhibit consistent with the firm's information security standards, and a clear allocation of responsibility for breach notification. Where outside counsel guidelines speak to AI use or to confidentiality of client data, follow them and obtain client consent in writing where required.

3.4 Duty to Supervise (Rules 5.1 and 5.3)

A partner is responsible for the work of associates and for the conduct of nonlawyer assistants and tools.[19] The 2012 amendments to Rule 5.3 made clear that the rule applies to nonlawyer services, which includes software tools. Practically, this means a firm should designate a lawyer responsible for AI tool selection, configuration, and review, should establish written work-product standards for AI-assisted output, and should ensure that supervising lawyers actually review output before it goes out the door. "AI did it" is not a defense to malpractice or to a Rule 5.3 violation.

3.5 Duty of Candor to the Tribunal (Rule 3.3)

A lawyer cannot knowingly make a false statement of fact or law to a tribunal and cannot offer evidence the lawyer knows to be false. Filing a brief with hallucinated citations violates this duty whether or not the lawyer used AI. A growing number of federal and state judges have issued standing orders requiring disclosure or certification regarding AI use.[20] Compliance is a matter of (a) checking the docket for any standing order at the case opening stage and (b) following firm policy on disclosure of AI assistance in court filings.

3.6 Communication and Reasonable Fees (Rules 1.4 and 1.5)

Whether to disclose AI use to a client is increasingly fact-specific. Formal Opinion 512 declines to mandate disclosure across the board, but it makes clear that disclosure is required when (a) the client has asked, (b) outside counsel guidelines require it, (c) confidential information will be processed by a third-party tool that requires consent, or (d) AI use materially affects the basis of the lawyer's fee.[21]

On fees, the rule is unchanged: a fee must be reasonable.[22] Two things follow. First, a lawyer may not bill the time it would have taken to perform a task manually when AI performed it in minutes. Second, charging a separate cost recovery fee for AI tools is permissible only if (a) the engagement letter clearly authorizes it and (b) the charge is consistent with the firm's general practice of cost recovery. Many firms are migrating toward fixed fees and outcome-based pricing for AI-leveraged work, which sidesteps the issue entirely.[23]

3.7 Bias, Fairness, and Unauthorized Practice

Lawyers should be aware that LLM output can reflect biases in the underlying training data. This matters most in client-intake screening, employment practices, and any context in which a model is used to make

or support decisions about people. Firms should not deploy AI in those contexts without a human in the loop and a documented review process. Lawyers should also avoid deploying client-facing tools that could be construed as providing legal advice without lawyer involvement, both for unauthorized practice reasons and because the output cannot meet the standard of care.[24]

Part IV. Building a Governance Framework

Most firms have not yet adopted a written generative AI policy.[25] That is the single highest-leverage gap. A workable governance framework has six components: a policy, an intake and approval process, vendor diligence, training and certification, monitoring and audit, and a designated committee. Each is described below in enough detail to draft from.

4.1 Component One: A Written AI Use Policy

The policy should be short, in plain English, and binding on every lawyer and staff member. It should be referenced in the firm's engagement letters where appropriate. The minimum content is the following.

1. Scope and definitions, including what counts as "generative AI" and which categories of tools the policy reaches.
2. Approved tools, by name, with the use cases each is approved for.
3. Prohibited uses, including the use of consumer-grade or unapproved tools for any client matter content.
4. Confidentiality rules, including the requirement that no client confidential information be entered into any tool that has not been approved through the firm's vendor process.
5. Verification standards, including the requirement that all citations and factual assertions in any AI-assisted work product be verified by a human before delivery to a client or filing with a tribunal.
6. Disclosure rules, including (a) when AI use must be disclosed to clients, (b) how to handle outside counsel guidelines that address AI, and (c) compliance with court standing orders.
7. Billing and fee rules, including the prohibition on billing the time AI saved at hourly rates and the rules for cost recovery.
8. Training and certification requirements before a lawyer or staff member is permitted to use any approved tool on a client matter.
9. Reporting obligations, including the obligation to report errors, near-misses, and suspected breaches to the AI Governance Committee.
10. Sanctions for violations.

