Teaching AI to Lawyers

Colin S. Levy
2026

This document is for informational and educational purposes only and does not constitute legal advice.

About the Author

Colin S. Levy is a legal technology advocate, writer, and advisor who works at the intersection of law, technology, and business. With experience spanning in-house legal roles, legal technology companies, and legal operations, he brings a practical perspective to how legal teams can adopt and govern emerging technologies responsibly.

Colin writes and speaks extensively on legal innovation, artificial intelligence in legal practice, and the evolving role of legal professionals in a technology-driven landscape. His work focuses on helping legal teams move beyond the hype cycle to make sound, informed decisions about the tools they use and the workflows they build.

He is the author of *The Legal Tech Ecosystem* and editor of the *Handbook of Legal Tech*, and a regular contributor to publications covering legal technology and operations. He advises organizations on responsible AI adoption, legal operations strategy, and the practical governance frameworks that make innovation sustainable.

This guide is part of a series on AI and law that includes *Addressing AI Concerns*, *Managing AI Hallucinations*, *AI for Lawyers*, *AI for Legal Teams*, *AI Agents Data Handling and Cybersecurity Guide*, *AI in the Courtroom*, *Contracting with AI Vendors*, *Human Judgment and AI in Legal Practice*, and the *AI Implementation Playbook for Legal Teams*.

Table of Contents

Part One

Why This Guide Exists

The Teaching Gap

Every law firm, in-house department, and law school is now in the business of teaching artificial intelligence to lawyers, whether they admit it or not. The question is no longer whether to teach. It is whether the teaching will be deliberate or accidental. Two years of generative AI in the profession has produced a generation of lawyers who use these tools daily, often without understanding what the tools do, what they cannot do, and what professional duties travel with their use.

The consequences of accidental teaching are now visible in the published cases. Lawyers who never had a structured introduction to large language models filed fabricated citations. Associates who learned prompting from a colleague's screen pasted privileged material into consumer chatbots. Partners who skimmed a vendor demo signed contracts that quietly authorized training on client data. None of this is malice. It is the predictable outcome when a powerful technology spreads faster than the teaching that should accompany it.

What This Guide Is, and Is Not

This guide is a framework for teaching AI to lawyers. It is written for the people who carry that responsibility: law professors building a course, partners running a CLE, learning and development leaders inside firms and in-house teams, and bar associations mapping out continuing education. It draws on the published guidance from the American Bar Association, several state bars, and regulators in the United Kingdom, Canada, Australia, Singapore, and the European Union.[1] It draws on the body of case law in which courts have addressed lawyer use of AI directly. It draws on the practical experience of programs that have been running long enough to have something to show for it.

It is not a textbook. It is not a vendor comparison. It is not a survey of every possible AI tool. The technical landscape changes too quickly for a guide to be useful as a snapshot. What does not change is the underlying question: what does a competent lawyer need to know, and how should we teach them to know it.

Who This Guide Is For

The audience splits into two overlapping groups. The first is law professors, clinical instructors, and continuing legal education designers who teach lawyers and law students for a living. The second is the working lawyer who has been asked, often without warning, to lead training inside a firm or department. The needs of those two groups overlap more than they differ. Both need a defensible curriculum. Both need to know what to teach, in what order, and to what depth. Both need to evaluate whether learners actually walk away with useful skill.

How the Guide Is Organized

Part Two sets out what lawyers actually need to know. Part Three discusses how lawyers actually learn, which is not always how technologists assume they learn. Parts Four through Six work through the three substantive content areas: foundations, practical skills, and ethics and professional responsibility. Parts Seven and Eight address two distinct teaching contexts, the law school course and the practitioner training, each of which calls for different design choices. Part Nine addresses assessment, materials, and a maturity model that any program can use to find its current position. Part Ten closes with what comes next.

What Has Changed Since 2023

The teaching agenda this guide describes did not exist in its current form three years ago. The trajectory matters because curriculum decisions are easier to defend when they are anchored to the specific developments that produced them. ChatGPT became broadly available in late November 2022. By June 2023 a federal court had imposed sanctions on counsel in *Mata v. Avianca* for filing fabricated citations. The Florida Bar issued Advisory Opinion 24-1 in January 2024. The California State Bar issued its Practical Guidance in November 2023. The ABA issued Formal Opinion 512 in July 2024. The European Union's Artificial Intelligence Act entered into force in August 2024. The Supreme Court of New South Wales issued Practice Note SC Gen 23 in late 2024, effective February 2025. The Stanford HAI study by Magesh and colleagues, published in 2025, gave the field its first rigorous empirical measure of hallucination rates in legal research tools.

Two years of that pace produces a curriculum problem that does not exist in subjects that move more slowly. A constitutional law syllabus does not require quarterly revision. An AI syllabus does. The pedagogical implication is that program design has to be modular, with stable foundational elements and a clearly demarcated layer of current developments that gets refreshed every term. A program that does not draw that line ages badly because the foundational material gets contaminated by the moving parts. A program that does draw the line ages predictably because only the moving parts move.

> **Key Principle:**
>
> Lawyers are already learning AI. The question is whether anyone is teaching them, and whether the teaching is grounded in the duties they already owe. The lawyers who get hurt are not the ones who studied carefully and proceeded with caution. They are the ones who picked up the tool from a colleague and never looked under the hood.

Part Two

What Lawyers Actually Need to Know

The Four Layers of AI Literacy

Effective AI training for lawyers covers four layers, in roughly this order. Each layer rests on the one below it. Skipping the foundation is the most common pedagogical mistake, and the one that produces the most expensive errors.

The first layer is conceptual. What is a large language model. How does it generate output. Why does it produce confident text that turns out to be wrong. What does retrieval-augmented generation do, and what does it not do. The second layer is practical. How do you write a prompt that produces useful work. How do you verify what comes back. How do you know when a tool is not appropriate for the task in front of you. The third layer is ethical and regulatory. What do the rules of professional conduct require. What does the ABA's Formal Opinion 512 actually say. What court orders and bar opinions apply where you practice. The fourth layer is governance. How does an organization translate individual competence into a defensible system, with policy, an approved-tool list, training records, and incident reporting.[2]

The "Enough to Ask the Right Questions" Standard

The mistake some programs make is teaching lawyers to be junior machine learning engineers. The mistake others make is teaching nothing technical at all and pretending the tools are magic. Both miss the standard the profession actually needs. The standard is functional. A lawyer should know enough to ask the right questions of a vendor, recognize when an output is suspicious, push back on an associate who cannot explain what tool they used, and read a data processing addendum without glazing over.

That standard is closer to what a litigator already brings to a damages model from an opposing expert. The litigator does not need to derive the regression. The litigator needs to know that a regression has assumptions, that the assumptions are testable, and that the right questions will surface them. AI literacy is the same kind of skill, applied to a different artifact.

What Most Curricula Get Wrong

The most common curriculum failure is to lead with the tools. A two-hour CLE that walks through Harvey, CoCounsel, Lexis+ AI, Westlaw Precision AI, and one or two general-purpose models will teach learners what to click and nothing about what the click is doing. Six months later the tools will have changed and the learners will be lost. A curriculum built on tools is a curriculum with a short half-life.

The second common failure is to lead with horror stories. *Mata v. Avianca* is a useful case, but a curriculum that opens with sanctions and never gets past them produces lawyers who fear the technology rather than understand it. Fear is not literacy. Fear without understanding produces avoidance, which produces a different professional risk: the lawyer who falls behind the standard of care because the lawyer never learned what was possible.[3]

The third common failure is to skip the ethics layer entirely. Some technical trainers treat professional responsibility as a footnote, a slide near the end. The lawyers in the room are not learning AI in a vacuum. They are learning AI as a tool they will use under existing fiduciary duties, with existing supervisory obligations, with existing duties of confidentiality and candor. The ethics layer is not the dessert. It is the table the rest sits on.

A Working Definition of AI Competence

For purposes of curriculum design, treat AI competence as a lawyer's reliable ability to do four things. Identify whether a given task is appropriate for AI assistance, and at what risk tier. Select an approved tool that matches the task. Use the tool with prompts and verification protocols that meet professional duties. Document the use sufficiently to defend it if the work product is later questioned. A lawyer who can do all four, repeatedly, across different practice areas, is competent in this sense. A lawyer who can do only one or two of the four is exposed.

Competency Benchmarks by Role

The four-part competency above is the floor. Different roles require different ceilings. A junior associate working under close supervision needs reliable competence in identification and verification. A partner who supervises associates and signs the brief needs all four parts and the additional capacity to evaluate whether the team's work meets the standard. An in-house counsel acting as the company's AI buyer needs the four parts plus enough familiarity with vendor diligence and data processing terms to negotiate from a position of understanding. A law professor designing assessments needs the four parts plus the pedagogical capacity to identify what students actually retained.

The benchmarks below are pragmatic, not aspirational. They describe what a working program should expect each role to demonstrate at the end of an annual training cycle. Programs that publish their benchmarks tend to produce learners who meet them. Programs that leave the benchmarks implicit produce learners who guess at what they were supposed to learn.

- **Junior associate.** Can articulate why a model hallucinates, can run a citation-by-citation verification pass on AI-assisted work product, can identify when a task does not belong in AI at all, and can describe the firm's policy in plain language.
- **Mid-level associate or counsel.** All of the above, plus the ability to choose among approved tools for a given task and to supervise a junior's AI-assisted work as if the junior had drafted it without AI.
- **Partner or supervisory lawyer.** All of the above, plus the ability to evaluate the firm's policy and approved-tool list against current guidance, identify when an incident requires reporting, and lead a meaningful conversation with a client about AI use.
- **In-house counsel.** All of the above, plus the ability to read a vendor data processing addendum, identify training and retention terms that need negotiation, and articulate the company's AI risk posture to non-legal stakeholders.

- **Government and public sector lawyer.** All of the above, plus competence with FOIA and transparency implications, due process when AI affects adverse decisions, and any sector-specific procurement and disclosure rules.
- **Law professor or instructor.** All of the competence above relevant to the audience, plus the pedagogical capacity to design assessments that distinguish surface familiarity from durable skill.

