AI for Lawyers

Drafting, Research, and Client Work in the Age of LLMs

Marcus Hale

AI for Lawyers

Drafting, Research, and Client Work in the Age of LLMs

by Marcus Hale

CONTENTS

I What's Actually Changed — and What Hasn't

II The Hallucination Problem: Mata v. Avianca and What We Learned

III The Tools That Matter: General-Purpose vs. Legal-Specific

IV Drafting: Contracts, Briefs, and Memos

V Legal Research with AI: Speed, Gaps, and the Verification Imperative

VI Discovery and Document Review: Where AI Pays for Itself

VII Due Diligence: M&A, Real Estate, and Regulatory Deals

VIII Client Communication: Drafting Summaries, Emails, and Status Updates

IX Deposition Preparation: Outlines, Chronologies, and Impeachment Sets

X Litigation Strategy: AI as a Thinking Partner, Not a Decision-Maker

XI Compliance, Ethics, and the Model Rules: What the Bar Has Actually Said

XII Confidentiality and What Not to Paste: Rule 1.6 in the Age of LLMs

XIII Billing and AI: The Candid Conversation

XIV The Next Five Years: What to Watch, What to Ignore, and How to Stay Competent

CHAPTER I

What's Actually Changed — and What Hasn't

Every few years, the legal profession convinces itself that this technology — this one, right now — will transform the practice of law. It almost never does. Not on the promised timeline. Not in the promised way. So when attorneys ask whether AI is different, the honest answer is: partly yes, and in ways that matter — but no, and in ways that matter just as much.

That tension is what this book is about. Not the hype. The actual inventory.

The Machine You Are Dealing With

Before you evaluate any AI tool, you need a working model of what it is. Not a technical deep-dive — just enough to stop you from making category errors that will cost you, or your clients, badly.

A large language model is a statistical prediction engine. It was trained on an enormous corpus of text — books, websites, legal databases, forum posts, court opinions, everything — and it learned, at a very deep level, which words tend to follow which other words in which contexts. When you type a prompt, the model generates a response token by token, each token chosen based on the probability distribution its training encoded. That's it. There is no database being queried. There is no understanding of truth. There is no memory of your last conversation (unless the system explicitly implements one). There is no legal reasoning in the sense a lawyer would recognize.

What this means in practice: the model will produce text that sounds like the answer to your question. Whether it is the answer is a separate question entirely. The two properties — plausibility and accuracy — are essentially unrelated outputs of the same process. A well-trained model will write you a confident, fluently structured legal argument citing cases that do not exist with the same ease that it writes you an argument citing cases that do. It cannot tell the difference. Neither can the text it produces.

This is not a bug that will be fixed next quarter. It is a structural feature of how these systems work. The researchers call it "hallucination." The better term might be "confabulation" — the model is not malfunctioning; it is doing exactly what it was built to do. Generating plausible text. For legal work, that distinction is critical.

Keep that model in your head. Everything else follows from it.

What Has Actually Changed

With that foundation laid, let's be fair. Something has changed. Several things, actually — and practitioners who dismiss LLMs entirely because of the hallucination problem are leaving real efficiency gains on the

table.

Drafting Speed

The most unambiguous gain is speed in first-draft production. Not quality. Speed.

A mid-level associate at a transactional firm who previously spent three hours drafting a confidentiality agreement template from scratch can now generate a serviceable first draft in fifteen minutes. That draft will need to be reviewed, edited, corrected, and adapted — but the blank-page problem, the one that burns the first ninety minutes of every drafting assignment, is largely gone. For routine documents — NDAs, demand letters, straightforward employment agreements, client update memos, deposition summaries — this compression is real and measurable.

The 2023 study by Fabrizio Dell'Acqua and colleagues at Harvard Business School (involving 758 consultants at BCG, but the findings map) showed knowledge workers using GPT-4 completed tasks 25% faster and produced work rated 40% higher in quality by blind evaluators. The legal analogy is imperfect — legal drafting has higher error costs than business consulting deliverables — but the speed finding holds across multiple domains and multiple tools. You will draft faster. That is not a projection; it is well-documented.

The mechanism is mundane. AI is extraordinarily good at pattern-completion for documents it has seen many times before. Standard contract clauses, boilerplate motion language, routine memo structure — all of this is densely represented in training data. The model has, in effect, read millions of documents of those types and can reproduce their structure and language fluently. For high-volume routine work, that is genuinely useful.

Large-Document Pattern Recognition

The second real capability gain is in pattern recognition across large text volumes. This is where AI stops being a faster typist and starts being something qualitatively different.

Feed a 200-page commercial lease into a capable model with a well-constructed prompt, and it will pull every tenant termination right, flag every non-standard liability cap, and surface every consent-to-assignment clause — faster than a junior associate reading linearly, and without the degradation in attention that comes at page 140 of a dense document at 11 p.m. before closing. For due diligence work, discovery document review, and contract portfolio analysis, this pattern-matching at volume is the most significant operational shift in legal practice since keyword search was introduced to e-discovery.

This is not without its own risks. The model may miss something. It may characterize a clause incorrectly. It may fail on genuinely novel or complex drafting that doesn't match its training patterns. But as a triage and organizing tool — to sort a document set before human review, to flag potential issues before a lawyer reads for substance — it performs at a level that justifies adoption. Chapter 6 addresses discovery review in detail, and Chapter 7 handles due diligence. The point here is simply: this capability is real, it is available now, and ignoring it is not professional caution. It's professional lag.

Synthesis and Issue-Spotting

A third genuine gain is synthesis. Ask a well-designed legal AI tool to survey the split in circuit authority on mandatory arbitration of PAGA claims, and it will give you a structured, readable synthesis of the landscape — faster than a Westlaw search-and-read workflow, and often more usable as a starting framework. Issue-spotting from a fact pattern is similar: describe a client's situation and ask what legal questions it raises, and a current model will surface issues that a harried junior associate might not flag.

Notice the careful framing there. Starting framework. Might not flag. Because here is where the capability gain and the permanent limitation collide.

What Has Not Changed

The Model Does Not Know the Law

This sentence deserves to stand alone, because lawyers consistently underestimate it.

The model does not know the law. It was trained on text that contains legal information, but it has no access to current statutes, current regulations, or current case law as of any fixed date. It has no index to binding authority. It cannot distinguish a majority rule from a minority one unless that distinction appears explicitly in its training data. It has no concept of jurisdiction beyond what it can infer from context. It does not know what happened in a circuit after its training cutoff — and every major model has a training cutoff, typically six months to a year behind its release date.

In May 2023, attorneys Steven Schwartz and Peter LoDuca submitted a brief in Mata v. Avianca, Inc. (S.D.N.Y. No. 22-cv-1461) that cited, among other cases, Varghese v. China Southern Airlines Co., Shaboon v. EgyptAir, Petersen v. Iran Air, and Martinez v. Delta Air Lines — none of which existed. ChatGPT had fabricated them, complete with plausible-sounding holdings, case numbers, and party names. When the court ordered the attorneys to produce the decisions, they couldn't. The case became the most prominent early demonstration that AI legal research without verification is not research. It is fiction dressed in citation format.

The full anatomy of that case gets its own chapter. The reason it appears here, in the first chapter, is this: Mata was not a story about reckless attorneys or defective software. It was a story about a category error. The attorneys believed the model knew things it did not know. They assigned it the epistemic authority of Westlaw. That misunderstanding — the conflation of fluent output with verified knowledge — is the single most dangerous mistake a lawyer can make with this technology.

The model doesn't know. It generates text that sounds like knowing.

Citations Are Not Reliable

Related but distinct: even when a model produces a real case citation, you cannot trust that it has characterized the holding correctly. LLMs do not read cases the way lawyers do. They pattern-match against training text that described or summarized cases. The description may have been incomplete. The training data may have included secondary sources that misstated the holding. The model may blend two cases. It may correctly identify a case but misstate the circuit, the year, or whether the decision was affirmed or reversed on appeal.

The ABA's Formal Opinion 512, issued in July 2024, is direct on this point. The Opinion addresses AI use under the Model Rules and observes that lawyers "must take steps to ensure that any information provided by" generative AI tools "is accurate" — which, when applied to legal authority, means independent verification through authoritative sources. The Opinion does not say "verify if you have doubts." It implies verification as a baseline obligation whenever AI-generated legal content goes into a work product that will be used in a matter.

The bar is not treating this as aspirational guidance. It is treating it as a competence requirement.

Professional Responsibility Stays With You

This one some lawyers have absorbed intellectually without absorbing it practically.

No AI tool carries professional responsibility. Not Harvey. Not Lexis+ AI. Not Claude. Not whatever ships next year. The professional responsibility — for the accuracy of the work, for the supervision of non-lawyer assistance, for candor to the tribunal, for the protection of client confidences — belongs to the attorney of record. Every bit of it.

ABA Formal Opinion 512 maps this clearly across several Model Rules. Under Rule 1.1 (competence), a lawyer using generative AI must understand, at least at a functional level, what the tool does and doesn't do, and must review AI-generated work with the judgment and diligence the representation requires. Under Rule 5.3 (supervision of non-lawyer assistance), the lawyer must supervise AI-generated output as they would supervise a junior associate or contract attorney — which means reviewing it, not just forwarding it. Under Rule 3.3 (candor toward the tribunal), submitting a brief with AI-generated citations the lawyer did not verify is not a technology failure. It is a candor failure.

The Opinion's framing is useful: generative AI is a form of non-lawyer assistance, and the rules governing supervision of non-lawyer assistance apply. You would not let a paralegal's first draft of a brief go to the court unreviewed. The same standard applies when the paralegal is an LLM.

The Speed Gain Has a Liability Shadow

Here is the uncomfortable corollary. The speed gain is real — but it comes with a corresponding risk of confidence inflation. When a junior associate spends three hours drafting a motion, she reads it multiple times in the process. She catches her errors. When an AI produces that same draft in three minutes, the temptation is to review it less thoroughly because it looks polished. The grammar is clean. The structure is logical. The

citations look right.

That look is the hazard. Polished presentation is a property the model produces automatically. Accuracy is not. The two are decoupled in a way that has no analog in human-drafted work, where prose quality and legal accuracy tend to correlate (imperfectly, but noticeably). With LLM output, you can receive a beautifully written, perfectly structured, completely wrong document. The editorial quality is no evidence of legal quality.

Build that into your review process. The chapter on drafting will give you specific protocols. For now, just internalize the principle: more polished does not mean more accurate.

The State of the Tools in 2026

By mid-2026, the market has split into two rough categories. General-purpose large language models — Claude (Anthropic), GPT-4-class models (OpenAI), Gemini (Google) — operate without native access to verified legal databases. They are powerful, flexible, and available at relatively low cost, but they do not retrieve from Westlaw or Lexis in real time, and their citations are generated, not retrieved. Legal-native platforms — Harvey, Lexis+ AI, Westlaw Precision, and their competitors — are built on top of authoritative legal databases, use retrieval-augmented generation (RAG) architectures to ground responses in actual documents, and are designed specifically to reduce (though not eliminate) citation hallucination.

The distinction matters. It doesn't resolve the verification problem entirely — RAG-based systems can still mischaracterize retrieved documents — but it changes the risk profile meaningfully. Chapter 3 covers this landscape in detail, including data retention policies and confidentiality implications that general-purpose tools raise under Model Rule 1.6. What you need at this stage is simply the frame: the tool categories are not equivalent, the risks are not identical, and choosing the right tool for the right task is itself an exercise of professional judgment that the Model Rules now require you to make.