4.2 Component Two: Intake and Approval

No tool should be used on client matters without going through a documented intake. The intake form should require the requesting lawyer to identify the use case, the categories of data the tool will process, the proposed vendor, and the expected user population. A standard intake will trigger (a) information security review, (b) privacy review for any tool that processes personal data,[26] (c) outside counsel guideline

review, (d) ethics review against the firm's policy, and (e) commercial review of the contract terms, including data retention, training, indemnification, and termination.

Vendor Diligence Questionnaire

- Where is data processed and stored, and in which jurisdictions?
- Is firm or client data used to train any model? If so, can that be disabled by contract?
- What is the data retention period, and is "zero data retention" available for sensitive matter work?
- What encryption standards are applied at rest and in transit?
- Has the vendor completed a SOC 2 Type II audit, an ISO 27001 certification, or equivalent?
- What are the breach notification timelines and the obligations to cooperate?
- Which underlying models are used, and how are upgrades managed?
- What logging is available to the firm, and for how long are logs retained?
- What administrative controls are available to enable or disable features by user group?
- What are the indemnification terms, especially for intellectual property claims arising from output?
- What is the exit plan, including data return and deletion?

4.3 Component Three: Training and Certification

The single most useful internal investment is mandatory training. The curriculum can be delivered in two hours and should cover (a) what generative AI is and is not, (b) hallucination and verification, (c) the firm's policy and approved tools, (d) prompt patterns that work and patterns to avoid, (e) confidentiality rules, and (f) ethics and disclosure obligations. Completion should be tracked, refreshed annually, and required before access to any approved tool is provisioned. Practice-group-specific deep dives are useful additions but do not replace the firmwide baseline.

4.4 Component Four: Monitoring and Audit

Approved enterprise tools provide administrative dashboards. Use them. The governance committee should receive at least quarterly reports on user counts by tool, prompt volume, types of use cases, security events, reported errors, and any outside counsel guideline conflicts. A targeted spot audit of a small sample of AI-assisted work product, conducted by the firm's general counsel office or a designated partner, should occur on a defined cadence.

4.5 Component Five: Incident Response

Errors will happen. The firm needs a defined path. The minimum is (a) a single internal email alias for reporting, (b) a triage protocol that includes the firm's general counsel and risk management lead, (c) a

decision tree for client notification under Rule 1.4 and applicable breach notification laws, and (d) a post-incident review that feeds back into the policy.

4.6 Component Six: The AI Governance Committee

A small, standing committee owns this. A workable composition is the firm's general counsel as chair, a litigation partner, a corporate or regulatory partner, the chief information officer or chief information security officer, the director of knowledge management or innovation, and a senior associate or counsel responsible for day-to-day administration. The committee should meet at least quarterly, approve new tools, review the policy annually, and report to firm leadership.

4.7 Mapping to External Frameworks

A firm whose clients include regulated entities should expect questions about whether the firm aligns with recognized frameworks. The two most important are the National Institute of Standards and Technology AI Risk Management Framework, including the 2024 Generative AI Profile,[27] and the European Union's AI Act, which began phased application in 2025 and 2026.[28] Several U.S. states have enacted AI-specific statutes, including the Colorado AI Act, New York City's automated employment decision tools law, and amendments to the Illinois Human Rights Act.[29] Even firms not directly subject to these laws will increasingly be asked to demonstrate that their internal controls are consistent with them.