Practical Takeaway:

Teach in this order: foundations, practical skills, ethics and professional responsibility, governance. Skipping the foundation produces the most expensive errors. Leading with tools produces lessons with a six-month half-life.

Concern Resolved:

The right thing to teach is not the latest tool. It is the four-layer base that makes any tool legible to the lawyer in front of it. A curriculum built on the four layers ages well. A curriculum built on this quarter's vendor demo does not.

Part Three

How Lawyers Actually Learn

Adult Learners with Existing Frameworks

Lawyers are adult learners with deeply formed mental models of how knowledge work happens. Many came of age before generative AI existed in any practical form. Their working theory of research, writing, and analysis was built around physical books, then around Westlaw and Lexis, then around document automation, then around the internet. AI is the fourth or fifth wave they have absorbed. The teaching that works treats the prior frameworks as assets to build on, not obstacles to push past.

This means leading with analogies the audience already trusts. A junior associate is the most common one. The reasons we verify a junior's first draft, supervise the junior's research, assign the junior to tasks that match the junior's experience, and never let the junior sign a brief unread are the same reasons we verify, supervise, and oversee an AI tool. The analogy is not perfect. Junior associates learn from feedback. Models do not, at least not in the same way. But the supervisory frame travels well, and ABA Formal Opinion 512 explicitly adopts it by treating generative AI tools as nonlawyer assistants for purposes of Model Rules 5.1 and 5.3.[4]

Skepticism Is a Pedagogical Asset

The most common complaint about lawyer audiences from technical trainers is that they are too skeptical. The complaint is misplaced. Skepticism is the professional reflex these audiences have built over decades of cross-examining experts, scrutinizing contracts, and reading regulations against their grain. The teaching that fails treats skepticism as resistance to be overcome. The teaching that works treats skepticism as the lever that gets the lawyer engaged with the material. Show a hallucinated citation and ask the room how they would have caught it. Show a data processing addendum with a quiet training-rights clause and ask what the negotiation looks like. The audience comes alive when the material rewards the skill they brought in the door.

Audience Segmentation

One curriculum does not fit every lawyer audience. The needs and constraints diverge sharply across five recognizable segments. Each segment shifts what to teach, in what depth, and what to leave for another day.

Segment	What the Curriculum Has to Add or Adjust
Solo and small firm	Time-compressed and opinionated. Prescribe one or two tools, a working contractual posture, and a pre-filing checklist the lawyer can adopt the same week.
BigLaw associates and partners	Tools and policy already exist at the firm level. Focus on the firm's approved configuration, the supervisory architecture above and below the lawyer, and the firm's verification routine.
In-house counsel	Add commercial AI risk: vendor diligence, data processing terms, indemnities, EU AI Act exposure, and the lawyer's role inside enterprise governance.
Government and public sector	Procurement constraints, FOIA implications, due process when AI affects adverse decisions, and statutory transparency obligations the private sector does not face.
The bench	Not typically AI users for opinions, but increasingly recipients of AI-assisted submissions and authors of standing orders. Curriculum focuses on what to ask of counsel and what to require in practice notes.[5]

The Demonstration Imperative

Adult learners absorb a demonstrated workflow far more durably than they absorb a described one. The implication for curriculum design is that the centerpiece of any module should be something the instructor does in front of the audience, not something the instructor talks about. A live verification pass on a hallucinated citation, performed in real time, with the audience watching the search return no results, teaches more in three minutes than thirty minutes of slides on the importance of verification. A live walk through a vendor's data processing addendum, with the instructor reading aloud the training-rights clause and pausing to ask the audience what it means, teaches more than a separate module on contract review.

Demonstrations work for the same reason cross-examination works. The audience watches the artifact behave, sees what the artifact does, and arrives at conclusions the audience can defend because the audience drew them. The instructor who tries to substitute description for demonstration, often to save time, ends up requiring more time in the long run because the durability of the learning suffers.

Handling Pushback

The lawyer audience is articulate. Pushback will happen. Three patterns recur, and each repays preparation. The first is the senior practitioner who insists the technology is overhyped and that the existing tools are sufficient. This pushback is often partially correct, and the instructor who concedes the partial correctness while distinguishing the part that is not correct gets further than the instructor who treats the pushback as resistance. The second is the early adopter who has been using consumer AI for a year and resents being told to slow down. This pushback is often partially correct as well, and the answer is to acknowledge the head start while surfacing the specific risks the adopter has not yet encountered. The third is the technically skeptical practitioner who wants more technical detail than the room can absorb. This pushback rewards a candid reading list at the end of the session rather than an attempt to satisfy it in the moment.

Across all three patterns, the principle is the same: the audience's pushback is information. It tells the instructor where the curriculum has not yet landed for that learner. Treating pushback as feedback, rather than as obstacle, produces a better next session.

Time, Attention, and Return on Investment

Lawyer audiences will give a curriculum the attention they think it deserves. The attention budget is small at the start and grows only as the material proves itself. Open every module with a concrete payoff. A two-hour CLE that ends with the participant able to write three working prompts, run a verification pass on a generated brief, and list the contract terms to ask their next vendor about has earned its time. A two-hour CLE that ends with the participant able to recite a definition of "transformer architecture" has not.

> **Teaching Note:**
>
> Do not start by trying to make the audience excited. Start by respecting the skepticism the audience already brought into the room, and use it. The lawyer who pushes back on the first slide is the lawyer who will use the material on Monday.

> **Concern Resolved:**
>
> Lawyers do not learn AI the same way computer science students do. They learn it as adult professionals with prior frameworks, real time pressure, and a working sense of risk. The curriculum that respects all three meets the audience where it actually is.

Part Four

Teaching the Foundations

What a Large Language Model Actually Does

The single most useful sentence in any AI training for lawyers is some version of this: a large language model produces text by predicting the most probable next token, given the text that came before. That sentence is technically rough but functionally correct, and it answers more questions than a chapter of architecture diagrams. It explains why the output sounds confident even when it is wrong. It explains why two prompts that look almost the same can produce very different answers. It explains why the model cannot tell you what it does not know, because not-knowing is not a state the model has access to. It explains why citations get fabricated: a citation is just a text pattern, and the model is good at producing patterns.[6]

Once that sentence is in the room, almost everything else fits around it. Hallucination is no longer mysterious. Verification is no longer optional. Prompting becomes a discipline rather than a magic incantation, because the prompt is the prior context the model is conditioning on.

Foundation Models, Wrappers, and Retrieval

The next foundational concept is the stack. A foundation model, such as the GPT, Claude, Gemini, or Llama families, is a large general-purpose model trained at enormous expense by a small number of companies. The legal-specific tool the lawyer is using is rarely a model the vendor built from scratch. It is a wrapper around a foundation model, often supplemented with retrieval over a curated corpus of cases, statutes, or firm documents. That is what retrieval-augmented generation, or RAG, refers to.

Teaching the stack matters because it teaches where the risks live. If the foundation model is hosted by a third party in another country, the data goes there. If the wrapper has its own retention policies on top of the foundation model's, both apply. If the corpus the tool retrieves over is incomplete or stale, the answers will be incomplete or stale, and no amount of clever prompting fixes that.[7]

Why Models Hallucinate

The Stanford HAI study by Magesh and colleagues, published in the *Journal of Empirical Legal Studies*, remains the cleanest empirical reference point for any training. Across federal case-law queries, the study found hallucination rates of approximately seventeen percent for Lexis+ AI, thirty-three percent for Westlaw's AI-Assisted Research, and forty-three percent for GPT-4. Purpose-built legal tools using RAG hallucinate substantially less than general-purpose models, but they still hallucinate.[8]

That is the empirical headline. The mechanism behind it is the one already introduced. The model is generating text that fits the statistical pattern of legal writing. A fabricated citation is structurally indistinguishable to the model from a real one. The retrieval layer narrows the gap by grounding the output in retrieved sources, but it does not close the gap, because the model can still misread the retrieved text, conflate two cases, or confabulate a quotation that does not appear in the source.

Context Windows, Training Cutoffs, and System Prompts

Three more concepts repay the time spent introducing them. The context window is the amount of text the model can attend to at once. Material outside the window is invisible to the model regardless of how important it is. The training cutoff is the date past which the model has not seen new data, and it explains why the model does not know about last week's opinion or this morning's regulation. The system prompt is the instruction layer the vendor has wrapped around the model, which the user typically does not see and which shapes everything the model does. Lawyers who understand these three concepts ask better questions about the tools they are evaluating.

Common Foundational Misconceptions

Lawyer audiences walk into a foundations module carrying recognizable misconceptions that the curriculum has to surface and dislodge. Four recur most often, and each is worth a few minutes of explicit attention.

The first misconception is that the model has access to the internet in real time. Most foundation models do not, except in specific configurations the user has to invoke. The model knows what it was trained on, plus whatever is in the prompt, plus whatever a retrieval system has placed in front of it. Surface this early, because the lawyer who assumes live access will trust the output for current questions the model cannot actually answer. The second misconception is that the model remembers prior conversations across sessions. Most consumer interfaces do not retain that state by default, and the ones that do should be understood explicitly because they create their own confidentiality issues. The third misconception is that "fine-tuning" the model on the firm's documents would make the model produce only firm-correct answers. Fine-tuning shifts behavior at the margins. It does not bound the model to the training corpus. The model can still hallucinate, and hallucinations from a fine-tuned model carry the additional risk of sounding more authoritative because they sound more like the firm. The fourth misconception is that the model has reasoning capacities equivalent to a human professional. The model has fluent reasoning-shaped output. Whether that constitutes reasoning is an open question, and the safer working assumption for a legal practice is that the output is a draft requiring evaluation, not a conclusion requiring deference.

Five Metaphors That Work, and One That Does Not

Pedagogically, metaphors do most of the heavy lifting in foundational teaching. Five that hold up under scrutiny:

- **The very fluent intern.** Smart, articulate, often plausible, sometimes wrong, never embarrassed when caught. Verify every cite.
- **The autocomplete that reads your sentence and finishes it.** Useful for highlighting that the model is producing the next likely thing, not retrieving a fact.
- **The very large library with no librarian.** The information may be in there. The system has no reliable way to tell you whether it is, or where.