Why This Moment Feels Different (And Why That's Partly Correct)

It is fair to ask: haven't lawyers heard this before? Legal research software, then Boolean search, then e-discovery platforms, then predictive coding, then "legal analytics" — each wave brought vendor promises of transformation. Most practitioners learned to wait out the hype and adopt selectively.

This moment is genuinely different in one specific way. Prior legal technology innovations changed how lawyers performed tasks that still required trained legal judgment at every step. Westlaw made case law accessible. E-discovery platforms made document review tractable. Neither could produce professional work product. LLMs can produce something that looks like professional work product — a drafted brief, a research memo, a client letter — at a level of fluency that creates a plausible substitution pressure that prior tools never created.

That substitution pressure is why bar associations are issuing ethics opinions. It's why federal judges are adopting AI disclosure rules. It's why this book exists. The threat is not that AI will replace lawyers. The threat is that lawyers will stop doing the parts of their job that AI cannot do — the independent judgment, the case theory, the credibility assessment, the client relationship — and assume the machine has covered them. It hasn't. It can't.

The lawyers who will practice well with AI in 2026 and beyond are not the ones who adopt every tool enthusiastically. They are the ones who understand precisely what each tool can and cannot do, build verification workflows that are proportionate to the risk, and apply their own trained legal judgment to everything the machine produces. That is the same standard the profession has always required. AI just makes it harder to remember to apply it.

A Framework for Reading This Book

The chapters that follow address specific task categories: drafting, research, discovery, due diligence, client communication, deposition prep, strategy. Each chapter gives you concrete workflows and prompt architectures. But every chapter is built on the same underlying structure, which you can carry into any new tool, any new task, and any new AI development that emerges after this book goes to print.

What is the model actually doing here? Is it generating from training (higher hallucination risk) or retrieving from a verified source (lower, but not zero)? What is the training cutoff? Does the platform have a data retention policy that implicates Rule 1.6?

What are the error consequences? A factual error in a client summary letter is bad. A citation error in a brief filed with a federal court is sanctionable. Calibrate your verification effort to the consequence, not to your trust in the tool.

Who is supervising? Under Rule 5.3, you are. That means you read it. Not skim it — read it. You compare citations to their sources. You confirm holdings. You apply your own legal judgment to the analysis the model produced. AI is a first-draft engine and a thinking partner. It is not a supervising attorney.

What does the client know? This will be litigated over the next several years. Some jurisdictions are moving toward disclosure requirements when AI was substantially used in client work product. Even where disclosure isn't yet required, the competence and communication obligations of Rules 1.1 and 1.4 apply. The client's interests, not your efficiency preferences, set the standard.

One more thing before we move on. The profession has a tendency, when confronted with a new technology it doesn't fully understand, to polarize. Some practitioners adopt uncritically — impressed by the output, unaware of the failure modes. Others reject categorically — protecting their judgment by refusing to engage with the tools at all.

Neither posture serves clients.

The attorney who uses AI without understanding hallucination is a liability waiting to happen. The attorney who refuses to engage with AI while competitors compress document review costs by sixty percent is not being cautious. She's being left behind in ways that will eventually affect her ability to serve clients competitively.

What the moment calls for is neither enthusiasm nor rejection. It calls for the same disciplined skepticism that good lawyers have always applied to expert witnesses, to opposing counsel's representations, and to their own first drafts. Trust, but verify. Except with AI, skip the first part. Just verify.

The rest of this book is about how.

CHAPTER II

The Hallucination Problem: Mata v. Avianca and What We Learned

The brief looked fine. That was the problem.

On March 1, 2023, attorney Steven Schwartz of the New York personal injury firm Levidow, Levidow & Oberman filed an opposition brief in Mata v. Avianca, Inc., a straightforward foreign airline liability case pending before Judge P. Kevin Castel in the Southern District of New York. Roberto Mata had been injured by a metal serving cart on an Avianca flight. Avianca moved to dismiss on statute-of-limitations grounds. Schwartz, who had practiced law for over thirty years, needed case authority on the Montreal Convention and related airline liability questions. He turned to ChatGPT. The tool returned a confident, well-formatted list of cases — complete with reporters, volume numbers, page numbers, and parentheticals summarizing each holding. The brief was filed.

Six of the cases were invented. None of them existed.

What happened next became the most consequential cautionary event in the first wave of legal AI adoption — not because the outcome was catastrophic in dollar terms (it wasn't), but because it happened in open court, under a federal judge's eye, with the full paper trail preserved in a public docket. Mata v. Avianca is the legal profession's first hallucination case study. It will not be the last.

What Actually Happened: The Docket Record

When Avianca's counsel searched for the cases Schwartz had cited, they found nothing. They flagged the discrepancy to the court. Judge Castel issued an order to show cause, demanding that Schwartz produce copies of the cited opinions.

Schwartz asked ChatGPT whether the cases were real. ChatGPT confirmed they were. He submitted that exchange to the court — apparently without recognizing that asking the tool that had invented the citations to verify its own inventions was not a verification method. It was asking the counterfeiter to authenticate the bill.

The fabricated citations included Varghese v. China Southern Airlines Co., purportedly reported at 925 F.3d 1339 (11th Cir. 2019); Shaboon v. EgyptAir, supposedly at 2013 WL 12322439 (N.D. Ill.); Petersen v. Iran Air, at 905 F.2d 1424 (10th Cir. 1990); and three others with equally precise-looking but entirely fictional coordinates. The citations had volume numbers. They had page numbers. They had plausible courts and plausible years. Two even came with accurate-sounding parenthetical descriptions of holdings in cases that had never been decided because the cases had never been filed.

On June 22, 2023, Judge Castel sanctioned Schwartz, his supervising partner Peter LoDuca, and the firm jointly in the amount of $5,000. More painful than the fine was the public record. The opinion ran thirty-one pages. It reproduced Schwartz's ChatGPT chat logs. It quoted his affidavit, in which he stated he had been "unaware of the possibility that its content could be false." The order was covered by the New York Times, the Wall Street Journal, Reuters, and every legal trade publication with a website.

None of this needed to happen. The $5,000 sanction is almost beside the point. What the case reveals about how LLMs actually work — and what that means for every attorney who uses one — matters far more than the dollar amount.

Why Language Models Confabulate

The word "hallucination" has become the standard term of art, and it is mildly misleading. Hallucination implies a mind that perceives things that aren't there. An LLM is not perceiving anything. It is predicting.

A large language model is, at its structural core, a next-token prediction engine. It was trained on an enormous corpus of text — the internet, books, legal databases, code repositories, digitized libraries — and through that training it learned the statistical relationships between words, phrases, sentences, and documents. Given a prompt, it produces the continuation that, according to those statistical patterns, is most likely to follow. It does not retrieve facts from a database. It does not look up citations. It generates text that looks like the kind of text that typically follows the kind of prompt you gave it.

This distinction is not a technicality. It is everything.

When you ask a lawyer to cite authority for a proposition, the lawyer searches Westlaw, retrieves actual opinions, reads them, and reports what they say. The citation is a pointer to a real object that exists in the world. When you ask an LLM for the same thing, it generates text that is statistically consistent with what legal citation text looks like. It has seen millions of legal citations during training. It knows that citations have the structure Reporter v. Defendant, [volume number] [reporter abbreviation] [page number] ([court] [year]). It knows that Eleventh Circuit opinions cite 11th Cir. cases and that 2019 is a plausible year for such a citation. So it generates a citation in that form. Whether any actual document sits at those coordinates is simply not part of the computation.

The result is text that satisfies every surface feature of a real citation — format, specificity, authoritative tone — while pointing to nothing.

This is not a bug. It is an architectural consequence of what these systems are. A language model's job is to produce fluent, coherent, contextually appropriate text. Fluency is orthogonal to factuality. The model that invents "925 F.3d 1339" with perfect confidence is doing exactly what it was trained to do. The problem is that attorneys, reading that output, imported the cognitive shortcuts they apply to human research assistants: if it sounds that specific, someone must have verified it.

They haven't. There is no one to verify. The specificity is generated, not retrieved.

Why Legal Citations Are Particularly High-Risk Output

Not all LLM errors are equally dangerous. Ask an AI to draft a party introduction clause in a contract, and an error will likely be caught on review because you're reading prose. Ask it to suggest a tone for a client email, and the worst outcome is a bad first draft. But ask it for case citations, and the failure mode is nearly invisible — and the consequences of filing unverified citations can be severe.

Citations are dangerous AI output for three specific reasons.

First, precision masquerades as accuracy. A citation to 925 F.3d 1339 is either right or it doesn't exist. There is no partially correct citation the way there is a partially correct argument. But the level of specificity in a citation — the volume number, the exact page, the parenthetical — reads to the human brain as evidence of research. "How would it know the page number if it hadn't looked it up?" Easily. It learned the distribution of page numbers in legal reporters and generated one in range.

Second, citations are binary facts embedded in discursive text. When you read a brief and encounter a sentence like "Courts have consistently held that…, Varghese v. China Southern, 925 F.3d 1339 (11th Cir. 2019)," the citation sits in the flow of prose that you are actually reading for its argument. Your eye moves past it. You are not, in the course of reading a passage for legal logic, simultaneously running each citation against a database. The verification step has to be deliberate and separate, and it is easy to skip.

Third, the training data creates a particularly perverse problem for legal AI. LLMs have absorbed enormous quantities of legal text — law review articles, briefs, judicial opinions, bar publications. All of that text contains real citations. The model has learned not just the format but the flavor: which courts cite which types of authority, which doctrines appear in which reporters, what a plausible Montreal Convention case looks like. The fictional citations in Mata v. Avianca were not random. They were thematically coherent. They were about the right subject matter. The courts were plausible venues. The years fit the doctrinal timeline. The model had learned enough legal structure to be convincingly wrong.

The Problem Has Not Been Solved

This is where attorneys who have moved on from the Schwartz story sometimes make an error. "That was ChatGPT 3.5," the reasoning goes. "We're using GPT-4o now, or Claude 3.7, or a legal-specific platform. Newer models are better." This is partly true and entirely insufficient.

Newer models do hallucinate less frequently on certain tasks. They are better calibrated — meaning they are more likely to say "I'm not certain" when the underlying data is sparse. Some legal-specific platforms, including Lexis+ AI and Westlaw Precision, have built retrieval architectures that ground answers in specific retrieved documents before generating output, which reduces (not eliminates) fabrication. These are real improvements. None of them remove the fundamental risk.

In 2024, researchers at Stanford's RegLab and elsewhere published benchmarking studies on legal AI performance that documented persistent citation hallucination rates across multiple commercial models, including models marketed specifically for legal use. The rates varied — some platforms performed meaningfully better than others — but no model tested was hallucination-free. Zero. And the failures that slipped through shared a consistent feature: they were the most confident-sounding outputs. The ones that were obviously uncertain got flagged. The ones that were fluently, specifically, authoritatively wrong got filed.

The ABA's Standing Committee on Ethics and Professional Responsibility addressed this directly in Formal Opinion 512, issued in July 2024. The Opinion noted that generative AI tools "may generate text that appears authoritative but is factually incorrect" and stated that the attorney's competence obligation under Model Rule 1.1 requires understanding those limitations before deploying the tools in client work. The Opinion did not distinguish between consumer and enterprise AI tools, between old and new models, or between general-purpose and legal-specific platforms. The competence obligation attaches regardless of which tool you use. You are responsible for verifying what the tool produces.