Part V. Getting Started: A 180-Day Roadmap

This roadmap is for a firm that has not yet formally launched. It assumes ordinary firm management resources and an appetite to begin with low-risk pilots before expanding. The goal of the first one hundred and eighty days is to ship a written policy, train every lawyer, run two structured pilots in two practice groups, and present a recommendation for broader rollout. Each phase ends with a deliverable.[30]

Days 1 to 30: Foundations

1. Form the AI Governance Committee. Set a regular meeting cadence and a published agenda template.
2. Conduct a current-state inventory. Identify any AI tools already in use across the firm, including unsanctioned use of consumer chatbots. This is almost always more than leadership expects.
3. Adopt an interim policy. A short, plain-English statement that prohibits use of unapproved tools on client matters and announces that approved tools and training are coming.
4. Communicate firmwide. A note from the managing partner setting expectations and inviting interest in the upcoming pilots.
5. **Deliverable:** a written interim policy circulated to every lawyer and staff member.

Days 31 to 60: Procurement and Training Design

1. Shortlist two enterprise tools. A general-purpose enterprise assistant (for example, a vendor that offers a private deployment of a leading frontier model) and one legal-specialized platform that fits the firm's practice mix.
2. Run vendor diligence. Use the questionnaire in section 4.2.
3. Negotiate the master services agreement, the data processing addendum[31], and the security exhibit. Insist on a "no training on firm data" commitment, zero data retention for matter work, and audit rights.
4. Build the training. Two-hour mandatory module, with a short knowledge check at the end.
5. Update the engagement letter and outside counsel guideline review process to address AI use, disclosure, and cost recovery.
6. **Deliverable:** signed enterprise contracts for two pilot tools and a training curriculum ready to deploy.

Days 61 to 120: Pilots

1. Choose two practice groups for the pilots. A litigation group and a transactional group is a natural pairing because the use cases are different.

2. Define each pilot tightly. For the litigation pilot, deposition summarization and document review on a single matter. For the transactional pilot, contract abstraction across a defined population of vendor agreements.
3. Set baseline metrics. Hours per task before AI, error rate, client satisfaction, and lawyer satisfaction. Measure them again at the end.
4. Train the pilot teams first. Provision tool access only after training is complete.
5. Hold a weekly thirty-minute pilot check-in. Capture issues, prompt patterns, and unexpected uses.
6. **Deliverable:** a pilot results memo with the metrics, the lessons, and a recommendation on whether to scale each tool.

Days 121 to 180: Decision and Scale

1. Present the pilot memo to firm management. Decide whether to scale, modify, or stop each tool.
2. Adopt the final policy. Replace the interim policy with the full version described in section 4.1.
3. Roll out training firmwide. Tie tool access to training completion in the firm's identity system.
4. Stand up the dashboards. Configure the administrative reports the governance committee will use.
5. Communicate to clients where appropriate. Where outside counsel guidelines or fee arrangements call for it, send a written notice describing the firm's approach.
6. **Deliverable:** final firmwide policy, completed firmwide training, two production tools, and a publicly visible governance posture.

Common Pitfalls

Several patterns recur in firms that struggle. Treating AI as an information technology project rather than a professional practice change leaves lawyers behind. Skipping training to "save time" produces the very errors the policy was designed to prevent. Buying tools before defining use cases produces shelfware. Allowing partners to opt out of policy compliance because they are senior creates the malpractice risk. And, most commonly, deferring action because the technology is "moving too fast" cedes the field to less careful competitors and to ad-hoc adoption that is harder to govern after the fact.[32]

Appendix A. Sample Engagement Letter Language

The clause below is a starting point. Tailor for jurisdiction, practice area, and the firm's actual practice. Do not use without firm legal review.

"In providing legal services to you, the firm may use generative artificial intelligence and similar technology tools to support tasks such as document review, summarization, drafting, and research. We use these tools only through enterprise platforms that we have evaluated for security and confidentiality and that contractually prohibit the use of your information to train any underlying model. Lawyers at the firm remain responsible for all work product, and the firm verifies factual and legal assertions before delivering work to you or filing it with any tribunal. We do not separately bill you for the time saved by these tools, and any cost recovery for the tools, if applicable, is set out in the schedule attached. If you have specific instructions, restrictions, or disclosure requirements regarding the use of these tools on your matters, please let us know in writing and we will accommodate them or, where we cannot, discuss the matter with you before proceeding."