- **The translator who makes things sound natural.** Useful for explaining why tone and structure are often excellent even when the substance is wrong.
- **The seasoned actor reading a script you wrote.** Useful for explaining why prompt engineering matters: the script the model performs depends on the script you handed it.

The one metaphor to avoid is "the digital brain." Brains do many things models do not, and models do many things brains do not, and the metaphor produces wrong intuitions in both directions. It also makes lawyers either over-trust or over-fear the tool depending on which direction their imagination runs.

> **Teaching Note:**
>
> If the audience leaves understanding only one thing, make it the next-token sentence. Everything else in the curriculum is easier to teach to a room that has internalized it.

Teaching Agents and Agentic Systems

The foundations module has to introduce one further distinction that is becoming consequential as tools mature. A traditional generative AI tool waits for a prompt, produces an output, and stops. An agentic system takes goals or tasks and executes a sequence of steps autonomously, sometimes including writing files, sending emails, calling external APIs, or invoking other tools. The pedagogical importance of the distinction is that supervision is no longer a step the lawyer chooses to take after the output. Supervision must be designed into the workflow at the moment work product becomes consequential, because there is otherwise no human in the loop at the consequential moment.

For most lawyer audiences a brief introduction to the agentic concept is enough at the foundations stage. The audience needs to know that "AI" no longer refers to a single category of tool, that the supervisory question depends on which category they are deploying, and that any tool that takes consequential action without a human checkpoint is a different supervisory problem than a tool that merely produces a draft for review. Programs running ahead of the average curve will return to agentic systems in a separate module or follow-up session, because the supervisory architecture they require is not a small addition to existing controls. It is a different architecture.

> **Concern Resolved:**
>
> Foundations are teachable in roughly an hour to any lawyer audience that respects the standard of care. The lawyers who understand why the model produces what it produces stop being surprised by it. Surprise is the first place errors live.

Part Five

Teaching the Practical Skills

Prompting as a Skill, Not a Hack

Most published prompting advice is folklore. Some of it is useful, some of it is performative, and some of it has the shelf life of last season's vendor demo. Teaching prompting well means teaching the underlying logic rather than a list of tricks. The logic comes back to the foundational sentence. The model conditions on what came before. So a good prompt does three things. It gives the model a clear role and audience. It supplies the relevant material the model would otherwise have to fabricate. And it specifies the form of the output in enough detail that the answer is usable when it arrives.

Beyond that, four habits do most of the practical work. Provide examples of the output you want. Ask the model to show its reasoning where reasoning matters. Break long tasks into smaller ones. And iterate, treating the first response as a draft and refining the prompt rather than trying to argue with the output. Lawyers familiar with how a partner gives an assignment to a junior pick up these habits quickly, because the parallels are exact.

Common Prompt Patterns in Legal Practice

A few prompt patterns recur often enough in legal work to be worth teaching as patterns rather than as ad hoc constructions. Each is a structure the lawyer can adapt to a wide range of tasks. Each can be illustrated in a single demonstration and then handed to the audience for adaptation in a lab.

The role-context-task-format pattern is the workhorse. The prompt opens by stating what role the model should play (a defense litigator preparing for a motion hearing, a transactional lawyer drafting a representation), supplies the relevant context (the fact pattern, the jurisdiction, the procedural posture), states the task in clear terms, and specifies the form of the output (a memorandum, a numbered outline, an annotated checklist). The few-shot pattern adds one or two worked examples of the desired output before the actual request, on the principle that the model performs the script better when it has seen the script. The chain-of-thought pattern asks the model to walk through its reasoning before stating a conclusion, which is useful in analytical tasks where the lawyer needs to evaluate the steps rather than only the answer. The constrained-output pattern instructs the model to respond only in a defined format, which reduces variability for tasks that will be parsed by another tool or compared across iterations.

None of these patterns is novel. All of them are teachable in fifteen minutes. The reason to teach them as patterns is that the audience can then adapt the pattern to a thousand tasks rather than memorize a thousand prompts.

A Worked Iteration: Refining a Single Prompt

One of the most useful demonstrations a curriculum can offer is a live iterative refinement. Start with a deliberately weak prompt, run it, and walk the audience through three rounds of refinement with the

audience suggesting the next change. The pedagogical value is in watching how small adjustments produce large differences in output.

A typical sequence on a research task might look like this. Version one is a single sentence: "Summarize the law on the duty to mitigate damages in breach of contract cases." The output is generic and probably partly wrong on jurisdiction. Version two adds the role and the jurisdiction: "You are a litigator practicing in California. Summarize the California law on the duty to mitigate damages in breach of contract cases, citing only California cases." The output improves but may still cite cases the lawyer cannot verify. Version three constrains the output and the source: "Using only the cases I provide below, summarize how California courts have applied the duty to mitigate damages. For each case, state the year, the holding in one sentence, and the page citation. If a case I have provided is irrelevant, say so and skip it." The output is now usable, because the model is grounded in the lawyer's curated source set rather than in its own statistical guesses.

The teaching point of the worked iteration is not the specific prompt. It is the discipline of moving from underspecified to specified, from open-ended to grounded, and from trusting the model's recall to providing the model's source material. Every audience that watches the iteration leaves with a working pattern they can apply that day.

The Verification Habit

The single most important practical skill is verification. A curriculum that does not produce reliable verification habits has not done its job, regardless of how well it teaches anything else. The verification standard ABA Formal Opinion 512 sets is task-specific, but the operating principle is consistent across guidance: every citation, every quotation, every material factual claim in any AI-assisted work product must be confirmed against an authoritative source by a lawyer before reliance.[9]

Teach verification as a routine, not as a virtue. Routines survive the busy day. Virtues do not. The routine that works in practice is short enough to fit on a sticky note: pull every cite, compare every quote, confirm every key fact, and read the whole document one more time before signing. The lawyers who do this will not be the ones in the next published opinion.

Choosing the Right Tool for the Task

Not every task belongs in an AI tool. The triage framework from the companion guide on hallucinations works well as a teaching artifact: classify use cases into Red, Yellow, and Green tiers based on the consequence of an error and the verifiability of the output. Red tier work, which would cause material harm if the output were wrong and which cannot be easily verified, should not be delegated to AI at all. Yellow tier work belongs in AI only with rigorous verification. Green tier work, where the consequence of error is low and verification is straightforward, is where AI delivers most of its real value. Teach the framework and let the learners place their own daily tasks on the grid. The exercise is more useful than any list of approved use cases the trainer could supply.

A Hands-On Lab Plan

A lecture about AI does not produce competence. A hands-on lab does. The lab does not need expensive infrastructure. A working laptop, a single approved AI tool, three or four short legal scenarios, and a structured worksheet are enough. The structure of a productive lab follows a consistent shape:

- Begin with a five-minute concept review of the foundational sentence and the verification routine.
- Hand the learner a fact pattern and a specific work product to produce, such as a research memo on a narrow question, a deposition outline, or a draft cease-and-desist letter.
- Have the learner draft a prompt, run it, and capture the output.
- Have the learner verify each citation, quotation, and factual claim against authoritative sources, marking on the worksheet which checks passed and which failed.
- Debrief in small groups. The hallucinations almost always surface, and the conversation about them does more teaching than any lecture would.

Sample Lab: Drafting a Motion to Compel

A useful first lab uses a discovery dispute the learners can grasp without specialist knowledge. Hand each participant a one-page fact pattern: a defendant has produced redacted documents in response to a request that the plaintiff believes calls for unredacted production, the meet-and-confer has failed, and counsel needs to draft a motion to compel under the relevant rule. Ask the learner to use a single approved AI tool to produce a first draft. Then ask the learner to run a verification pass.

The verification pass surfaces predictable issues. The model will sometimes cite cases that do not exist, or that exist but stand for different propositions. It will sometimes quote rule language that does not match the current rule. It will sometimes produce arguments that read well but rest on inapplicable jurisdictional precedent. Each error is a teaching moment. The learner who sees it once, on a low-stakes practice exercise, is dramatically more likely to catch it on a real filing.

Sample Verification Worksheet

For any AI-assisted work product completed during this lab, the participant initials each step and retains the worksheet.

☐ Every case, statute, and rule citation independently retrieved from Westlaw, Lexis, or the official source.

☐ Every quotation compared to source text and confirmed verbatim.

☐ Every factual assertion confirmed against the underlying record or a primary source.

☐ The whole work product read in full by the participant before submission.

☐ Verification time, errors found, and tool used recorded for the post-lab debrief.

Participant: _____ Date: _____

Common Lab Failure Modes and How to Use Them

Five failure modes will surface in almost every hands-on session, and each is a structured teaching opportunity rather than an embarrassment to be smoothed over.

The first is the participant who copies the entire fact pattern, including names and details, into a consumer-tier AI tool with no enterprise terms. This is the moment to surface the confidentiality issue, the Samsung incident, and the difference between consumer and enterprise tiers of the same product family. The second is the participant who accepts the first output and never iterates. This is the moment to teach iteration as a discipline. The third is the participant who skips the verification pass because the output "looks good." The fabricated citations almost always come from this group, and the debrief writes itself. The fourth is the participant who over-prompts, packing the prompt with so much instruction that the model collapses under conflicting demands. The fifth is the participant who insists the model is wrong when it is right, because the answer does not match the participant's prior belief. This last one is rarer but the most useful, because it teaches the symmetric error: AI tools are not always wrong, and the lawyer who treats them as a foil rather than a collaborator is making a different kind of mistake.

What the Lab Should Not Do

A few things are tempting in a lab and should be resisted. Do not use real client matters, even sanitized. The hypothetical-but-realistic scenario produces the same learning without the risk. Do not turn the lab into a tool comparison. The point is the skill, not the brand. Do not let the lab become a coding exercise. Lawyers need to know how to use the tool, not how to build it. Do not skip the structured debrief. The debrief is where the lab does its actual teaching. A lab that ends when the timer runs out leaves most of the learning on the table.

> **Concern Resolved:**
>
> Practical skills are taught with practice, not with slides. A two-hour module that includes a structured lab with a verification worksheet teaches more than a full day of lecture. Build the lab first, and the lecture content will sort itself around it.