Retrieval-augmented generation — the architecture most legal platforms use, in which the model is given retrieved documents to cite before generating output — helps significantly. But it does not make verification optional. Models using retrieval can still misread what they retrieved, can still construct citations that don't match the retrieved text, and can still generate analysis that misrepresents the holding of a case they were handed. The retrieved document is evidence the model was given access to something real. It is not evidence the model understood and reported it accurately.

The honest framing is this: the gap between "sounds authoritative" and "is accurate" has narrowed. It has not closed. It will not close for structural reasons that current architectures cannot resolve, because generating fluent text and retrieving verified facts are different cognitive operations.

A Verification Protocol That Actually Works

The following protocol is not complicated. It is, however, non-negotiable if you are using any AI tool for legal research.

Step one: Never file or serve a citation you have not independently verified in Westlaw or Lexis. Not Casetext. Not the AI's own "citation check" feature. Westlaw or Lexis, using the citation as a direct search query, confirming the case exists and says what you think it says. This takes ninety seconds per citation if the case is real. If the case doesn't appear in those databases, stop. Do not assume the AI was right and the database is wrong.

Step two: Read the actual opinion, not the AI's summary of it. The parenthetical the model generated may be accurate, misleading, or wrong. The only way to know is to pull the opinion and read the relevant passage. This is the step attorneys skip most often, and it is the step that catches misrepresentation even when the citation is real. AI summaries of holdings are frequently accurate in tone while being imprecise or technically

wrong about scope.

Step three: Shepardize or KeyCite everything before you rely on it. A citation can be real and still be bad law. An AI will not tell you that Miranda v. Arizona was distinguished in a controlling circuit opinion last February. You have to check.

Step four: Build the verification step into your workflow as a gate, not an afterthought. Many of the post-Mata incidents — and there have been others, though few as publicly documented — happened because attorneys were treating AI research output as a draft to clean up rather than raw material to verify. The psychology matters. If your mental frame is "ChatGPT found these cases and I'm polishing the brief," you will miss things. If your mental frame is "ChatGPT generated citation candidates that I now need to verify independently before I can use them," you will not.

Step five: If you are supervising a junior attorney or paralegal who is using AI tools, the verification obligation is yours. Model Rule 5.1 is explicit on supervisory responsibility. ABA Formal Opinion 512 reinforces it. Delegating the AI research task does not delegate the competence obligation. Peter LoDuca, who supervised Schwartz and filed the brief under his own name, was sanctioned alongside Schwartz. He had not used ChatGPT himself. He had signed the brief.

What Schwartz Got Wrong — and What He Got Right

It is easy to use Schwartz as a cautionary tale, and this chapter has done exactly that. It is worth adding nuance, because the lesson only works if you understand where the failure actually occurred.

Schwartz did not act in bad faith. Judge Castel said so explicitly. He was an attorney who encountered a new tool, found it impressive, and extended it the professional trust he would have extended to a competent research assistant. That trust was misplaced, but the instinct — to rely on confidently delivered specific information — is human and normal.

The mistake was not using ChatGPT. The mistake was treating ChatGPT output as equivalent to Westlaw output. These are not equivalent. A Westlaw result is a retrieved document that exists in a curated legal database. A ChatGPT result is generated text that matches the statistical pattern of legal research output. The interface looks similar. The epistemological status is completely different.

Schwartz's second error compounded the first: when the citations were challenged, he asked ChatGPT to confirm their validity. This reveals a misunderstanding of what ChatGPT is. It has no access to the docket. It has no memory of what it said before. It cannot verify its own outputs by checking them against reality. When it told Schwartz the cases were real, it was generating the most probable response to a question about whether it had been accurate — which, given its training on human text, is "yes, they are real." Asking the model to check its own citations is like asking the person who forged the documents to appraise them.

The Lesson, Stated Plainly

Fluency is not accuracy. Confidence is not authority. Specificity is not evidence of research.

These are principles that experienced attorneys intuitively apply to junior associates who are new to a practice area. You don't file the first-year's brief without checking the cases. The same standard applies to AI output, and it applies more rigorously, because a first-year associate at least knows when they haven't looked something up and will tell you. The model does not know what it doesn't know. It will give you six invented citations in the same confident tone it uses for six real ones.

Mata v. Avianca was a $5,000 sanction and a thirty-one page public dressing-down. For a personal injury attorney at a small firm, the reputational cost was larger than the dollar amount. For the profession, the cost was a jolt of public awareness that AI tools produce output which "sounds authoritative" and "is accurate" are phrases that can describe entirely different documents.

The case was filed in February 2022. The AI-assisted brief was filed March 1, 2023. The sanctions order issued June 22, 2023. The full opinion, Mata v. Avianca, Inc., No. 22-cv-1461, is publicly available on PACER. Read it before you use an AI tool for research. Not because it will frighten you into avoiding useful technology, but because it will give you an accurate mental model of the failure mode — and that mental model is the only thing standing between you and a similar order to show cause.

Every tool in the next chapter has a hallucination risk profile. None of them have eliminated it. Your professional responsibility does not shift based on the tool. It rests with you.

CHAPTER III

The Tools That Matter: General-Purpose vs. Legal-Specific

The market for AI tools marketed to lawyers is, charitably, confusing. In early 2026, you can choose from a few dozen products promising to transform your practice, a handful of established platforms you may already pay for, and two or three general-purpose assistants that your junior associates are almost certainly already using whether you know it or not. Every vendor claims accuracy. Every demo is polished. Almost none of the sales materials explain what actually matters for professional risk.

This chapter cuts through that. It lays out a practical framework for evaluating AI tools across five dimensions — citation grounding, retrieval architecture, hallucination mitigations, data retention policy, and pricing — then applies that framework to the tools most likely to appear in your firm's procurement conversation. No winners are declared here. What you get instead is the vocabulary to ask the right questions and the judgment to match a tool to your practice's actual risk profile.

The Fundamental Split: Generative Versus Retrieval-Grounded

Before comparing individual products, you need to understand the architectural difference that separates general-purpose AI assistants from legal-native platforms. It is the most important technical distinction in this space, and it is the one most consistently glossed over in vendor presentations.

General-purpose LLMs — GPT-4-class models from OpenAI, Claude from Anthropic, Gemini from Google — are, at their core, next-token predictors trained on enormous text corpora. When you ask one of them what the standard for summary judgment is in the Ninth Circuit, it does not look anything up. It generates a response based on statistical patterns learned during training. If those patterns encode a correct answer, you get a correct answer. If they encode a plausible-sounding but wrong answer — a citation that never existed, a standard that was overruled in 2019 — you get that instead, delivered with identical confidence.

This is not a bug that gets fixed in the next version. It is a property of the architecture.

Legal-native platforms take a different approach. They layer LLM capabilities on top of retrieval systems backed by authoritative databases — Westlaw's corpus, LexisNexis's database, case-specific document sets you upload. When Lexis+ AI or Westlaw's CoCounsel answers a legal research question, the response is grounded in documents actually retrieved from those databases, with citations pointing back to retrievable sources. The model synthesizes; the database supplies the facts. This is retrieval-augmented generation, commonly called RAG, and it is the architectural reason legal-native platforms carry meaningfully lower hallucination risk for citation-dependent tasks than their general-purpose counterparts.

Lower. Not zero. Remember that.

The Tools: A Plain-Language Inventory

General-Purpose Assistants

Claude (Anthropic) excels at long-context document analysis. The 200,000-token context window in Claude 3.5 and later versions can hold a full commercial agreement and an entire deposition transcript simultaneously. It reasons carefully through complex instructions. Its prose output is clean enough that its drafts often require less editing than comparable tools produce. What it cannot do reliably is cite caselaw. Ask Claude to identify controlling authority on a standing issue in Illinois courts and it will frequently produce citations that look correct — proper format, plausible case names, accurate-sounding procedural posture — and are fabricated. It knows this about itself. Ask it directly and it will tell you. The danger is when you do not ask, or when the attorney reviewing the output does not think to verify.

Anthropic's enterprise offering, Claude for Enterprise, runs under a data processing agreement that commits Anthropic not to train on your inputs. That distinction matters enormously for confidentiality analysis under Rule 1.6. The consumer claude.ai product does not carry the same contractual protections. Different tools. Different risk profiles. Same interface — which is exactly why the distinction is so easy to miss.

GPT-4-class models (OpenAI) — specifically GPT-4o and the o-series reasoning models — are the most widely deployed general-purpose AI in legal practice right now. The underlying models are capable, sometimes more precise than Claude on structured data extraction tasks, sometimes weaker on long legal documents with complex cross-references. The more important variable, for professional responsibility purposes, is which interface you are using. ChatGPT consumer accounts have terms that allow OpenAI to use your inputs for model improvement unless you actively opt out. The OpenAI API accessed under an enterprise agreement carries different terms. This gap matters if a paralegal is pasting excerpts from a client's acquisition agreement into a free ChatGPT tab at 10 p.m. before a deadline. That is happening right now at firms that have not issued AI use policies.

The GPT ecosystem is also the engine under most Microsoft 365 Copilot integrations. If your firm runs on Microsoft, the Copilot features embedded in Word, Outlook, and Teams are GPT-based. They inherit Microsoft's data governance commitments under your existing enterprise agreement, which is generally favorable — but you still need to confirm that client data processed through Copilot does not flow to Microsoft's training pipelines. Reading the enterprise terms, or having IT confirm the configuration, is not optional.

Legal-Native Platforms

Harvey is the most discussed legal AI startup entering 2026. Founded in 2022 by Winston Weinberg and Gabriel Pereyra, it raised capital at a reported $1.5 billion valuation and counts firms including A&O Shearman, Linklaters, and PricewaterhouseCoopers Legal among its clients. Harvey is not a model in the

base sense — it is an application layer built on top of frontier models (primarily GPT-4-class and Claude, depending on the task) combined with firm-specific retrieval and legal-domain fine-tuning.

What Harvey offers over raw general-purpose access is workflow integration and, critically, a data processing agreement structured for law firm use. Harvey has been explicit in its enterprise contracts that client matter data is not used to train models. That is the baseline requirement for any legal AI platform worth deploying, but confirm it in writing every time. Every legal-native competitor says the same thing. "Says it" and "commits to it contractually with indemnification provisions" are different things.

Harvey's citation grounding depends heavily on your configuration. For firms that upload their own precedent libraries or connect Harvey to live Westlaw or Lexis feeds, outputs are meaningfully grounded. For firms using Harvey against a blank retrieval setup and asking legal research questions, they are not. Harvey's team will tell you this openly — but it is the kind of detail that gets lost between the sales demo and the associate working through a research memo at midnight.

Lexis+ AI is LexisNexis's integrated AI product, available within the Lexis+ subscription tier. Its architecture retrieves from the Lexis database before generating responses, which means citations in research queries are sourced from Lexis-indexed cases, statutes, and secondary sources rather than confabulated from training data. LexisNexis has published internal comparisons claiming materially lower fabrication rates on citation-dependent queries against general-purpose models. Independent audits of those claims are limited in number. Treat the vendor comparisons as directionally informative rather than precise.

The practical advantage of Lexis+ AI is that it lives inside the same interface many practitioners already use for research. Workflow integration is underrated in the AI conversation. A tool you use because it is already open in your browser is functionally superior to a theoretically better tool that requires a separate login, a separate window, and a separate habit to form. Friction kills adoption, and unused tools protect no one.

Pricing runs roughly $200–$250 per month per attorney for Lexis+ subscriptions inclusive of AI features, though firm-negotiated rates vary substantially. Solo practitioners and small firms should ask specifically whether AI features are included in their current subscription tier before assuming access.