Appendix B. Sample Court Filing Certification

"The undersigned counsel certifies that, to the extent any portion of this filing was drafted with the assistance of a generative artificial intelligence tool, every quotation, citation, and factual assertion has been independently verified by a human attorney admitted to practice in this Court."

Appendix C. Recommended Reading

- ABA Standing Committee on Ethics and Professional Responsibility, Formal Opinion 512 (2024).
- State Bar of California Standing Committee on Professional Responsibility and Conduct, Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law (2023).
- Florida Bar Ethics Opinion 24-1 (2024); New York State Bar Association Task Force on Artificial Intelligence Report (2024); New Jersey Supreme Court Preliminary Guidelines on the Use of Artificial Intelligence by New Jersey Lawyers (2024); D.C. Bar Ethics Opinion 388 (2024).
- NIST, AI Risk Management Framework 1.0 and Generative AI Profile (2023, 2024).
- Stanford RegLab and HAI, Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools (2024).
- The Sedona Conference Working Group 13 on Artificial Intelligence and the Law, materials and ongoing publications, including the inaugural meeting (2025).
- Thomson Reuters Institute, Generative AI in Professional Services Reports (2024 and 2025 editions).

Appendix D. Glossary Summary

See the table at section 1.3 for the working definitions used throughout this guide.

Endnotes

The numbered references below correspond to the superscript markers in the body of the guide.

1. ABA Standing Committee on Ethics and Professional Responsibility, Formal Opinion 512, "Generative Artificial Intelligence Tools" (July 29, 2024).
2. Thomson Reuters Institute, "2024 Generative AI in Professional Services Report" (May 2024); Thomson Reuters Institute, "2025 Generative AI in Professional Services Report" (Apr. 2025). The 2024 study found that 12 percent of law firms and corporate legal teams were already using legal-specific AI, with another 43 percent planning to do so within three years; the 2025 follow-up report shows continued sharp growth.
3. See Mata v. Avianca, Inc., 678 F. Supp. 3d 443 (S.D.N.Y. 2023), where counsel were sanctioned $5,000 for filing a brief citing six fictitious judicial opinions generated by ChatGPT.
4. See ABA Formal Op. 512 (July 29, 2024); State Bar of California Standing Committee on Professional Responsibility and Conduct, "Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law" (Nov. 16, 2023); Florida Bar Ethics Op. 24-1 (Jan. 19, 2024); New York State Bar Association Task Force on Artificial Intelligence Report and Recommendations (Apr. 6, 2024); New Jersey Supreme Court, "Preliminary Guidelines on the Use of Artificial Intelligence by New Jersey Lawyers" (Jan. 25, 2024); D.C. Bar Ethics Op. 388 (Apr. 2024).
5. Model Rules of Professional Conduct r. 1.1 cmt. 8 (Am. Bar Ass'n 2012) (a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology).
6. A large language model, or LLM, is a neural network trained on enormous volumes of text that predicts the next token (a word or word fragment) given the prior context. See generally Stanford HAI, "Artificial Intelligence Index Report 2024," ch. 2.
7. Retrieval-augmented generation is described in Patrick Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," Advances in Neural Information Processing Systems 33 (2020).
8. The phenomenon is commonly called "hallucination." See Ziwei Ji et al., "Survey of Hallucination in Natural Language Generation," ACM Computing Surveys 55(12) (2023).
9. See Daniel E. Ho et al., Stanford RegLab and HAI, "Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools" (May 2024) (finding meaningful hallucination rates even in tools marketed as essentially error-free).
10. See Goldman Sachs Global Investment Research, "The Potentially Large Effects of Artificial Intelligence on Economic Growth" (Mar. 26, 2023), estimating that 44 percent of legal work tasks could be automated by current generative AI capabilities, the highest share of any occupation studied.
11. Mata v. Avianca, Inc., 678 F. Supp. 3d 443 (S.D.N.Y. 2023).