Part Six

Teaching Ethics and Professional Responsibility

Anchoring on ABA Formal Opinion 512

The ABA Standing Committee on Ethics and Professional Responsibility issued Formal Opinion 512, "Generative Artificial Intelligence Tools," on July 29, 2024. It is the first comprehensive ABA opinion on lawyer use of generative AI, and it is the natural anchor for any ethics module. The opinion addresses six existing duties: competence under Model Rule 1.1, confidentiality under Rule 1.6, communication under Rule 1.4, candor under Rules 3.1 and 3.3, supervision under Rules 5.1 and 5.3, and reasonable fees under Rule 1.5.[10]

The pedagogical point is that the opinion does not invent new duties. It applies existing duties to a new tool. That framing does important work in a classroom or CLE. It tells the audience that the obligations they already understand are the obligations that apply here. The mental model the audience already carries is mostly correct. What changes is the mechanism through which the duties show up in daily practice.

The Six Duties Lawyers Already Owe

A teachable summary of how each of the six duties applies in the AI context fits comfortably on one classroom slide and rewards extended discussion.

- **Competence (Rule 1.1).** The duty to keep abreast of the benefits and risks of relevant technology, recognized in Comment 8 to Rule 1.1, was adopted by the ABA in 2012 and has been adopted in some form in a substantial majority of U.S. states. It now plainly extends to a reasonable understanding of any AI tool the lawyer uses.[11]
- **Confidentiality (Rule 1.6).** Rule 1.6(c) requires reasonable efforts to prevent inadvertent or unauthorized disclosure. Submitting confidential information to a tool that trains on inputs, retains prompts, or routes data through unknown sub-processors is the failure mode the rule was written to prevent.
- **Communication (Rule 1.4).** Where AI use is material to the representation, including the use of a tool to generate output the client will rely on, the duty of communication may require disclosure even where confidential information is not involved.
- **Candor (Rules 3.1 and 3.3).** A signed filing represents to the tribunal that the contentions in it are warranted. AI-generated content that the lawyer has not verified does not satisfy that representation.
- **Supervision (Rules 5.1 and 5.3).** Formal Opinion 512 treats generative AI tools as nonlawyer assistants for supervisory purposes. Supervisory lawyers owe the same kind of oversight to AI-assisted work that they owe to paralegal-assisted work.

- **Reasonable fees (Rule 1.5).** A lawyer cannot bill the client for time saved by using AI, cannot bill for time spent learning a generally applicable AI tool, and may charge for the tool's actual cost only with adequate disclosure and consent.

Sanctions Cases as Teaching Tools

The published sanctions cases are now a curriculum in themselves. Researcher Damien Charlotin maintains a public database of legal decisions in which courts have addressed lawyer reliance on AI-generated content. As of early 2026 the database contains more than 1,300 documented decisions across multiple jurisdictions, with new decisions added at a steady pace.[12] A focused module can be built around four or five anchor cases.

Mata v. Avianca, Inc., 678 F. Supp. 3d 443 (S.D.N.Y. 2023), remains the leading case. Plaintiff's counsel filed an affirmation citing six fabricated decisions generated by ChatGPT, then provided fabricated opinion texts when challenged. Judge Castel found subjective bad faith and imposed a $5,000 sanction.[13] *Park v. Kim*, 91 F.4th 610 (2d Cir. 2024), is the natural follow-up: the Second Circuit referred counsel to its Grievance Panel for citing a nonexistent case generated by ChatGPT and making no inquiry into its validity.[14] *Noland v. Land of the Free, L.P.*, decided in September 2025 by California's Court of Appeal, was the first published California Court of Appeal opinion to address AI-fabricated authority. The court imposed a $10,000 sanction and reported counsel to the State Bar after finding briefs "replete" with fabrications generated using multiple consumer AI tools.[15] *Zhang v. Chen*, 2024 BCSC 285, demonstrates that the same pattern recurs in common-law jurisdictions outside the United States. The Supreme Court of British Columbia ordered counsel to personally compensate opposing counsel for time spent unwinding fabricated authority, and the Law Society of British Columbia opened an investigation.[16]

A Case-Based Ethics Module

The most effective ethics module is built around a small set of cases that the audience reads in advance and discusses in class. Three or four cases, paired with the relevant rule provisions and Formal Opinion 512, fill a productive sixty to ninety minutes. The pedagogical aim is not to make the audience afraid. It is to make the audience capable of asking, on their own next filing, what the lawyers in the published cases failed to ask. That question is the same in every case: what would have surfaced this error if anyone had run a verification pass.

Sample Discussion Questions

For an ethics module built around *Mata*, *Park*, *Noland*, and *Zhang*:

- At what step in the lawyer's workflow could each error have been caught?
- What firm-level controls would have made that step routine rather than discretionary?
- Which of the six ABA Formal Opinion 512 duties is most directly implicated by each case, and why?
- If you were sanctioned today and had to defend your AI use, what documentation would you produce?

Building the Confidentiality Module

Confidentiality is the duty most likely to be misunderstood by AI users, and a brief subsection inside the ethics module is rarely sufficient to dislodge the misunderstandings. A working confidentiality module is forty-five to sixty minutes of its own, anchored to Model Rule 1.6(c) and the cluster of state bar opinions addressing cloud and third-party processing, and built around three teaching artifacts.

The first artifact is the Samsung incident from March 2023, in which engineers exposed confidential source code by pasting it into a consumer AI tool. The incident is useful precisely because the engineers were not malicious. They were trying to do their jobs faster, and the data left the building anyway. The pedagogical question to put to the audience is direct: what would have prevented this in your firm or department, and is that thing actually in place. The second artifact is a redacted vendor data processing addendum. Hand the audience the addendum and ask them to find the training-rights clause, the retention period, the sub-processor list, and the cross-border transfer mechanism. Lawyers who have never read one of these documents are surprised by how much is in them and how easily the consequential terms hide. The third artifact is a worked privilege analysis. Walk the audience through how the existing privilege analysis on cloud and outside vendors applies to AI processing, where it breaks down, and what the conservative position requires. The Florida Bar's Advisory Opinion 24-1, with its recommendation of informed consent before submitting confidential information to a third-party generative AI tool, is the natural anchor.

Building the Supervision Module

The supervision module is shorter and more structural. The goal is not to make every lawyer a compliance officer. The goal is to make every supervising lawyer capable of recognizing whether their team's AI use sits inside the supervisory architecture the rules contemplate. The module covers Model Rules 5.1 and 5.3 in the AI context, the framing in ABA Formal Opinion 512 that treats AI tools as nonlawyer assistants, and the practical question of what the supervising lawyer should be able to demonstrate if a regulator or insurer asked. The artifact at the center of the module is the firm's own AI use policy, walked through provision by provision, with the audience identifying which provisions implicate supervision and what the supervisor's role is for each.

Common Pitfalls in the Ethics Module

Two pitfalls recur in ethics modules taught by people who care deeply about the technology. The first is the urge to spend most of the module explaining the technology, on the theory that the duties only make sense once the audience understands the model. That order is backwards. Lawyers already understand the duties. The duties are the lawyer's home turf. Lead with the duty, then introduce just enough technology to show how the duty applies. The second pitfall is the urge to over-discipline the discussion, treating the published cases as morality plays in which the sanctioned lawyer becomes a cautionary villain. The audience does not respond to that. The audience responds to a sober reading in which the lawyer in *Mata* is recognizably a competent practitioner who fell into a trap the audience could fall into themselves. Empathy with the sanctioned lawyer produces more learning than contempt for them does.

> **Concern Resolved:**
>
> Ethics is not a chapter at the end of an AI training. It is the table the rest of the training sits on. ABA Formal Opinion 512 gives the module its anchor. The published cases give the module its texture. The audience leaves capable of asking, on their own next filing, what the lawyers in those cases failed to ask.

Part Seven

Designing the Law School Course

Where AI Fits in the Curriculum

Law schools are working out where AI belongs in the JD program, and the answers vary widely. Some schools are running a 2L or 3L elective, often in conjunction with an existing legal technology or legal innovation program. Others are running short modules embedded in the 1L Legal Research and Writing sequence, on the theory that students using AI on their first memo should have at least a structured introduction to it. A small number of schools have built clinics or capstones around AI-related work, especially in areas like access to justice and technology law. Each placement has trade-offs.

The 2L or 3L elective allows depth and assumes a base of legal knowledge to push against. The disadvantage is that students who do not enroll never receive structured AI instruction, and most of those students are using these tools every day. The 1L module reaches every student but has limited time and asks 1Ls to absorb conceptual material before they have built the legal frameworks the material plays against. The clinic or capstone produces the deepest competence in a small number of students. The pragmatic answer at most schools is some combination of these.

Standalone Course or Integrated Across the JD

The case for a standalone course is that AI is now substantial enough to support its own treatment, with its own assessments, materials, and faculty expertise. The case for integration is that AI literacy is more like legal writing or legal research than it is like antitrust: a foundational skill that should appear in many courses rather than be cordoned off in one. The strongest programs do both. They run a standalone course for depth and require touchpoints in core JD courses for breadth, with consistent vocabulary and frameworks across all of them.

Working with Technologists and Practitioners

A law school course on AI is rarely well taught by a single faculty member working in isolation. The course benefits from team teaching, with at least one faculty member with a technical background and one with substantial professional responsibility expertise. Practitioner guests with current courtroom or transactional experience add the operational reality the academic frame does not always supply. The integration is not always smooth, but the audience benefits when the perspectives push against each other rather than reinforcing a single narrative.

Sample Fourteen-Week Syllabus

A workable fourteen-week syllabus, designed to fit a standard semester, looks something like the outline below. Each week assumes a two-hour seminar with reading assigned in advance, plus a hands-on lab or assessment in roughly half of the weeks.

- **Week 1.** Foundations and the next-token sentence. What an LLM is and what it is not.