Westlaw Precision and CoCounsel are now effectively one platform, following Thomson Reuters' acquisition of Casetext in August 2023 for $650 million — one of the largest deals in legal tech history. CoCounsel, originally Casetext's AI product, has been integrated into the Westlaw ecosystem. The product answers legal research questions with citations grounded in the Westlaw database, supports document review and contract analysis, and can run deposition preparation and contract comparison workflows.

Thomson Reuters has stated in its enterprise agreements that CoCounsel does not train on client data. The retrieval architecture grounds citation outputs in the Westlaw corpus. The interface is more polished than the early Casetext days, though some practitioners who used the original CoCounsel product note that certain workflow steps added layers of interface complexity in the migration.

Pricing for Westlaw Precision with CoCounsel starts around $600 per month per seat for small-firm arrangements, with volume discounts at larger firms. That price point effectively excludes solo practitioners

unless you can negotiate access through a bar association arrangement. Several state bars, including the New York State Bar Association, have negotiated member-rate access to major research platforms. Checking whether your bar offers similar arrangements is worth thirty minutes of research before you assume you cannot afford it.

The Five Dimensions: A Practical Scorecard

When evaluating any AI tool for legal work — not just the platforms named above but anything a vendor puts in front of you over the next 18 months — measure it against these five questions.

Citation Grounding. Does the tool retrieve citations from a verified, curated database, or generate them from training data? This is the single most critical question for litigation and research work. Ask the vendor directly: where do citations in research outputs come from? A retrieval-grounded tool will describe a specific database integration and a linking mechanism back to source documents. A generation-first tool will say something about the model being trained on "extensive legal data" or having "access to legal information." Those phrases describe different architectures. One is materially safer for citation work.

Even retrieval-grounded tools occasionally confuse documents, surface outdated authority, or generate language that drifts from the retrieved source. The verification obligation does not disappear. But the baseline error rate is meaningfully different, and that gap has professional responsibility consequences.

Retrieval Architecture. The quality of a RAG system depends on three things: the quality of the underlying database, the precision of the retrieval mechanism, and whether the model actually stays tethered to what was retrieved rather than wandering into generative territory mid-response. Ask the vendor: is the retrieval database updated in real time, or on a lag? Lexis and Westlaw update their primary source databases continuously. A legal-native platform operating on a static snapshot of caselaw that is six months old is not a research tool. It is an expensive reference book.

Hallucination Mitigations. No tool eliminates hallucination. What separates platforms is how they structure the risk. Look for: citation sourcing with direct hyperlinks to underlying documents; confidence indicators that flag uncertain outputs; stated behavior when a query falls outside the retrieval scope. Does the tool refuse to speculate, or does it generate anyway? Harvey has built explicit handling for out-of-scope queries. Whether those guardrails hold under adversarial prompting is a separate question — and one worth testing before deployment.

Some vendors publish independent accuracy evaluations. Thomson Reuters has commissioned studies on CoCounsel accuracy. Read them as marketing materials that contain useful data, not as neutral audits. The Mata v. Avianca sanctions — an attorney filing ChatGPT-fabricated citations in federal court in 2023, covered in Chapter 2 — arose from output that sounded authoritative. Sounding authoritative is what LLMs do. It is not evidence of accuracy.

Data Retention Policy. This dimension governs your confidentiality exposure directly. Under Model Rule 1.6, you must make reasonable efforts to prevent unauthorized disclosure of client information. The analysis

breaks into three tiers. Consumer tools — free ChatGPT, consumer claude.ai, Google Gemini without a Workspace agreement — carry the highest risk. Their terms may permit training on your inputs, and those terms change without prominent notice. Enterprise API access under a signed data processing agreement, from any major provider, carries meaningfully lower risk: the vendor commits not to train on your inputs and provides contractual grounding for your reasonable-efforts defense. On-premise or firm-hosted deployments carry the lowest third-party disclosure risk because client data never leaves your infrastructure. The tradeoff is cost, IT burden, and capability — self-hosted open models as of 2026 still trail frontier commercial models on complex multi-step legal reasoning.

ABA Formal Opinion 512, the most substantive AI ethics opinion the bar has produced, requires that competent use of AI tools include "an understanding of what happens to information you input." That standard is not satisfied by assuming the tool is private. You need to read the terms, or have someone read them for you.

Pricing and Total Cost. Published list prices are not what large firms pay. Per-seat pricing for any platform you plan to deploy across 20 or more attorneys is negotiable; the rates on the vendor website are ceilings. What matters for realistic budget planning is the all-in cost across three buckets: the tool itself, integration work (connecting it to your document management system, your existing Westlaw or Lexis accounts, your conflicts platform), and attorney and paralegal training time. A tool that is cheap per seat but requires eight hours of onboarding per person is more expensive than it looks.

How to Choose

There is no universal right answer. That is not a hedge — it is an accurate description of a market where the relevant variable is your practice type.

If you are a litigator at a firm with an active federal docket, CoCounsel or Lexis+ AI should be your primary research interface. The citation grounding and database currency justify the cost at scale. General-purpose tools remain useful for drafting, for argument structuring, for translating a complex procedural history into a client-readable summary — not for the authoritative research you file.

If you are a transactional attorney focused on contract review and drafting, Harvey or a well-configured Claude or GPT-4 deployment with your firm's precedent library as retrieval context may serve you better than a litigation-focused research platform. The risk profile for contract drafting is different: you are not filing citations with a court, you are generating language that your counterparty will redline and you will review again. General-purpose tools handle contract drafting tasks respectably. The verification burden shifts from citation accuracy to substantive legal judgment — which was always the attorney's job anyway.

If you are a solo practitioner or a two-attorney shop, the economics of legal-native platforms are difficult. A defensible configuration: Claude for Enterprise at approximately $30 per month per seat under the Teams plan, or GPT-4 access via API, combined with a bar-association-subsidized Fastcase subscription for citation verification, plus a personal protocol that every citation gets independently confirmed before it leaves your

office. That workflow is not the same as CoCounsel. But it is professionally defensible if the verification step is real and consistently applied.

What the Bar Actually Requires

Bar guidance on tool selection is thinner than practitioners need, but it is not absent. ABA Formal Opinion 512 identifies competent AI use as requiring understanding of the benefits and risks of the tools you deploy. Not how transformer architectures work. How the tool behaves, where it is known to fail, and what your workflow does to catch those failures before they reach a court or a client.

The New York State Bar Association's AI task force report, issued in early 2024, went further, recommending that attorneys understand "at a functional level the distinction between retrieval-augmented and purely generative outputs." California's State Bar AI guidance calls out citation verification as a specific competence obligation rather than a general aspiration. The Florida Bar has emphasized supervision obligations when AI tools are used by non-lawyer staff. These documents are readable. The NYSBA report runs 84 pages, which sounds long, but the operative guidance fits in roughly 15.

The practical compliance posture requires three things: documented rationale for the tools you select, a written AI use policy that specifies which platforms are approved for client matter work and which are not, and verification checkpoints built into any workflow producing citable outputs. The policy does not need to be a treatise. It needs to be real, it needs to be followed, and it needs to address the specific tools your attorneys are actually using — not the ones you think they are using.

A Final Calibration

The tool conversation in legal circles tends to drift toward capability benchmarks — which model scores highest on simulated bar exam questions, which platform produced the most accurate research summaries in a vendor-commissioned study. Those numbers are not useless. But the right frame for a practicing attorney is not which tool is smartest. It is which tool fails in ways you can manage before the failure reaches a client, a court, or opposing counsel.

General-purpose tools fail by confabulating with confidence. Legal-native platforms fail by retrieving the wrong document, surfacing outdated authority, or generating language that drifts subtly from what the retrieved source actually says. Both failure modes require human review to catch. The difference lies in probability and type of error — and in how visible the failure is before it does damage.

You are not looking for a tool that never fails. You are looking for a tool whose failures surface before they matter. Carry that question into every vendor demonstration: show me where this breaks, and show me how I find out in time to fix it. If the sales team cannot answer concretely, with examples and a description of the error mode, the tool is not ready for professional deployment. Or the sales team is not being straight with you. Either way, walk away.

CHAPTER IV

Drafting: Contracts, Briefs, and Memos

The blank page has always been the enemy. Ask any lawyer who has billed four hours to produce a two-page letter — most of that time staring at the cursor — and you understand why AI drafting tools spread through law firms faster than any prior legal technology. The speed gain is real. Genuinely, measurably real. A skilled attorney using a well-prompted AI assistant can produce a competent first draft of a mid-complexity NDA in under twenty minutes; the same document might have taken ninety minutes cold.

But here is what that sentence buries: first draft. And competent. Not final. Not verified. Not filed. The productivity unlocked by AI drafting tools belongs entirely to the revision stage that follows. If you skip that stage — if you treat AI output as a finished product — you have not saved time. You have borrowed time from your client's matter and hidden the interest payment inside a document that may misstate the law, omit a required provision, or reflect a jurisdiction you did not actually specify.

This chapter is about how to use AI to make drafting faster without making it worse. That means step-by-step workflows for each of the three document types lawyers draft most — transactional contracts, litigation briefs, and internal research memos — with annotated prompt templates you can adapt to your practice. It also means being explicit about the division of labor: what AI can supply, and what only you can.

The Division of Labor: What Only You Can Supply

Before a single word of AI output touches a client matter, the attorney must answer four questions. No AI tool can answer them for you, and the quality of everything downstream depends on your getting them right.

Jurisdiction and governing law. An AI model trained on American legal text will confidently generate contract provisions, reference Model Act defaults, and cite statutory schemes — without knowing whether your deal is governed by Delaware law, Texas law, or the law of a country that uses a civil code entirely. You must specify this. Explicitly. In the prompt.

The client's actual facts. AI generates plausible text. Plausible is not the same as accurate to your client's situation. The AI does not know that your seller has a consent right buried in a prior preferred stock agreement, or that your defendant was in a different city on the date the complaint alleges. You do. Those facts belong in the prompt, supplied by you.

The strategic posture. Contracts are negotiated documents. Briefs argue a position. Even memos reflect a framing. AI will produce balanced, anodyne language unless you tell it whose interests you are protecting and how aggressively. "Draft an NDA favoring the disclosing party, with a broad definition of Confidential Information and no carve-out for independently developed information" produces a different — and more useful — document than "draft an NDA." The strategy is yours.

What the document must not say. AI has no knowledge of prior representations, pending regulatory matters, privileged conversations, or the specific factual record in your case. It can easily generate language that contradicts a prior filing, breaches a standstill agreement, or waives an argument you are preserving for appeal. You are the only person in the room who knows what the document cannot contain.

ABA Formal Opinion 512, issued in August 2024, treats AI-assisted work product as materially identical, for supervision purposes, to work delegated to a junior associate or paralegal. The opinion states that "the lawyer's ultimate professional responsibility for the work product cannot be delegated." That is not a formality. It means every clause in the output is yours, professionally, before you deliver it.

Contracts

Transactional drafting is where AI assistance makes the most intuitive sense, and where the failure modes are also most predictable. Take a mutual NDA as the working example — not because it is the hardest document you will draft, but because it is common enough that you have a calibrated sense of what correct output looks like.

Step 1: Build your prompt with all four inputs.

Do not open Claude, Harvey, ChatGPT, or any other tool and type "draft an NDA." That instruction produces a generic document reflecting the average of every NDA in the training corpus, governed by no particular state, protecting no particular party's priorities. It will look like an NDA. It will read like an NDA. It may be missing a tolling provision, a specific injunctive-relief clause, or a return-of-materials obligation your client requires. You won't know until you compare it against your checklist.