12. Park v. Kim, 91 F.4th 610 (2d Cir. 2024). The court referred the attorney to its Grievance Panel after she submitted a brief with non-existent authorities apparently produced by a generative AI tool.
13. ABA Formal Op. 512 (July 29, 2024).
14. See sources collected at endnote 4 above.
15. Model Rules of Professional Conduct r. 1.1 cmt. 8 (Am. Bar Ass'n 2012).
16. Model Rule 1.6(a), (c). See also ABA Formal Op. 477R (May 2017) on electronic communications and reasonable efforts.
17. Most enterprise vendors now offer "zero data retention" or "no training" contractual commitments. Verify them in writing rather than relying on marketing materials. Compare current enterprise terms for OpenAI, Anthropic, Google, and Microsoft against the firm's outside counsel guidelines and vendor management policies.
18. A "data processing addendum," or DPA, is a contract supplement defining the parties' respective roles and obligations with respect to personal data, typically required under privacy laws such as the GDPR (Regulation (EU) 2016/679) and the CCPA (Cal. Civ. Code section 1798.140 et seq.).
19. Model Rules 5.1 and 5.3. The Rule 5.3 comments were amended in 2012 to clarify application to nonlawyer services and tools.
20. See, e.g., Standing Order of Judge Brantley Starr, N.D. Tex., "Mandatory Certification Regarding Generative Artificial Intelligence" (May 30, 2023), requiring filers to certify that no portion of any filing was drafted by generative AI, or, if any was, that every word was checked for accuracy by a human.
21. ABA Formal Op. 512 (July 29, 2024) at sections III(B) and III(D).
22. Model Rule 1.5(a) (a lawyer shall not make an agreement for, charge, or collect an unreasonable fee).
23. See Adam Smith, Esq. (Bruce MacEwen), "The Billable Hour and [vs.?] Pricing AI-Driven Legal Services" (Sept. 2024); see also ABA Formal Op. 512 (July 29, 2024) (cautioning that, when GenAI enables a lawyer to complete a task much more quickly, charging the same flat fee may be unreasonable under Rule 1.5).
24. See generally Colorado AI Act, S.B. 24-205, codified at Colo. Rev. Stat. Title 6, Article 1, Part 17 (effective Feb. 1, 2026); New York City Local Law 144 (automated employment decision tools); Illinois HB 3773 (amending the Illinois Human Rights Act effective Jan. 1, 2026).
25. See Thomson Reuters Institute, "2024 Generative AI in Professional Services Report" (May 2024) (reporting that only about 10 percent of law firms had implemented a written generative AI policy, with 8 percent treating AI under an existing technology policy and roughly three-quarters reporting no policy); see also ILTA, "2024 Technology Survey" (Sept. 2024) (reporting that 34 percent of firms had no official AI policy and 26 percent had a policy under development).

26. U.S. state privacy laws now in effect include those of California, Virginia, Colorado, Connecticut, Utah, Texas, Oregon, Montana, Iowa, Tennessee, Indiana, Florida, New Jersey, Delaware, and others. The European Union's General Data Protection Regulation continues to apply to processing of personal data of EU data subjects.

27. National Institute of Standards and Technology, AI Risk Management Framework 1.0 (NIST AI 100-1) (Jan. 2023), and Generative AI Profile (NIST AI 600-1) (July 2024).

28. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (the "AI Act"), O.J. L, 2024/1689.

29. See endnote 24 above on the Colorado AI Act, NYC Local Law 144, and Illinois HB 3773.

30. See Lucy Bassli, "The Simple Guide to Legal Innovation" (ABA 2020), describing change-management patterns observed in legal departments and law firms.

31. See endnote 18 above on DPAs.

32. See LexisNexis, "Generative AI and the Future of the Legal Profession" survey reports (2023, 2024); LexisNexis, "Investing in Legal Innovation Survey" (2024), reporting growing client expectation of AI-driven efficiency and a sharp increase in lawyer use of generative AI tools year over year.