- **Week 2.** Foundation models, wrappers, and retrieval-augmented generation. The stack and why it matters.
- **Week 3.** Hallucination, bias, and the limits of explainability. The Magesh study and its implications.
- **Week 4.** Prompting as a discipline. Hands-on lab.
- **Week 5.** Verification protocols and the routine that survives. Hands-on lab.
- **Week 6.** Use case triage and the Red, Yellow, Green framework. Workshop.
- **Week 7.** The six duties and ABA Formal Opinion 512. Case discussion.
- **Week 8.** Sanctions cases module. Mata, Park, Noland, Zhang and one current addition.
- **Week 9.** Confidentiality, privilege, and the AI vendor stack. Reading data processing terms.
- **Week 10.** Cross-border issues. EU AI Act, GDPR, and the multi-jurisdictional reality.
- **Week 11.** Supervisory architecture and policy design. Drafting a model AI use policy.
- **Week 12.** AI in litigation: standing orders, disclosure, and verification certifications.
- **Week 13.** AI in transactions and counseling: deal use cases, advisory work, and client-facing risk.
- **Week 14.** Capstone presentations. Each student presents an AI use case for a stated practice setting, with a verification protocol, an applicable rules analysis, and an implementation plan.

> **Curriculum Note:**
>
> Resist the urge to assign a single textbook. The field moves too quickly. Build the syllabus around primary materials: the ABA opinion, the state bar opinions, two or three published cases, the EU AI Act recitals, the Magesh study, and current practice notes. Refresh the readings every term.

Common Pitfalls in Law School AI Courses

Three pitfalls recur in academic AI courses often enough to deserve their own warning. The first is the temptation to treat the course as a survey of the AI industry, with guest speakers from vendors and product walkthroughs filling much of the syllabus. Industry context matters, but the course is not a trade show. Time spent on a vendor demo is time not spent on the underlying skill the student needs.

The second pitfall is the inverse: a course that treats AI as a pure ethics seminar, with extensive discussion of the published sanctions cases and almost no contact with the technology itself. Students leave that course with strong opinions and weak skills. The strongest courses keep ethics central but ground every ethics discussion in actual outputs from actual tools the students have used in lab.

The third pitfall is the assessment problem. A course graded on a single seminar paper produces students who write competently about AI but cannot use it. A course graded on a single hands-on project produces students who can use one tool but cannot generalize. The strongest assessment combines short scenario-based written work, a hands-on lab portfolio, and a capstone that requires the student to design an AI use case for a stated practice setting. Each component tests something the others do not.

Assessment in the Academic Setting

Academic assessment in this subject is harder than in subjects with more stable answers. The instructor cannot rely on a single examination to capture competence, because competence in this domain has multiple components that do not test cleanly under exam conditions. Three formats together do most of the work.

The first is a short scenario-based written exercise, completed in class or as a take-home with limited time, asking the student to identify the appropriate tool tier, draft a working prompt, and identify the rules of professional conduct most directly engaged. The second is a verification lab portfolio, in which the student documents three or four hands-on labs with the verification worksheet completed, errors found, and a short reflection on what was learned. The third is a capstone in which the student presents an AI use case for a stated practice setting, with a verification protocol, an applicable rules analysis, and an implementation plan. The capstone is the closest analog to the work the graduate will do in practice, and it is the assessment that tells the instructor whether the underlying skill transferred.

Existing Programs to Borrow From

A small number of law schools have been running structured legal technology programs long enough to have something to teach. Suffolk University Law School's Institute on Legal Innovation and Technology, Vanderbilt Law School's Program on Law and Innovation, and Stanford Law School's CodeX Center for Legal Informatics each pre-date the generative AI moment and have adapted their existing curricula to incorporate it. Outside the United States, Bucerius Law School's Center for Legal Technology has run a sustained Legal Tech Summer Program. None of these programs is a direct template for any other school. The reason to study them is not to copy a syllabus. It is to see how a sustained institutional commitment to legal technology produces a curriculum that does not reset every time the technology changes.

> **Concern Resolved:**
>
> A law school course on AI is not a one-off elective. It is part of a multi-year curricular shift in how the JD trains lawyers to work with technology. Schools that build the course around primary materials, hands-on labs, and team teaching produce graduates the profession will recognize as ready. Schools that wait will produce graduates the profession will have to retrain.

Part Eight

Designing CLE and In-House Training

Length, Format, and Frequency

Practitioner training operates under different constraints than a law school course. The audience cannot be assembled for fourteen weeks. The competing demands on attention are real and immediate. The successful CLE designer treats time as the dominant constraint and builds backward from it. A one-hour CLE can introduce the foundational sentence, the verification routine, and the duties in Formal Opinion 512, with a single concrete takeaway the participant will use that week. A three-hour CLE can add a hands-on lab and a structured walk through the firm's policy. A full day, available primarily as an annual or onboarding investment, can move the audience through the full curriculum laid out in this guide. Anything beyond a full day in a single sitting is wasted on practitioner audiences.

Frequency matters as much as length. A single CLE, however well designed, decays quickly. The firms that have measurable outcomes from their training are the firms that schedule a substantial annual session, supplemented by short refreshers when policies change or when a notable case appears. Treat the training program the way the firm treats its anti-money-laundering or its conflicts training. It is a recurring obligation, not a one-time event.

Tying Training to Policy and Approved Tools

The strongest CLE programs are tied directly to the firm's own policy and approved-tool list. Generic AI training delivered in isolation is limited. Training that walks the participant through the firm's actual policy, the firm's actual approved tools, the firm's actual verification expectations, and the firm's actual incident reporting process produces compliance the firm can document. When the supervisory rules ask whether reasonable measures were taken, a documented training program tied to the actual operating rules is the answer.

The Supervisory Hook

Model Rules 5.1 and 5.3 give the firm-level training program its strongest internal justification. Partners and lawyers with managerial authority owe a duty to ensure that subordinate lawyers and nonlawyer assistants conform to the rules of professional conduct. Where AI tools are nonlawyer assistants, the supervisory architecture that meets that duty must include training, policy, and oversight. A firm that runs no training is not just a firm with weaker lawyers. It is a firm whose supervisory architecture has gaps the regulator can describe with precision.[17]

Model CLE Outline: Three Hours That Move the Needle

The outline below is a starting point for a three-hour CLE that has the core content most lawyer audiences need. It assumes a single instructor or two co-instructors, a working laptop for each participant, and a single approved AI tool the firm has already deployed.

Three-Hour CLE Outline

Hour One: Foundations and Practical Skills. Open with the next-token sentence and the foundational stack. Ten minutes on hallucinations and the Magesh study. Fifteen minutes on prompting as a discipline. Twenty minutes of structured hands-on prompting. Five-minute break.

Hour Two: Verification and Triage. Open with the verification routine and the worksheet. Twenty minutes of structured verification practice on a fact pattern. Fifteen minutes on the Red, Yellow, Green triage framework. Fifteen minutes of group placement of common practice tasks on the triage grid. Five-minute break.

Hour Three: Ethics, Policy, and Action Items. Twenty minutes on the six duties and ABA Formal Opinion 512. Twenty minutes on the firm's own AI policy, approved-tool list, and incident reporting process. Twenty minutes of case discussion using *Mata*, *Noland*, and one current case. Close with three concrete action items each participant will complete that week.

Reaching the Partner Audience

The hardest audience in most firm-level AI training is the senior partner cohort. Partners have schedule pressure, decades of professional habit that did not include any of this technology, and a strong incentive to delegate to associates rather than learn the tool themselves. A training designed to convert partners into power users will not work. A training designed to give partners enough fluency to supervise the associates who are the power users will work, because that is the role the rules already assign them.

The partner-targeted module that produces the best outcomes runs ninety minutes, not three hours. It opens with the supervisory exposure the partner already faces under Rules 5.1 and 5.3, anchored to the published cases. It walks through the firm's policy and approved-tool list, with the partner identifying where each provision interacts with their practice. It closes with a short practical demonstration of the partner's least-favorite AI capability, so the partner has a working mental model of what the team is using. What the partner does not need is a hands-on prompting lab. The lab belongs in the associate-targeted training. The partner needs the architecture, the supervisory frame, and enough familiarity with the artifact to ask the right questions.

Just-in-Time Training Inside the Matter

Some of the most effective training is not training in any formal sense. It is short, focused guidance delivered inside a live matter, at the moment a lawyer is about to use AI on something consequential. A practice group leader who runs a fifteen-minute session at the start of a complex matter, walking through which AI tools will be used and what the verification expectations are, produces immediate behavior change. A pre-filing checklist that names AI specifically and is reviewed by the signing lawyer at the moment of filing produces immediate verification. These interventions are not substitutes for the structured curriculum. They are the field-level reinforcement that keeps the structured curriculum alive between sessions.

A program that wants to take just-in-time training seriously builds three artifacts and embeds them in matter workflows. A short matter-opening guidance template that surfaces AI use questions at the start of new matters. A pre-filing or pre-send checklist that names AI verification specifically. A brief monthly note from the AI governance committee, distributed firm-wide, summarizing one new case, one new guidance development, and one practical tip. The artifacts are inexpensive to maintain and produce more behavior change than a much longer formal session would produce in isolation.

What CLE Should Not Try to Do

A CLE cannot turn a lawyer into a machine learning engineer. A CLE cannot replace a written policy. A CLE cannot make every participant a confident AI user in three hours. The pragmatic measure of success is more modest: the participant leaves with a foundational understanding, a working verification habit, and enough familiarity with the firm's policy to operate inside it. The ambitious goals are a multi-year program. The CLE is one step in that program.

> **Concern Resolved:**
>
> Effective CLE is not a longer CLE. It is a CLE tied to the firm's actual tools, actual policies, and actual cases. The supervisory rules give the firm-level program the structural support it needs. The firms that build the program now will not be the firms in the next published opinion.

Part Nine

Assessment, Materials, and Maturity

The Curriculum-to-Outcomes Map

Every layer in the four-layer model produces a specific outcome the learner should be able to demonstrate when the training is complete. The map below pairs each layer with the outcome and the assessment artifact that confirms the outcome was reached. The map is the worklist for any program designer and the rubric for any reviewer.

Curriculum Layer	Learner Outcome	Assessment Artifact
Foundations	Can explain how a large language model produces output and why it hallucinates, in plain language.	Short written explanation; oral check during seminar.
Practical skills	Can draft a working prompt for a defined task and run a verification pass that catches seeded errors.	Completed lab worksheet with verification log.
Ethics and professional responsibility	Can identify which of the six ABA Formal Opinion 512 duties are engaged by a given fact pattern and what the duty requires.	Scenario-based written response or class discussion.
Governance	Can read a vendor data processing addendum, identify the training and retention terms, and articulate what to negotiate.	Annotated vendor agreement; mock negotiation memo.