Here is an annotated prompt that works:

> *"Draft a mutual non-disclosure agreement governed by Delaware law. The parties are a seed-stage software company (the "Company") and a strategic investor conducting preliminary diligence. The Company is the primary disclosing party; draft the agreement to favor the Company's interests. Include: (1) a broad definition of Confidential Information capturing source code, customer data, and financial projections; (2) a two-year confidentiality period post-disclosure; (3) an explicit provision that independently developed information is NOT excluded from the confidentiality obligation — note this is an aggressive position, flag it for negotiation; (4) a specific performance and injunctive relief clause with no bond requirement; (5) Delaware as exclusive jurisdiction. Format as a clean draft with defined terms in bold."*

Notice what that prompt does. It names the jurisdiction. It identifies the parties and their asymmetric positions. It specifies the strategic posture. It calls out the aggressive provision and flags it as a negotiation point — because you want to know when AI has generated something that may not survive counterparty review.

Step 2: Review the output against a checklist, not just your memory.

Every practice area has a set of provisions that should appear in a given document type. Before AI, you probably had this checklist in your head or in a firm standard. Now, you need it on paper — because the cognitive shortcut of "this looks right" is exactly what AI output exploits. Something can look right and be wrong. The NDA checklist for a technology company might include: definition of Confidential Information, exclusions, permitted disclosures, return-and-destruction obligations, residuals clause (or explicit absence of one), term and survival, injunctive relief, governing law, and dispute resolution. Work through the list line by line.

Step 3: Redline for substance, not just style.

Most attorneys receive AI output, read it once for flow, fix a few clunky phrases, and call it done. That is proofreading, not revision. Real review means asking, for each substantive provision: is this the correct standard for this jurisdiction? Does this reflect my client's actual agreement? Could this clause be read in a way that harms my client? The redline pass is professional work. It is not optional.

On redline integration: if you are working in Microsoft Word, the Compare Documents function lets you baseline an AI draft against a prior firm precedent and see exactly what changed. This is faster than reading for differences line by line. Use it. Legal-specific platforms — Harvey, Ironclad's AI tools, ContractPodAi — offer native redlining with version control, which is worth the additional cost in high-volume transactional practices. The workflow should be: AI draft → compare against precedent → su client-specific fill-in → final attorney sign-off. Five steps, none of them opti

Step 4: Fill in what AI cannot generate.

Specific deal economics, party-specific representations and warranties, schedules, and exhibits all require your input. AI can scaffold; you populate. An AI-generated Stock Purchase Agreement that contains "[REPRESENTATIONS AND WARRANTIES TO BE ADDED]" in brackets is not a draft. It is an outline. Your job is to know the difference, and to know that the bracketed language is not the only thing that needs your attention.

Briefs

Litigation briefs present a different problem than contracts. Contracts involve negotiated language. Briefs make representations to a court — and under Rule 3.3, those representations must be truthful and must not involve the submission of false statements of law. The stakes for AI error are higher.

The Mata v. Avianca post-mortem belongs in Chapter 2, but the operational lesson belongs here. Steven Schwartz used ChatGPT to generate case citations supporting a personal injury plaintiff's opposition to a motion to dismiss. The citations — Varghese v. China Southern Airlines, Shaboon v. Egypt Air, Zicherman v. Korean Air Lines, among others — were fabrications. Schwartz filed them in the Southern District of New York in May 2023. When opposing counsel could not locate the cases, Judge P. Kevin Castel ordered briefing.

Schwartz submitted a declaration stating he had not understood that ChatGPT could generate false information. He and his supervising partner were each sanctioned $5,000. The client's case was dismissed.

What produced that outcome was not negligence in the traditional sense. It was a categorical misunderstanding of what a language model does. You will not repeat that specific mistake. But the principle generalizes: AI will produce case citations, statutory quotations, and holdings that are wrong in ways that are not immediately detectable. Every case cite generated by an AI tool must be verified in Westlaw or Lexis before the brief is filed. Every time. No exceptions. Not "usually." Not "for citations I don't recognize." Every one.

With that baseline fixed, here is the brief-drafting workflow.

Step 1: Give AI the record, the standard, and your theory — explicitly.

> *"I am drafting the argument section of an opposition to a motion for summary judgment in a breach of contract action in the Southern District of New York. Governing law is New York. The movant argues the implied covenant of good faith and fair dealing does not apply because an express provision covers the disputed conduct. My theory is that the implied covenant can supplement but not contradict an express term, and that the movant's conduct breached both independently. Draft a 1,500-word argument section. Use [CASE CITATION] placeholders wherever you would cite a case — I will verify and insert actual citations. Do not generate case names or docket numbers. Structure the argument: (a) the legal standard for the implied covenant under New York law, (b) application to our facts, (c) rejection of the movant's argument."*

The instruction to use placeholders is deliberate. It forces the structural work AI does well — organizing a legal argument — while preventing the hallucination risk AI does worst. Some attorneys prefer to let AI suggest case names and then verify each one; others prefer placeholders throughout. Either is defensible. What is not defensible is accepting AI-generated citations without verification.

Step 2: Supply record excerpts directly in the prompt.

This is where brief-writing AI assistance becomes genuinely powerful. Paste the relevant deposition excerpts, the key contract provision, and the movant's brief into the context window, and the AI can draft argument sections grounded in specific record evidence rather than generic legal principles. Three pages of record excerpts in your prompt saves an hour of writing. The output will still require verification, but it will be orders of magnitude closer to a usable first draft than anything generated without that grounding.

Step 3: Fact-check the legal standard against primary sources before filing.

AI will state the summary judgment standard, the pleading standard, the abuse-of-discretion standard — all with confidence. Some will be correct. Some will reflect circuit splits, superseded doctrine, or standards from the wrong court. Before you file, pull the current standard from a case decided within the past three years in your specific court. If AI's articulation matches, you have confirmed it quickly. If it doesn't, you have caught an error before it became a Rule 3.3 problem.

Step 4: Apply judge-specific knowledge in the revision pass.

AI does not know your judge's preferences, prior rulings, aversion to block quotations, or tolerance for forty-page briefs. That knowledge is yours. It is worth more than any AI efficiency gain. Apply it.

Memos

Internal research memos are where AI drafting is least risky and most consistently useful. The audience is internal. The memo is not filed with a court and is not delivered to a counterparty as a representation. It is a working document, and errors can be caught before they propagate into advice or filings.

This does not mean memos require no verification. It means the stakes of an uncaught error are lower than in a brief or a contract, and the workflow can be calibrated accordingly. You are not cutting corners. You are right-sizing the review effort to the risk profile.

The memo prompt template:

> *"Draft a research memo analyzing the enforceability of non-compete agreements under California law for a B2B SaaS company with employees in San Francisco. Format: Question Presented, Brief Answer, Applicable Law [use [CITATION] placeholders throughout], Analysis, Conclusion. Include the current statutory framework under Business and Professions Code §16600 and note any material amendments enacted in 2023 or 2024. Note any unsettled areas or factual variables that could affect the outcome. Assume the employee signed a two-year non-compete at hire in 2022."*

The explicit instruction to flag unsettled areas matters. AI tends to present the law as more settled than it is — partly because training data skews toward secondary sources that smooth over uncertainty, and partly because confident prose is statistically rewarded in text generation. Asking for flags on contested doctrine nudges the model toward useful epistemic honesty.

What the AI-generated memo gives you:

A structured outline of the relevant doctrine, organized in memo form, with placeholders for citations you will verify. It surfaces secondary issues you might not have flagged immediately — a memo on California non-competes might identify the Business and Professions Code §16600 line of cases, the specific treatment of sale-of-business exceptions, and the extraterritorial application questions raised when California employees sign agreements under another state's governing law. Some of those issues will send you to Westlaw for deeper research. Good. That is the workflow functioning as designed.

What the memo does not give you:

Current law. The law on California non-competes changed materially in 2023 and 2024. SB 699, effective January 1, 2024, voided non-competes regardless of where they were signed, including agreements with choice-of-law clauses selecting other states. AB 2288, passed the same year, imposed notice requirements on employers. An AI with a training cutoff prior to those enactments will state the prior rule with confidence. You will not know it is wrong unless you check. Check.

Prompt Discipline as Professional Habit

Three rules. Stated plainly.

Your prompt is a professional document. The quality of your instructions determines the quality of the output. Vague prompts produce vague drafts. Specific prompts — jurisdiction named, party identified, strategic posture explicit, prohibited content flagged — produce specific, usable drafts. Spend three minutes writing a precise prompt rather than thirty seconds writing a lazy one. You will recover that time in the revision pass, easily, and the output will be better.

Never paste client-identifying information into a consumer AI product. Chapter 12 covers Rule 1.6 in full, but the drafting-workflow version of the rule is simple: if the AI tool operates without an enterprise data-processing agreement, use anonymized or hypothetical facts in your prompt. "A Delaware corporation selling its software business" is sufficient. "[Client Name], Inc., the seller in the [matter name] acquisition" is a confidentiality breach waiting to be discovered on a data-retention audit. The legal question does not require the identifying information. Don't include it.

Document your AI use in the matter file. ABA Formal Opinion 512 does not mandate disclosure to courts or clients in every instance, but it does require competent supervision and the ability to demonstrate that the work product was reviewed. A one-line note — "First draft of contract generated via [tool] on [date]; reviewed and revised by [attorney] on [date]" — is both a professional record and a malpractice defense. It takes thirty seconds. Make it a habit before the obligation becomes something your malpractice carrier argues about in discovery.

The Final Draft Is Yours

Every representation in a contract, brief, or memo that passes through your hands is your professional product. Not the AI's. The AI has no license, no client relationship, no professional obligations, and no accountability. You have all four.

ABA Formal Opinion 512 states that "a lawyer using generative AI must ensure the work product is accurate, complete, and not misleading." That is not a standard the AI can meet. It is a standard you must apply, after the AI has done the scaffolding.

The attorneys sanctioned in Mata v. Avianca were not sanctioned because they used a new tool. They were sanctioned because they filed documents containing false statements of fact and failed to correct those statements when challenged. The tool was incidental. The professional obligation was not.

AI drafting assistance, used correctly, is a genuine productivity multiplier. The attorney who builds a solid prompt, reviews output against a substantive checklist, verifies every citation, and revises for jurisdiction and client-specific accuracy will produce better first drafts faster than the attorney who starts cold. That is the legitimate value here. Not magic. Not replacement. A faster path to the revision that has always been your

job.

Use the tools. Then do your job.

CHAPTER V

Legal Research with AI: Speed, Gaps, and the Verification Imperative

The pitch arrives in a hundred different forms. "Just ask it what the standard is." "Have the AI pull the cases on this." "Why spend four hours in Westlaw when you can get an answer in thirty seconds?"

None of that advice is wrong, exactly. It's incomplete in ways that can cost you your bar card.

AI is genuinely useful in legal research. It is faster than a first-year associate at generating issue lists, faster than most paralegals at summarizing secondary sources, and surprisingly capable at identifying the contours of a legal question — the doctrinal landscape, the competing tests, the circuit tensions — before you go hunting in authoritative databases. That is real value. Use it.

But here is the thing the pitch omits: the same model that maps the contours of a legal question with elegant fluency will, with similar fluency, hand you a citation that does not exist. Or one that exists but says nothing like what the model claims. Or a case reversed, distinguished into irrelevance, or superseded by statute before the model's training data was assembled. The prose is identical in both instances. The model's confidence is identical. You cannot tell the difference from the output alone.

That is the research problem in a single paragraph. What follows is a workflow for capturing the genuine speed advantages of AI while building in the verification steps that professional responsibility demands.

What AI Actually Does Well

Start with a clear inventory. Not everything in legal research carries the same risk profile. Some tasks are genuinely well-suited to current LLMs, and grinding through manual processes where AI performs reliably is its own kind of inefficiency.

Issue spotting. This is where AI earns its keep. Given a dense fact pattern, a well-prompted LLM will surface potential claims, defenses, and doctrinal theories faster than most humans. Not because it knows the law — it does not — but because it has processed an enormous volume of legal text and can pattern-match a set of facts against common legal frameworks. Give it your client's situation. Ask what issues a cautious plaintiff's attorney would identify. Then ask what defense counsel would argue in response. A good model generates a list that takes a competent associate ninety minutes to replicate from scratch. You'll get there in five.

The critical caveat: treat everything it generates as a hypothesis, not a conclusion. The issue list is a starting point for your research. If the AI identifies a potential RICO theory in a commercial fraud matter, that is your signal to go look at the actual elements, the circuit authority, the pleading requirements. The model does not verify its own suggestions. Neither should you take them on faith.

Secondary-source synthesis. Ask an LLM to explain the current state of the economic loss doctrine, and you will get a readable, reasonably accurate summary of the general framework. Same with anticipatory repudiation, equitable estoppel, or the standards for preliminary injunctions. These are areas of stable, heavily published law that are well-represented in any model's training data. The summary is roughly equivalent to a Nutshell guide — useful for orientation, not for citation.

The operational use: when a partner hands you a research assignment in an unfamiliar area, run a secondary-source query in AI before you open Westlaw. Get your bearings. Learn the terminology. Understand the structural framework. You will be a faster, better researcher in the database afterward because you know what you are looking for.

Jurisdiction survey work. Suppose you need to know, quickly, how thirty states treat non-compete agreements in the wake of the FTC's 2024 rulemaking. An LLM can produce a rough-and-ready state-by-state summary in minutes. That summary will have errors — some states will be mislabeled, some recent legislative changes will be absent, a few will reflect legal positions that were accurate in 2022 but not today. But as a research framework — a list of which states you need to verify, organized by legal regime — it is dramatically faster than starting cold. The AI gives you the map. You still have to walk the territory.

Where AI Fails, and Why It Matters More in Law

Now the hard part.

LLMs are not retrieval systems. When you type a research query into ChatGPT, GPT-4o, or Claude, the model is not searching a legal database. It is not querying Westlaw or Lexis or any authoritative source. It is generating text that fits your query based on statistical patterns learned during training. The text sounds like legal research because legal research is well-represented in its training data. The citations look real because real citations were in its training data.

Generation versus retrieval. That distinction is the core of the hallucination problem, and it matters more in legal practice than in almost any other professional context.

In most fields, a plausible-but-wrong answer is a problem. In law, a plausible-but-wrong case citation filed in federal court is a sanctions event. We know this because it happened. Steven Schwartz, an attorney at Levidow, Levidow & Oberman in New York, submitted a brief in Mata v. Avianca (S.D.N.Y. 2023) containing citations to cases fabricated wholesale by ChatGPT. Varghese v. China Southern Airlines. Shaboon v. Egypt Air. Petersen v. Iran Air. None of them existed. Judge P. Kevin Castel's sanctions order, issued June 22, 2023, runs to thirty-one pages. Chapter 2 of this book covers the full post-mortem. The point here is narrower: if hallucination were merely a curiosity, Mata would be a footnote. It is the defining event in AI legal research so far. It happened because counsel trusted generation for a task that required retrieval.

The problem has not been solved by newer models. This bears repeating because vendors frequently imply otherwise. GPT-4o is more accurate than GPT-3.5. Claude 3.5 and later have better factual grounding than earlier versions. Legal-native platforms have built retrieval-augmented generation (RAG) architectures

specifically to reduce citation hallucination. These are genuine improvements. Not solutions. Research examining LLM performance on legal citation tasks has consistently found that even strong general-purpose models produce erroneous or fabricated citations at rates that would be professionally unacceptable in adversarial filings. RAG-based legal tools perform measurably better. No model reaches a reliability threshold that substitutes for human verification.

There is a second failure mode that gets less attention: currency. A model trained on data through, say, October 2024 does not know about a Supreme Court decision from March 2025. It does not know about the statutory amendment your state legislature passed last session. It does not know the regulatory guidance your agency issued last quarter. It will describe the law as it existed at some point in its training window — and do so in present tense, with no indication the landscape has shifted. For fast-moving areas — employment law, immigration, antitrust enforcement, tax regulations, anything touched by recent rulemaking — this is particularly treacherous.

Third failure mode: completeness. Even within its training window, a model's coverage of legal authority is uneven. Circuit splits may be underrepresented. Recent district court opinions are often absent. Slip opinions from last year? Frequently not. Unpublished decisions? Hit or miss. A model may accurately describe the prevailing rule in the Ninth Circuit while being entirely unaware of an en banc decision from the Sixth that cuts the other way. You would not know unless you looked.

The Sequenced Workflow

Given all of that, here is the framework that works.

The core principle: AI before databases, never instead of them. Use AI to accelerate the early, generative phase of research — issue identification, framework understanding, search-term development. Then confirm everything that matters in Westlaw, Lexis, or a legal-native AI platform before anything goes into a document. This is not a reluctant concession to AI's limitations. It is a workflow that is faster than either pure-AI or pure-database research, because each stage does what it does best.

Step 1: Frame the legal question with AI.

Before opening a database, describe your research problem to an LLM in plain language. Include the jurisdiction. Include the procedural posture. Include the governing law if you know it. Ask the model to identify the core legal issues, the controlling doctrines, and the key terms of art for authoritative searches. A prompt that works:

> *"My client is a commercial landlord in Illinois seeking to enforce a lease guarantee against a corporate guarantor whose principal entity filed Chapter 11 three months ago. The lease is a triple-net retail lease; the guarantee is a 'good guy' guarantee with a carve-out for bankruptcy. Identify the primary legal issues, the relevant doctrines, and the search terms I should use in Westlaw."*

The model's response is not your research. But it will identify quickly that you are dealing with issues at the intersection of landlord-tenant law, the bankruptcy automatic stay under 11 U.S.C. § 362, guarantor liability under state contract law, and potentially § 365 lease assumption questions. It gives you search terms. It may surface the Winn-Dixie retail bankruptcy cases as relevant context. You now have a structured research roadmap, built in five minutes instead of thirty.

Step 2: Use AI for secondary-source orientation.

Ask the model to explain the relevant doctrines in plain language. This step is optional for experienced practitioners in their core practice areas — you know the framework already. It matters most when you're stretching into adjacent territory or working under time pressure that doesn't allow proper secondary-source reading.

Step 3: Build your verification list from the AI output.

If the model mentions any specific cases, statutes, regulations, or secondary sources by name — write them down. Every single one becomes an item for Step 5. Do not use these citations in your work product before verification. Not in a memo. Not in a footnote. Not in an email to a partner. The model's citations are hypotheses.

Step 4: Conduct authoritative database research.

Now open Westlaw or Lexis. The AI-generated framework from Step 1 tells you what to look for; the database tells you what is actually there. Run your searches. Find the controlling authority. Shepardize or KeyCite everything. Check the current statutory text. For jurisdiction surveys, Lexis+ AI and Westlaw Precision are meaningfully better than general-purpose LLMs because they are built on grounded retrieval — the model's answers are tied to actual database content, and citations link to real documents. You can still get errors. RAG is not a hallucination cure. But the rate drops substantially, and you have a document to verify against. Spot-check a random sample of citations against the underlying source text before you rely on any survey output.

Step 5: Verify everything the AI said.

Return to your Step 3 list. Look up every case the model mentioned. Does it exist? Does it say what the model claimed? Is it still good law? Does it come from the right jurisdiction and court level for your purpose?

This step will occasionally produce a pleasant surprise — the AI mentioned a real, useful case you would have found anyway. More often, you will find at least one case that is misdescribed, one that the model characterized as binding when it is only persuasive, and — with non-trivial frequency — one that does not appear in Westlaw at all. That last category is the Mata category. Check it every time.

Prompting for Research Tasks

A few prompt architectures that produce better research outputs:

The skeptical respondent. After the model gives you an analysis, follow up: "What counterarguments would a sophisticated opposing counsel make to each of these points? Which argument is weakest if pushed hard in litigation?" This surfaces analytical gaps faster than re-reading the first response. Good attorneys do this with junior associates. Do it with AI too.

The jurisdiction-specific query. Always specify jurisdiction. "What is the law on a seller's duty to disclose latent defects?" is a worse prompt than "What is the duty-to-disclose standard for sellers of residential real estate in California, and has the Shapiro v. Sutherland line of cases been modified by subsequent legislation?" The first produces a generic national summary. The second forces engagement with a specific legal regime — and the specificity makes errors easier to catch, because you know enough to recognize inconsistencies with your actual experience of California law.

The gap analysis. After building your verified research results in Westlaw, bring the synthesized authority back to the AI and ask: "Here are the cases and statutes I've found on this issue. What gaps do you see? What arguments does this authority not address that opposing counsel might raise?" The model is now performing analysis on a foundation you have verified. This is one of the most productive AI applications in legal research, and it is low-risk precisely because you control the source material.

The plain-language test. Before finalizing a research memo, ask the model: "Here is a draft memo on this issue. Does the legal analysis have any internal inconsistencies? Are there doctrines I've cited that are in tension with each other?" This is an editorial check, not a citation check. The model reviews your work; it is not generating its own authority. Low risk, genuine value.

The Verification Imperative in Professional Terms

ABA Formal Opinion 512, issued July 29, 2024, addresses generative AI directly under Rule 1.1's competence standard. The opinion does not prohibit AI use. What it requires — and the language is plain — is that a lawyer using AI must understand "the benefits and risks" of the technology sufficiently to deploy it competently. That understanding includes knowing that AI-generated legal citations require human verification before appearing in any work product. Under Rule 3.3's candor obligation, submitting fabricated citations to a tribunal is a disciplinary event, as Mata demonstrated at length.

The New York State Bar Association's 2024 AI Task Force report emphasized the same point: reliance on AI for legal research without human verification does not satisfy the professional standard of competent representation. California's State Bar guidance from the same period was more granular, recommending written firm protocols for AI use — including explicit verification requirements for any AI-generated legal authority before it appears in client-facing work. The Florida Bar has issued parallel guidance along the same lines.

None of this guidance prohibits AI. All of it assumes verification as a non-negotiable step. That consistency across bar associations is not coincidental. For once, the bar has landed on a standard that actually tracks the technology's documented limitations.

A Note on Legal-Native Platforms

Harvey, Lexis+ AI, and Westlaw Precision occupy meaningfully different ground from general-purpose assistants, and their architecture matters specifically for research. Each uses some form of RAG — the model's output is anchored to retrieved documents from a curated legal database rather than generated purely from training weights. This reduces, but does not eliminate, hallucination. More importantly, it gives you something concrete to check: a cited document in Lexis+ AI is a real document in the Lexis database, which you can open and read against the model's characterization of it.

Chapter 3 covers the comparative architecture of these platforms in detail. For research purposes, the practical rule is this: use a legal-native platform with grounded retrieval for any research task where authority will be cited in work product. Use general-purpose models only for the framing, issue-spotting, and analytical tasks in the first two steps of the workflow above. The line is not about brand loyalty. It is about whether the model's citations are tethered to a database you can independently verify.