What Real Assessment Looks Like

Most AI training is not assessed at all. The participant attends, signs the attendance sheet, and the program assumes competence has been transferred. That assumption is increasingly untenable. The supervisory rules require reasonable measures, and reasonable measures include some way of confirming that the training did its job. Real assessment does not have to be elaborate. Three approaches do most of the work in practical settings.

The first is a short scenario-based written assessment, completed at the end of the training. Five or six fact patterns, each asking the participant to identify the appropriate tool tier, draft a working prompt or describe a verification step, and identify the rules of professional conduct most directly engaged. The assessment is not designed to fail anyone. It is designed to surface gaps the trainer can address. The second is a periodic spot check. The supervising lawyer reviews a small sample of recent AI-assisted work product, looking for evidence that the verification protocol was actually followed. The third is the incident log itself. A program that catches errors during the verification pass is a program working as

intended. A program where errors only surface when opposing counsel or a court catches them is a program with measurable gaps.

Building a Case Bank

Every effective program builds and maintains a case bank: a curated set of fact patterns, real or hypothetical, used in labs, discussions, and assessments. The case bank pays for itself within a year. New trainers can step in without reinventing materials. The same case can be used to test prompting, verification, ethics, and policy in different modules. And the case bank itself becomes a teaching artifact, because the variety of cases reinforces the breadth of practice contexts in which AI shows up.

Three sources fill out a working case bank. Sanctions cases from the Charlotin database, supplemented with the case names and key facts as published, form the ethics core. Practice-area scenarios drawn from common matters in litigation, transactional work, regulatory advice, and employment counseling fill out the practical labs. Hypotheticals built around vendor diligence, contract review, and policy drafting cover the governance layer.

Train-the-Trainer

Most programs hit the same constraint as they mature: there is more demand for instruction than there are competent instructors to deliver it. A train-the-trainer track is the fix, and the fix has to be deliberate because the skill of teaching this material is not the same as the skill of using the tools. The strongest programs identify two or three internal candidates with both the practice credibility to be heard and the pedagogical instincts to be effective, and invest a meaningful share of program budget in their development. The investment looks like a structured shadowing rotation through the existing modules, dedicated time to build instructor-facing materials, and supervised co-teaching of the first sessions before independent delivery.

The trap to avoid is the assumption that the firm's most technically fluent lawyer is automatically the best instructor. Often the most fluent lawyer has internalized so much that they cannot remember what was hard. A merely competent practitioner who teaches the material with patience and current contact with the lawyer audience usually produces stronger learning outcomes than a technical specialist with thinner pedagogical instincts. Both kinds of teachers are useful. They are useful for different audiences and different modules.

Measuring Whether the Program Is Working

Most AI training programs are not measured at all, beyond attendance counts. The supervisory rules increasingly require some way of telling whether the training did its job, and the case for measurement is not only regulatory. A program that cannot tell whether it is working is a program that cannot improve. Three categories of measure together provide a workable picture.

Direct outcome measures track what learners can actually do. Pre and post training scenario assessments, with results compared at the cohort level rather than the individual level, surface what the curriculum is and is not transferring. The percentage of learners who can correctly identify the appropriate

tool tier on a held-out scenario is a meaningful number. The percentage who can run a verification pass that catches seeded errors is a more meaningful one. Indirect outcome measures track what the lawyers do in their actual work. The incident log is the central artifact: trends in the rate at which AI errors are caught during verification rather than after submission tell the program whether the verification habit is producing the verification behavior. Program-health measures track the program itself. The freshness of the materials, the proportion of the tool list that has been reviewed in the past twelve months, the proportion of the policy that has been updated since the most recent ABA or state bar guidance change, and the time-to-update from a published court ruling to a curriculum revision, are all measurable and all matter.

None of these measures is perfect. None is meaningless. The point of measuring is not to prove the program works. It is to show the program where it does not work, so the program can change.

A Curriculum Maturity Model

Most programs can be placed on a four-stage maturity spectrum, parallel to the firm-level maturity model in the companion guide on AI concerns. The model is useful as a self-audit and as a planning artifact for the next budget cycle.

- **Reactive.** No formal training. AI use is informal. The first time the firm hears about a problem is when opposing counsel raises it. The maturity is incidental to the training because there is no training.
- **Compliant.** An annual CLE or training session is in place. Materials are mostly slide decks. There is no hands-on component. Assessment is attendance. The program meets the floor of the supervisory rules but does not yet produce measurable competence.
- **Proactive.** A multi-module program covers foundations, practical skills, ethics, and governance. Hands-on labs are part of the curriculum. The program is tied to the firm's policy and tools. Assessment includes scenario-based evaluation. Incidents are logged and inform program updates.
- **Strategic.** The program is a recognized internal capability with dedicated owners. Training is iterated based on incident data, evolving guidance, and changing tool capabilities. The case bank is mature. The firm publishes its approach externally and recruits on the strength of the program. The supervisory architecture is documented and would be defensible under examination.

Move one stage at a time. Most firms can advance one stage in a single calendar year if the work is owned by someone with authority and a small budget. Two stages in a year is rare. No stage is permanent. A strategic program that loses its owner reverts to compliant within eighteen months without anyone noticing until something goes wrong.

Maturity Stage	Typical Indicators
Reactive	No written curriculum. AI use is informal. Training, if any, consists of vendor demos. No assessment. No incident log.
Compliant	Annual training in place. Slide-deck format. Tied loosely to a written AI policy. Assessment by attendance.
Proactive	Multi-module curriculum. Hands-on labs. Scenario assessment. Tied to firm policy, tools, and an incident log. Refreshed when guidance changes.
Strategic	Owned program with dedicated capacity. Curriculum iterates from incident data. Mature case bank. Recruited and marketed on. Documented and defensible.

Concern Resolved:

You do not have to build the strategic program in the first year. You have to build the next program, the one that moves the firm or school one stage further than where it is now. The programs that survive the next decade are the programs whose maturity advances faster than the regulatory landscape changes.

Part Ten

What Comes Next

The teaching of AI to lawyers is in its first generation. The materials available now are better than they were two years ago and worse than they will be two years from now. Vendor offerings will continue to evolve. The hallucination rate of leading legal research tools will continue to fall, though probably not to zero. Bar guidance will continue to converge across jurisdictions, with disclosure rules becoming more uniform and supervisory expectations becoming more concrete. The EU AI Act will move from anticipated obligation to operational reality as the August 2026 deadline approaches, and lawyers in EU practice will have to integrate its requirements into their daily work.

None of this changes the structure of the curriculum. The four layers stay the same: foundations, practical skills, ethics and professional responsibility, and governance. The standard the curriculum points toward stays the same: a lawyer competent enough to ask the right questions, choose the right tool, run a verification pass, and document the work. The pedagogical principles stay the same: respect the audience's existing skepticism, teach the next-token sentence early, build hands-on labs that produce verification habits, and tie the training to the rules and the cases.

What changes is the urgency. The first generation of practitioners learning AI will set the patterns the next generation inherits. Schools and firms that build their programs deliberately now will produce lawyers who use these tools well for the rest of their careers. Schools and firms that wait will produce lawyers who learned by accident, with the predictable consequences. The choice is one a school or firm makes once. The compounding effect is permanent.

Teach the foundations. Teach the practical skills. Teach the ethics. Teach the governance. Tie the training to the policies and the tools. Build the case bank. Run the labs. Assess what learners actually retain. Move the program one stage further every year. The lawyers and law students who pass through a program built that way will be the ones the rest of the profession learns from.

Part Eleven

Reference Artifacts

This Part collects three artifacts that any program designer can adapt directly: a complete sample law school syllabus, a consolidated primary-sources reading list organized by module, and rubrics for the three assessment formats the guide recommends. The artifacts are written to be lifted, edited, and put into use, not to be read once and admired.

Artifact One: Sample Law School Syllabus

LAW XXX: Artificial Intelligence in Legal Practice

Course description. A two-credit seminar examining the use of artificial intelligence in legal practice. The course covers the technical foundations a competent practitioner needs, the practical skills that produce reliable use, the duties of professional conduct that apply, and the governance frameworks that translate individual competence into defensible institutional practice. Weekly seminar discussion is paired with structured hands-on labs using current AI tools.

Learning objectives. By the end of the course, students will be able to:

- Explain how a large language model produces output and why it hallucinates, in plain professional language.
- Draft prompts and run verification protocols that meet the standard ABA Formal Opinion 512 sets for AI-assisted legal work.
- Identify which of the six duties under the Model Rules of Professional Conduct are engaged by a given fact pattern involving AI use, and articulate what each duty requires.
- Read a vendor data processing agreement, identify training and retention terms, and describe what they would negotiate.
- Design a defensible AI use case for a stated practice setting, including the verification protocol, applicable rules analysis, and implementation plan.
- Place a given AI use case on the Red, Yellow, Green triage framework and defend the placement.

Assessment.

- Scenario-based written assessments (25%). Four short take-home assessments distributed across the term, each requiring the student to identify the rules engaged, draft a prompt or verification step, and propose a defensible course of action.
- Lab portfolio (25%). Documented evidence of four hands-on labs, each with the verification worksheet completed, errors identified, and a short reflection on what was learned.
- Capstone project (35%). A written memorandum and an oral presentation, described below.

- Class participation (15%). Includes attendance, preparation, and contribution to discussion.

Capstone deliverable. Each student designs an AI use case for a stated practice setting. The deliverable is a twelve-to-fifteen-page memorandum addressed to a senior partner or general counsel, plus a ten-minute oral presentation in the final class. The memorandum must include: a description of the use case, the practice setting and client type, the AI tool selected and the rationale, the verification protocol, an analysis of the rules of professional conduct most directly engaged, an implementation plan with training and oversight requirements, and a documented incident reporting and review process. Students may choose any practice area but must work in a setting different from any direct prior employment.

Weekly Schedule

Week 1. Foundations and the Next-Token Sentence. Read: instructor-selected non-technical primer on transformer architecture; ABA Formal Opinion 512, Sections I and II. In class: live demonstration of LLM behavior on a series of prompts; discussion of student mental models.