The Time Math

Here is what the sequenced workflow actually buys you.

An experienced associate at a mid-size litigation firm might spend four to six hours on a complex research question — building the framework, running searches, pulling cases, shepardizing, synthesizing. The same workflow with AI handling Step 1 and Step 2 compresses that to two to three hours on the database and verification work, because the framing and orientation phases are largely handled. That compression is real. It is significant on a per-matter basis. Across a busy practice, it adds up fast.

What it is not: a shortcut past verification.

The attorneys in Mata thought they had found a shortcut. They saved the time they would have spent in Westlaw. They then spent months in sanctions briefing, appeared at a show-cause hearing before a federal judge, paid financial penalties, and had their conduct described in a published opinion that will follow them indefinitely. The math on that trade is not close.

Verification takes time. Sometimes thirty minutes. On a complex research question with ten or twelve AI-surfaced authorities, longer. Budget for it. Build it into your workflow explicitly, not as an afterthought. The AI handles hypothesis generation. You handle proof.

That division of labor — creative and fast on one side, rigorous and final on the other — is how professional judgment has always worked. It happens to be the only version of AI-assisted legal research that is compatible with keeping your license.

CHAPTER VI

Discovery and Document Review: Where AI Pays for Itself

The economics are brutal. In large commercial litigation, document review can consume 70 to 80 cents of every dollar spent on e-discovery. A 2012 RAND Corporation study on litigant expenditures found that attorney and contract reviewer time spent reading documents dominated total discovery costs in cases with sprawling document populations — and that was before cloud storage made it trivially easy to accumulate millions of files across dozens of custodians. Linear review, in which a human eyes every document and makes a relevance call, is the most expensive, slowest, and (this is the part that should sting) least accurate method available. It is also, as of 2026, still the default in smaller firms and in matters where counsel has not made a deliberate choice about method.

That ends here. This chapter makes the affirmative case for AI-assisted review. Not uncritically, and not because vendors say so, but because the empirical record accumulated over fifteen years of controlled studies and court-supervised productions is, at this point, overwhelming. AI-assisted review, properly configured and properly supervised, is faster, cheaper, and more consistent than armies of first-year associates burning midnight hours in a document review warehouse. Your job is not to decide whether to use it. Your job is to use it defensibly.

The Baseline: What TAR Actually Is

Technology-assisted review — TAR, also called predictive coding or computer-assisted review — has been a known quantity in e-discovery since at least 2012. That year, Judge Andrew Peck of the Southern District of New York issued Da Silva Moore v. Publicis Groupe & MSL Group, 287 F.R.D. 182 (S.D.N.Y. 2012), the first federal court opinion explicitly approving TAR as an appropriate method of discovery review. He called it "the way of the future." He was right, and it took the profession another decade to notice.

TAR works by training a statistical model on a sample of documents that a human reviewer has coded. You — or a qualified reviewer — code a seed set. The system learns the pattern. It then applies that learned pattern to rank or categorize the remaining document population. The documents with the highest predicted relevance scores rise to the top. The documents with the lowest scores become candidates for withholding without full review.

Two principal workflows dominate. Simple passive learning (SPL, sometimes called TAR 1.0) uses a single training phase: code a static seed set, train the model, apply it. Continuous active learning (CAL, or TAR 2.0) is iterative — the system keeps learning as reviewers code documents throughout the production, presenting the most uncertain cases for human judgment and updating its predictions continuously. Maura Grossman and

Gordon Cormack, the researchers whose work has done more than any other to establish TAR's empirical credentials, published studies demonstrating that CAL substantially outperforms both SPL and linear review on recall. Their work showed CAL achieving recall rates above 75 percent while reviewing only a fraction of the document population that linear review would require. Numbers matter here: in a one-million-document production, that fraction can mean reviewing 150,000 documents instead of 1,000,000 — at equivalent or better accuracy.

That is the baseline. LLMs change the ceiling, not the floor.

What LLMs Add — and What They Do Not Fix

General-purpose large language models introduced something TAR never had: the ability to reason about document meaning rather than statistical pattern. A TAR model learns that certain words and phrases correlate with relevance. An LLM can read a contract clause and identify it as imposing a change-of-control obligation, even if that clause was drafted in unusual language that no seed-set document contained. It can read a chain of internal emails and summarize the discussion, identify the participants who appear to have known about the defect, and surface the one message that a keyword search would have missed entirely.

Relativity's aiR platform, Reveal's AI module, and Luminance — founded in 2016 by mathematicians from Cambridge and steadily refining its legal models since — all offer LLM-based review layers on top of existing infrastructure. The pitch is genuine: LLMs reduce the "I don't know what I don't know" problem that has always plagued keyword-based discovery. If your keyword list did not include "Project Falcon," you missed every document about the acquisition that blew up. An LLM reading for meaning finds Project Falcon even when it was never on the list.

But the hallucination risk does not disappear in the review context. It changes form. An LLM tasked with summarizing a document will, sometimes, misstate what the document says. Not invent a case citation — that particular failure mode, the one that cost the lawyers in Mata v. Avianca their reputations and their money, matters less in review — but mischaracterize the content of a real document in a way that causes it to be miscoded. A contract imposing a 30-day cure period gets summarized as having a 10-day cure period. A privilege log entry describes an email as involving outside counsel when it was entirely internal. Small errors. Consequential ones.

This is why no reputable legal technology vendor is selling fully automated LLM review without a human quality-control layer. Be deeply skeptical of any vendor who suggests otherwise.

Court Acceptance: Where the Law Actually Stands

Federal courts have been, on balance, permissive about TAR and increasingly open to AI-assisted methods — provided counsel can demonstrate that the method was reasonable, was disclosed, and was validated. The arc runs from Da Silva Moore in 2012 through Judge Peck's 2015 opinion in Rio Tinto plc v. Vale S.A., 306 F.R.D. 125 (S.D.N.Y. 2015), where he observed that "the use of predictive coding has now become an

acceptable and widely used method of review." By 2016, in Hyles v. New York City, Peck declined to compel a party to use TAR but wrote that the choice to use linear review instead was "a crazy thought" given the available technology. That sentence has been quoted in dozens of briefs.

Courts have not issued bright-line approval for LLM-based review specifically — the published opinions have not yet caught up with the 2023–2026 generation of tools. But the doctrinal framework in place is capacious enough to accommodate them. The operative standard, derived from Rule 26(g)'s certification requirement and the proportionality principle embedded in Rule 26(b)(1) since the 2015 amendments, is reasonableness. Was the method reasonable for the matter? Was it transparently disclosed? Was it validated?

In practice, this means three things. First, be prepared to explain your methodology to opposing counsel and, if challenged, to the court. "We used AI" is not an explanation. "We used continuous active learning on a population of 847,000 documents, trained on a seed set of 5,000 documents coded by a supervising associate, with elusion testing showing a recall rate of 81 percent at a 95 percent confidence interval" is an explanation. Second, disclose your methodology to opposing counsel — not because courts uniformly require it, but because courts increasingly expect it and because disclosure eliminates the most common challenges. Third, validate.

Rule 26, Proportionality, and the Quality-Control Obligation

Rule 26(b)(1), as amended in 2015, requires that discovery be "proportional to the needs of the case," considering the importance of the issues at stake, the amount in controversy, relative access to information, the parties' resources, and whether burden outweighs likely benefit. AI-assisted review is the proportionality argument made flesh. Reviewing 200,000 documents instead of 800,000, at a documented error rate comparable to human review, directly addresses the cost-benefit calculus the rule demands.

But proportionality cuts both ways. You cannot short-circuit quality control in the name of cost savings and claim proportionality as your defense. Rule 26(g) requires that each discovery response be signed by an attorney certifying it is "complete and correct as of the time it is made." That certification attaches to an AI-assisted production just as it does to a linear one. You sign it. You are responsible.

The practical implication is a quality-control sampling protocol you can document and, if necessary, produce to opposing counsel or the court. The standard approach — validated in TAR studies and accepted in multiple court-supervised productions — involves two measurements.

Elusion testing addresses the question you most need to answer: How many relevant documents did the system miss among the documents you are not going to review? Draw a random sample from the documents coded as non-responsive. A human reviewer codes that sample. The percentage of actually-responsive documents found in that sample is the elusion rate. A low elusion rate — industry consensus tends toward below 2 percent, though the acceptable threshold varies by matter and risk tolerance — confirms that the documents below your review cut-off are predominantly non-responsive.

Precision testing addresses the question your client cares about for cost: Of the documents coded as responsive, how many actually are? High precision means reviewers are spending time on genuinely relevant documents rather than false positives. A precision rate of 70 percent is considered acceptable in many matters; 85 percent or above is strong performance.

Both numbers go into your validation memo. Write the validation memo before production. Not after. If a challenge comes, you want documentation that preceded the production decision, not a reconstruction.

There is a third element courts have focused on for LLM-based review specifically: privilege. TAR systems have historically struggled with privilege identification — they can learn that documents involving certain people are often privileged, but they cannot read an email and independently recognize that it contains protected legal advice. LLMs are materially better at this task. Several platforms now offer LLM-based privilege screens that analyze document content for the presence of attorney-client communication or work product. Better — but not error-free. A privilege review error in production can become a waiver argument. A clawback dispute becomes motion practice that costs more than the review savings. Your privilege validation protocol needs to be at least as rigorous as your relevance validation protocol.

The Workflow in Practice

Here is a defensible workflow for an AI-assisted document review in a mid-size commercial litigation — say, a production population of 400,000 documents following initial processing, deduplication, and email threading.

Step one: Scope the universe. Before the AI touches a document, define the review population. Processing filters — date ranges, custodians, file types — narrow the 400,000 to, say, 280,000 after junk removal. Confirm your data retention and confidentiality settings with the platform vendor. This is not administrative box-checking: if you are using a cloud-based review platform, you need a data-processing agreement addressing your obligations under Model Rule 1.6. Chapter 12 covers that analysis in full. For now, confirm it exists before uploading a single document.

Step two: Decide on keywords, deliberately. Whether to run a keyword filter before training depends on your facts. In a focused, well-scoped matter, keywords efficiently reduce the training population. In a matter where the relevant issues are diffuse, or where the opposing party may have deliberately obfuscated communications, keywords create gaps. Know what you are trading off. The choice is yours to make and own.

Step three: Build and code the seed set. Your most senior available reviewer — ideally a supervising attorney who understands the case theory — codes the seed set. Not a paralegal in year one. Not a contract reviewer who received a two-paragraph case summary. The seed set codes are the intellectual foundation on which the entire model is built. Quality of the seed set determines quality of predictions. In CAL-based systems, the platform will surface documents for additional coding rounds; those rounds also require qualified reviewers.

Step four: Review the high-confidence responsive set. Documents scoring above your responsive threshold go to human review. Your team works through them. Privilege flags — whether generated by the LLM or by rule-based tools — are reviewed by an attorney. Not delegated.

Step five: Elusion test the non-responsive set. Draw your random sample. For a population of 280,000, 2,400 documents gives you a 95 percent confidence interval at a 2 percent margin of error. A qualified reviewer codes the sample. You calculate the elusion rate and document it.

Step six: Produce and file. Your production goes out with a Bates-stamped privilege log and the validation memo in your file. If opposing counsel challenges methodology, you respond from documentation.

The whole workflow, in a matter of this size, takes substantially less time than linear review. How much less depends on the platform, the facts, and the seed set quality. But the consistent empirical finding across Grossman and Cormack's work is that CAL-based review with proper training achieves equivalent recall to linear review while reviewing 30 to 50 percent of the documents. On 400,000 documents at $1.50 per document for contract reviewer time, that is a minimum $300,000 savings. Before you count attorney hours.