Week 2. Foundation Models, Wrappers, and Retrieval. Read: instructor-curated set of three current vendor whitepapers; Magesh et al. study with attention to methodology. In class: walk-through of the AI stack and where data flows.

Week 3. Hallucinations, Bias, and Limits of Explainability. Read: Magesh et al. (full); Solicitors Regulation Authority Risk Outlook on AI; Comment 8 to ABA Model Rule 1.1. In class: live hallucination demonstration; discussion of bias amplification in adverse decisions.

Week 4. Prompting as a Discipline (Lab 1). Read: ABA Formal Opinion 512, Section III; instructor-prepared memo on prompt patterns. In class: hands-on prompting lab on a research question, with iterative refinement.

Week 5. Verification Protocols (Lab 2). Read: Florida Bar Advisory Opinion 24-1; California State Bar Practical Guidance, sections on competence. In class: verification lab on a deliberately hallucinated brief.

Week 6. Use Case Triage (Workshop). Read: companion guide chapter on the Red, Yellow, Green framework. In class: workshop placing common practice tasks on the triage grid.

Week 7. The Six Duties and ABA Formal Opinion 512. Read: ABA Formal Opinion 512 (full); Florida Opinion 24-1 (full); California Practical Guidance (full). In class: structured discussion of each duty.

Week 8. Sanctions Cases Module. Read: *Mata v. Avianca*; *Park v. Kim*; *Noland v. Land of the Free, L.P.*; *Zhang v. Chen*; one current addition. In class: case discussion using the sample discussion questions.

Week 9. Confidentiality, Privilege, and the Vendor Stack. Read: ABA Formal Opinion 477R; Florida Opinion 24-1 (privilege analysis); a redacted vendor data processing agreement supplied by the instructor. In class: live walk-through of the vendor agreement.

Week 10. Cross-Border Issues. Read: GDPR Articles 5, 6, 28, 32, 44-50 (excerpts); EU AI Act Articles 6-15 (excerpts); Singapore Registrar's Circular No. 1 of 2024; NSW Practice Note SC Gen 23. In class: jurisdictional mapping exercise.

Week 11. Supervisory Architecture and Policy Design (Lab 3). Read: Model Rules 5.1 and 5.3; two model AI use policies supplied by the instructor. In class: policy drafting workshop.

Week 12. AI in Litigation. Read: Judge Brantley Starr's standing order; recent additions to the Charlotin database. In class: simulated pre-filing verification exercise.

Week 13. AI in Transactions and Counseling (Lab 4). Read: instructor-prepared materials on contract review and advisory work. In class: transactional lab on a deal use case.

Week 14. Capstone Presentations. Read: none. In class: ten-minute student presentations with structured peer feedback.

Artifact Two: Primary-Sources Reading List

The materials below are organized by module rather than by jurisdiction or document type. Each entry names the source and indicates briefly what it teaches. The list is a starting point. Refresh it every term, because in a fast-moving field the omitted current case is often the one the audience will hear about first.

I. Foundational and Technical Concepts

- An instructor-selected non-technical primer on how transformer-based language models generate text. The primer should support the next-token sentence without requiring mathematical fluency.
- Current vendor whitepapers from at least two foundation model providers, used to ground the architectural discussion in actual products rather than abstract concepts.
- Magesh, Surani, Dahl, Suzgun, Manning, and Ho, "Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools," 22 J. Empirical Legal Stud. 216 (2025). The cleanest empirical baseline available for hallucination rates in legal-specific AI tools.

II. ABA and U.S. State Bar Guidance

- ABA Standing Committee on Ethics and Professional Responsibility, Formal Opinion 512, "Generative Artificial Intelligence Tools" (July 29, 2024). The touchstone for U.S. lawyers and the natural anchor for any ethics module.
- Florida Bar Board of Governors, Ethics Opinion 24-1, "Lawyers' Use of Generative Artificial Intelligence" (January 19, 2024). Informed consent and confidentiality applied to AI tools.
- California State Bar, "Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law" (November 16, 2023). Among the most extensively developed state bar opinions, with concrete fee and billing guidance.
- Comment 8 to ABA Model Rule 1.1. The textual foundation for treating AI competence as an extension of existing duties.
- ABA Formal Opinion 477R, "Securing Communication of Protected Client Information" (May 22, 2017). The predecessor framework on cloud and electronic communications, useful for the privilege analysis applied to AI vendor processing.

III. Sanctions Cases (Charlotin Database)

- *Mata v. Avianca, Inc.*, 678 F. Supp. 3d 443 (S.D.N.Y. 2023). The leading case and the natural opening for any sanctions module.

- *Park v. Kim*, 91 F.4th 610 (2d Cir. 2024). Appellate-level sanctions reinforcing that the duty does not stop at the trial court.
- *Noland v. Land of the Free, L.P.*, Cal. Ct. App., 2d Dist. (Sept. 2025). The first published California Court of Appeal opinion addressing AI-fabricated authority.
- *Zhang v. Chen*, 2024 BCSC 285. The Canadian parallel that illustrates how the same pattern recurs across common-law jurisdictions.
- Damien Charlotin, "AI Hallucination Cases" database (damiencharlotin.com/hallucinations). The ongoing reference, refreshed continuously.

IV. International Guidance Outside the United States

- Solicitors Regulation Authority, "Risk Outlook Report: The Use of Artificial Intelligence in the Legal Market" (2023, with subsequent updates). The UK regulator's perspective, anchored in accuracy and accountability.
- Law Society of Ontario, "Licensee Use of Generative Artificial Intelligence" white paper (April 2024). The Canadian counterpart, with quick-start checklists through the LSO Technology Resource Centre.
- Supreme Court of New South Wales, Practice Note SC Gen 23, "Use of Generative Artificial Intelligence" (effective February 3, 2025). The Australian disclosure and verification regime.
- Singapore Registrar's Circular No. 1 of 2024, "Guide on the Use of Generative Artificial Intelligence Tools by Court Users." Places full responsibility for AI output on the user across the Supreme Court, State Courts, and Family Justice Courts.

V. European Union and Data Law

- Regulation (EU) 2024/1689 (Artificial Intelligence Act). Risk-based AI regulation, in force August 1, 2024, with most obligations applying from August 2, 2026.
- General Data Protection Regulation (Regulation (EU) 2016/679), Articles 5, 6, 28, 32, and 44-50 (excerpts). The privacy backbone any module on data flows has to ground itself in.
- Italian Garante per la protezione dei dati personali, decision halting ChatGPT processing of Italian data subjects (March 30, 2023), and subsequent administrative fine of fifteen million euros (December 2024). The first major EU enforcement against a generative AI provider.

VI. Court Standing Orders and Practice Materials

- Standing Order, Judge Brantley Starr, U.S. District Court for the Northern District of Texas (May 30, 2023). The first widely-cited U.S. federal standing order on AI use in filings.
- Local rules and standing orders from the practitioner's home jurisdiction. The instructor should supply a current set, refreshed each term.

VII. Companion Guides in This Series

- Levy, *Addressing AI Concerns* (2026). The companion volume on the firm-level governance and risk allocation framework.

- Levy, *Managing AI Hallucinations* (2026). The depth treatment of the verification protocol and the Red, Yellow, Green triage framework.
- Other titles in the series as relevant to the module under instruction.

Artifact Three: Assessment Rubrics

The rubrics below assess the three formats this guide recommends. Each rubric uses four performance levels and four dimensions. The descriptors are short to keep the rubric usable in practice; the rubric the instructor adopts can be expanded with examples drawn from the case bank.

Rubric for Scenario-Based Written Assessment

Dimension	Exemplary	Proficient	Developing	Not Yet
Issue identification	Identifies all engaged duties and any non-obvious ones.	Identifies the duties most directly engaged.	Identifies one or two duties; misses related ones.	Misidentifies or omits the duties at issue.
Rule analysis	Articulates the standard accurately and connects it to current authority.	Articulates the standard accurately.	States the rule with material gaps.	Misstates or omits the rule.
Application to facts	Applies each rule to specific facts with structured reasoning.	Applies the rule to the facts adequately.	Application is conclusory or partial.	Application is missing or incorrect.
Conclusion and remediation	Proposes a defensible course of action with documentation steps.	Proposes a workable course of action.	Offers a partial or unclear recommendation.	No actionable recommendation.

Rubric for Lab Portfolio

Dimension	Exemplary	Proficient	Developing	Not Yet
Verification thoroughness	Worksheet complete; all errors found, including subtle ones.	Worksheet complete; major errors found.	Worksheet partial; some errors missed.	Worksheet absent or perfunctory.
Reflection quality	Identifies what changed in the student's working approach.	Names what was learned with specificity.	Generic reflection without specifics.	Reflection absent or rote.
Tool selection rationale	Justifies tool choice against the triage framework.	Justifies tool choice with reasoning.	Tool choice asserted without reasoning.	Tool choice appears default or unconsidered.
Documentation	Audit trail sufficient to defend the work to a regulator.	Documentation complete for normal use.	Documentation has gaps that would matter on review.	Documentation insufficient.

Rubric for Capstone Project

Dimension	Exemplary	Proficient	Developing	Not Yet
Use case design	Clear, defensible, and well-tailored to the practice setting.	Workable use case appropriate to the setting.	Use case unclear or partially mismatched.	Use case undefined or not viable.
Verification protocol	Specific, repeatable, and tied to professional duties.	Adequate protocol with the necessary steps.	Protocol thin or partly missing.	Protocol absent or incoherent.
Rules analysis	Accurate, complete, and current; engages each duty in depth.	Accurate analysis of the engaged duties.	Analysis with material gaps.	Analysis missing or wrong.
Implementation feasibility	Realistic in the setting; addresses training and oversight.	Plan workable in the setting described.	Plan partially realistic.	Plan unrealistic or absent.
Presentation clarity	Organized, well-written, and clearly delivered.	Clear writing and adequate delivery.	Organization or clarity weakens the work.	Hard to follow or unfinished.

Concern Resolved:

A program that distributes the artifacts above with the curriculum has done most of the design work the next instructor would otherwise face. The artifacts are starting points. They become the program's own materials only after the first cycle of use, when the instructor knows what to keep, what to change, and what the audience required that the artifacts did not yet provide.

Glossary of Key Terms

Adult Learning.
The body of pedagogical practice oriented to learners with established mental models, professional experience, and limited time. Lawyer audiences are adult learners by definition, and curricula that ignore the implications produce poor outcomes regardless of how good the underlying material is.

Approved-Tool List.
A maintained inventory of AI tools the organization permits for use, classified by data sensitivity, with documented data processing terms, retention policies, and permitted use cases. The list is the operational artifact that turns a written policy into daily practice.

Case Bank.
A curated set of fact patterns, real or hypothetical, used in labs, discussions, and assessments across an AI training program. A working case bank is the single most reusable asset a program builds.

Competence (Model Rule 1.1).
The duty to provide competent representation, including the duty in Comment 8 to keep abreast of the benefits and risks of relevant technology. Adopted in some form in a substantial majority of U.S. states. Now plainly extends to a reasonable understanding of AI tools used in practice.

Foundation Model.
A large general-purpose AI model trained on broad data and adapted to many downstream tasks. Examples include the GPT, Claude, Gemini, and Llama families. Legal-specific tools are typically wrappers around foundation models, with retrieval or fine-tuning layers added.

Hallucination.
An AI output that is factually incorrect, fabricated, or unsupported by source material, presented with apparent confidence. The leading source of legal sanctions arising from AI use. Hallucination rates fall with retrieval-augmented architectures but do not reach zero.

Hands-On Lab.
A structured exercise in which learners use an actual AI tool to complete a defined task, run a verification pass on the output, and debrief the result. The single most effective training format for producing reliable practical skill.

Maturity Model.
A staged framework, typically four stages, used to describe the current state of a training program and to plan its next step. The reactive, compliant, proactive, and strategic stages map to recognizable patterns in actual programs.

Next-Token Prediction.
The mechanism by which a large language model produces text. The model selects the most probable next token given the text that came before, then repeats the process. The single most useful concept in foundational AI training for lawyers.

Prompt Engineering.

The discipline of constructing inputs to AI tools so as to produce reliable, useful outputs. The phrase has acquired some folkloric content, but the underlying skill is real and teachable, and it follows from a clear understanding of how the model conditions on prior context.

Retrieval-Augmented Generation (RAG).

An AI architecture that retrieves relevant documents from a knowledge base before generating a response, intended to ground outputs in source material. Reduces hallucination rates relative to general-purpose foundation models but does not eliminate them.

Sanctions Database.

A reference to the public Charlotin database of legal decisions in which courts have addressed lawyer reliance on AI-generated content. As of early 2026 the database contained more than 1,300 documented decisions and is the most useful single source for case-based ethics teaching.

Supervisory Hook.

The pedagogical and structural justification for firm-level AI training drawn from Model Rules 5.1 and 5.3. Treating AI tools as nonlawyer assistants, as ABA Formal Opinion 512 does, makes the supervisory architecture the rules contemplate the natural home for the training program.

Triage Framework (Red, Yellow, Green).

A use-case classification framework that sorts tasks by the consequence of an AI error and the verifiability of the output. Red tier work is unsuitable for AI delegation. Yellow tier work belongs in AI only with rigorous verification. Green tier work is where AI tools deliver most of their reliable value.

Verification Routine.

A short, repeatable procedure for confirming the accuracy of AI-generated work product before reliance, typically including independent retrieval of every citation, comparison of every quotation against source text, and confirmation of every material factual assertion. The single most important practical skill in any AI training for lawyers.

Endnotes

1. ABA Standing Committee on Ethics and Professional Responsibility, Formal Opinion 512, "Generative Artificial Intelligence Tools" (July 29, 2024); Florida Bar Board of Governors, Ethics Opinion 24-1, "Lawyers' Use of Generative Artificial Intelligence" (January 19, 2024); California State Bar, "Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law" (November 16, 2023); Solicitors Regulation Authority, "Risk Outlook Report: The Use of Artificial Intelligence in the Legal Market" (2023, with subsequent updates); Law Society of Ontario, "Licensee Use of Generative Artificial Intelligence" (white paper, April 2024); Supreme Court of New South Wales, Practice Note SC Gen 23 (issued November 21, 2024; effective February 3, 2025); Singapore Registrar's Circular No. 1 of 2024, "Guide on the Use of Generative Artificial Intelligence Tools by Court Users"; Regulation (EU) 2024/1689 (Artificial Intelligence Act).

2. The four-layer model in this guide is consistent with the structure of professional responsibility analysis in ABA Formal Opinion 512, which distinguishes underlying knowledge of the technology, the practical workflow, the duties of professional conduct, and the supervisory and governance layer.

3. Comment 8 to ABA Model Rule 1.1, adopted by the ABA in 2012, requires a lawyer to keep abreast of the benefits and risks associated with relevant technology. The duty to learn relevant technology has now been incorporated, in some form, in the rules of a substantial majority of U.S. jurisdictions. The relevance of this duty to AI tools is explicit in ABA Formal Opinion 512.

4. ABA Standing Committee on Ethics and Professional Responsibility, Formal Opinion 512 (July 29, 2024), addressing Model Rules 5.1 and 5.3 and treating generative AI tools as nonlawyer assistants for supervisory purposes.

5. Standing orders by individual U.S. judges, including Judge Brantley Starr in the Northern District of Texas (May 30, 2023), and practice notes such as Supreme Court of New South Wales Practice Note SC Gen 23, illustrate the active role of the bench in setting AI disclosure and verification expectations.

6. The "next-token prediction" framing is the standard non-technical description of how transformer-based language models generate text. It is consistent with the public technical descriptions of the GPT, Claude, Gemini, and Llama model families, and is the most pedagogically useful starting point for non-technical lawyer audiences.

7. The "stack" framing, separating the foundation model, the wrapper or application layer, and any retrieval corpus, is the standard architectural description used by enterprise AI vendors and tracks the data flow that informs the privacy, confidentiality, and cross-border analysis required by ABA Formal Opinion 512 and related state and international guidance.

8. Magesh, Surani, Dahl, Suzgun, Manning, and Ho, "Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools," 22 J. Empirical Legal Stud. 216 (2025). The study reported hallucination rates of approximately 17% (Lexis+ AI), 33% (Westlaw AI-Assisted Research), and 43% (GPT-4) across federal case-law queries.

9. ABA Standing Committee on Ethics and Professional Responsibility, Formal Opinion 512 (July 29, 2024); see also Singapore Registrar's Circular No. 1 of 2024 and Supreme Court of New South Wales Practice Note SC Gen 23, each imposing explicit verification requirements for AI-assisted submissions.

10. ABA Standing Committee on Ethics and Professional Responsibility, Formal Opinion 512, "Generative Artificial Intelligence Tools" (July 29, 2024). Addresses competence (Rule 1.1), confidentiality (Rule 1.6), communication (Rule 1.4), candor (Rules 3.1 and 3.3), supervisory responsibilities (Rules 5.1 and 5.3), and reasonable fees (Rule 1.5).

11. Comment 8 to ABA Model Rule 1.1, adopted by the ABA House of Delegates in 2012. State adoption has occurred in some form in approximately forty U.S. jurisdictions, with continued adoption activity since. The Comment

8 obligation has been the textual foundation for state bar guidance treating AI competence as an extension of existing duties rather than as a new obligation.

12. Damien Charlotin, "AI Hallucination Cases" database, available at damiencharlotin.com/hallucinations. As of early 2026 the database contained more than 1,300 documented decisions across multiple jurisdictions.

13. Mata v. Avianca, Inc., 678 F. Supp. 3d 443 (S.D.N.Y. 2023). $5,000 sanction imposed for filing six fabricated case citations generated by ChatGPT and providing fabricated opinion texts when challenged. Court found "subjective bad faith" in the failure to verify.

14. Park v. Kim, 91 F.4th 610 (2d Cir. 2024). Second Circuit referred attorney to its Grievance Panel for citing a nonexistent case generated by ChatGPT and making no inquiry into its validity.

15. Noland v. Land of the Free, L.P., Cal. Ct. App., 2d Dist. (Sept. 2025). $10,000 sanction imposed; counsel reported to the State Bar of California. First published California Court of Appeal opinion addressing AI-fabricated legal authority. The court found that briefs were "replete" with fabrications generated using multiple consumer AI tools.

16. Zhang v. Chen, 2024 BCSC 285 (Supreme Court of British Columbia). Court ordered counsel to personally compensate opposing counsel for time spent unwinding ChatGPT-generated fabricated authority. Law Society of British Columbia opened investigation.

17. ABA Model Rules of Professional Conduct 5.1 and 5.3, addressing supervisory responsibilities of partners, managers, and supervisory lawyers with respect to subordinate lawyers and nonlawyer assistants. ABA Formal Opinion 512 treats generative AI tools as nonlawyer assistants for purposes of these rules and grounds the firm-level training and policy obligation in the existing supervisory framework rather than in any AI-specific rule.

18. Florida Bar Board of Governors, Ethics Opinion 24-1, "Lawyers' Use of Generative Artificial Intelligence" (January 19, 2024); California State Bar, "Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law" (November 16, 2023); Law Society of Ontario, "Licensee Use of Generative Artificial Intelligence" (April 2024). These three documents, taken together with ABA Formal Opinion 512, are the most extensively developed bar-level guidance available in English as of the date of this guide and form the substantive backbone of the ethics module recommended in Part Six.

19. Regulation (EU) 2024/1689 (Artificial Intelligence Act), entered into force August 1, 2024; majority of obligations applicable from August 2, 2026. High-risk classification under Annex III applies to specified uses including those affecting access to justice and the administration of law. Lawyers practicing in EU jurisdictions, or whose work touches data of EU subjects, should treat the August 2026 deadline as a working planning horizon for curriculum updates.

20. The maturity model in Part Nine is parallel to, and consistent with, the firm-level maturity model in the companion guide *Addressing AI Concerns*. Both rest on the observation that programs advance one stage at a time, that no stage is permanent, and that the supervisory architecture and the curriculum architecture should reinforce each other rather than evolve independently.