The Supervision Obligation You Cannot Delegate

ABA Formal Opinion 512, issued in 2024, addresses generative AI directly and grounds the analysis in Model Rule 1.1's existing competence framework. The opinion's core point applies squarely in the document review context: the attorney's obligation to supervise non-lawyer assistance — including AI-assisted work — under Rule 5.3 does not diminish because the tool is sophisticated. If anything, sophistication increases the supervision burden, because plausible-sounding outputs are more likely to slip past an inattentive reviewer.

Supervision in the document review context means, at minimum, four things. One: the attorney who certifies the production under Rule 26(g) has personally reviewed the validation methodology and sampling results — not a summary prepared by the vendor. The actual numbers. Two: the attorney has reviewed a random spot-check of documents coded responsive, even ones the AI flagged clearly, to confirm coding aligns with the case theory. Three: the attorney has reviewed the privilege log for accuracy and established clear protocols for privilege calls. Four: the attorney knows what happens to the data after production. What is the vendor's retention period? Does the data leave the platform? Has the client consented?

Client consent is live. Formal Opinion 512 flags the disclosure question, and several state bars have gone further. The California State Bar's 2024 guidance on AI and professional responsibility notes that disclosure obligations may apply when AI is used in ways that could materially affect client interests. Using AI to make relevance decisions in a document production almost certainly qualifies. Have the conversation with your client. Add AI-assisted review disclosure to your engagement letter as of today.

The Error Rate You Are Actually Comparing Against

One pushback you will hear — sometimes from judges, sometimes from opposing counsel trying to impose linear review as a strategic matter — is that AI makes mistakes. True. But so does human review, at a rate that would be embarrassing if we measured it as carefully as we measure AI.

Grossman and Cormack's research on human review accuracy is not encouraging for the linear camp. Their work has documented reviewer agreement rates on relevance calls — how often two reviewers independently coding the same document reach the same conclusion — in the range of 60 to 70 percent on contested documents. Published research has found that human reviewers miss relevant documents at rates between 20 and 50 percent depending on matter complexity. We hold TAR and LLM systems to a recall standard of 75 or 80 percent and treat it as a limitation. We never applied that standard to the linear review it replaced.

This does not mean AI is infallible or that quality control is optional. It means the honest comparison point is real human review, not perfect human review. Perfect human review does not exist. It never did. The argument is not "AI is flawless." The argument is "AI is measurably better than the alternative we have been using for decades without ever auditing it."

A Note on Smaller Matters

Everything here scales. The same principles apply to a production of 15,000 documents in a commercial contract dispute. The seed set is smaller — perhaps 300 documents. The elusion sample is smaller — roughly 370 documents gives you a 95 percent confidence interval at a 5 percent margin of error for a population of 15,000. The workflow is the same. The supervision obligation is identical.

The mistake small-firm practitioners make is assuming AI-assisted review requires a Relativity subscription and an e-discovery vendor with a six-figure setup fee. It does not. Logikcull and Everlaw both offer accessible cloud-based platforms with AI-assisted review functionality designed for matters of this size. The economics work at 15,000 documents, not just 1.5 million.

If you are handling a case with more than 5,000 documents and you are doing linear review without an intentional reason for the choice, you are leaving time and money on the table. Your client's.

What to Tell Opposing Counsel

Transparency about review methodology has evolved from a courtesy to a near-expectation in federal litigation. The Sedona Conference, the nonprofit think tank that has done more than any other organization to develop e-discovery doctrine, recommends early collaborative discussion of review methodology as part of the Rule 26(f) conference. Practically, this means arriving at your 26(f) conference prepared to describe your proposed methodology at a level of specificity that lets opposing counsel understand it.

You do not need to reveal your training documents. The work product doctrine protects those. What you should be prepared to describe: the review platform and method, the approximate training population size, and your validation approach. If you propose to share recall and precision metrics after validation, say so.

Courts appreciate procedural transparency, and it cuts off most tactical challenges before they start.

If opposing counsel insists on linear review — sometimes for legitimate reasons, sometimes as a fishing expedition dressed up as a methodology objection — courts have generally declined to mandate it over a producing party's objection. Hyles is the clearest statement: the choice of review methodology is the producing party's to make, provided it is reasonable. You do not need permission to use AI-assisted review. You need to be able to defend it.

You will be able to defend it. That is what the validation memo is for.

CHAPTER VII

Due Diligence: M&A, Real Estate, and Regulatory Deals

The average M&A data room in a mid-market deal — say, a $200 million acquisition of a regional manufacturer — contains somewhere between 3,000 and 15,000 documents. A large-cap deal can push past 100,000. Commercial real estate portfolio acquisitions routinely involve 500 to 2,000 individual leases, each carrying its own riders, estoppel certificates, SNDAs, and amendments. Regulatory deals — HSR filings, FCC license transfers, healthcare acquisitions requiring CMS review — add a third layer of agency documents on top of the commercial contracts.

The math has never worked cleanly. A team of three associates, billing at 2026 rates and sleeping occasionally, cannot review 80,000 documents in a four-week exclusivity window with the attention each document theoretically deserves. What actually happens — and every transactional lawyer reading this knows it — is triage. Certain folder categories get deep review. Others get spot-checked. The methodology is reasonable and defensible, and clients have accepted it for decades. But it carries real risk.

AI-assisted due diligence does not eliminate that risk. It redistributes it. The goal is to get more documents into meaningful review faster, so attorney attention concentrates on the genuinely material exceptions rather than the routine. That goal is achievable. The workflow matters enormously.

The Data Room Reality

When a data room opens, you are not looking at organized information. You are looking at whatever the seller's team assembled under time pressure: PDFs scanned at odd angles, Excel schedules exported as images, NDAs executed in 2011 with handwritten margin amendments, and a folder labeled "Misc." containing 400 items. The seller's counsel has organized it according to their own logic, which may or may not correspond to your diligence checklist.

This is the first place AI helps, and it is not glamorous. Modern document-processing pipelines — including features native to platforms like Kira (now part of Litera), Luminance, and Harvey's enterprise offering — can ingest a data room's contents, classify documents by type (lease, loan agreement, material contract, employment agreement, regulatory license), and surface an inventory to your team within hours rather than days. You get a map before you start walking.

Classification is not review. That distinction matters. When an AI platform tells you that folder 12 contains 47 lease agreements, you know where to look. You do not yet know what is in those leases, whether any contain co-tenancy clauses that would affect deal economics, or whether three of them have already been

terminated and the seller simply forgot to remove them. Classification is the precondition for efficient review. Not the review itself.

Contract Abstraction: Where the Speed Gains Are Real

Contract abstraction is the extraction of specific data points from a document — parties, effective date, governing law, term length, termination rights, assignment restrictions, change-of-control provisions — and their population into a structured summary matrix. In a traditional diligence workflow, a paralegal or junior associate reads each contract and manually populates that matrix. That takes anywhere from fifteen minutes for a simple NDA to several hours for a complex commercial agreement with multiple schedules.

AI-trained abstraction tools process the same documents in seconds per contract. Kira's published case studies have described throughput that would require multiple paralegal-weeks compressed into hours. Luminance, built on machine learning research from Cambridge, was adopted early by several Magic Circle firms for precisely this use case. The speed is real.

But the accuracy deserves scrutiny. Current abstraction tools perform well on clean, standard provisions — governing law clauses, term dates, renewal options. They perform less reliably on negotiated language, bespoke definitions, and provisions where operative text is spread across multiple sections and cross-references. An AI platform that correctly identifies "the contract has a change-of-control provision" can still miss the fact that the definition of "change of control" was specifically narrowed in Amendment No. 3 to exclude the current transaction structure.

The rule is simple: treat AI abstraction output as a first-pass summary that an attorney must verify against the underlying document at any point where the abstracted term is material enough to go into a representation in the purchase agreement. Not every abstracted data point requires line-by-line verification. The governing law of a standard vendor contract for office supplies does not. The assignment restriction in a key customer agreement representing 40 percent of the target's revenue absolutely does.

Building the Red-Flag Report

Exception reporting — identifying contracts that contain unusual, adverse, or deal-threatening provisions — is the highest-value deliverable in most diligence engagements. Clients do not pay for a list of everything they already expected. They pay for the problems you found.

AI accelerates red-flag identification in two distinct ways. First, breadth: a well-configured abstraction platform running a red-flag query across 10,000 contracts simultaneously will surface every instance of "unlimited liability" or "automatic renewal absent a 30-day notice" faster than any team of lawyers. Second, consistency: the AI applies the same criteria to document 9,847 as it applied to document 1. Humans get tired. Humans skip the annex to the annex. The AI does not.

To make this work, you have to define your red flags explicitly before you run the analysis. This is the attorney's job, not the platform's. Common red-flag categories in M&A due diligence include:

- Change-of-control triggers: Does assignment require consent upon a change of control? Does the definition of "change of control" include an asset sale distinct from a stock transaction?
- Material adverse change definitions in existing credit facilities or material commercial agreements that could be triggered by the announcement itself
- Termination-for-convenience clauses counterparties could exercise on notice of the deal
- Most-favored-nation or exclusivity provisions limiting the target's ability to serve other customers post-close
- IP ownership and work-for-hire clauses in agreements with independent contractors
- Non-compete and non-solicit obligations that survive closing and bind the acquiring entity

You configure the platform with these categories. It runs. You review the hits. Then you verify the hits that matter.

Do not skip the verification step because the output sounded certain. The error mode here is subtler than in legal research — unlike the fabricated citations in Mata v. Avianca, the AI is working with real documents — but misreading complex cross-references, or flagging a provision that was superseded by a later amendment, is entirely possible. The attorney who signs off on the exception report is the attorney who owns it.

Schedule Population

Purchase agreements in M&A deals include disclosure schedules: lists of material contracts, exceptions to representations, employee benefit plans, litigation disclosures, and intellectual property registrations. Populating those schedules from the diligence data room is one of the most tedious mechanical tasks in a closing process. It is also an area of genuine, documented AI-assisted value.

An AI platform that has ingested the data room and abstracted the material contracts can, with appropriate configuration, generate a draft schedule of material contracts, a draft intellectual property schedule drawn from registration records, and a draft list of real property leases. Draft is the operative word. These outputs require attorney review for completeness, accuracy, and appropriateness — the seller's disclosure obligations are defined by the representations in the purchase agreement, and the scope of what must be disclosed requires legal judgment, not just document retrieval.

Schedule errors in purchase agreements generate post-closing disputes. Indemnity claims. In egregious cases, fraud claims. The efficiency gain from AI-assisted schedule population is worth capturing. The risk of treating an AI-generated schedule as a final schedule, without clause-level verification of the representation it supports, is not.

Real Estate: Lease Abstraction at Scale

Commercial real estate acquisitions present the abstraction problem in its purest form. A portfolio acquisition of 150 retail locations involves 150 leases, each with its own rent structure, renewal options, co-tenancy requirements, permitted-use clauses, exclusivity protections, and landlord-consent requirements. Add the amendments, SNDAs, estoppel certificates, and guaranties, and a single-location file can run to dozens of documents.

Lease abstraction platforms — including lease management modules built into Kira and Luminance, as well as standalone tools like Lease Accelerator — have been in commercial use for several years. Law firms with significant real estate practices were running AI-assisted lease abstraction by 2019. The workflow is mature relative to newer generative AI applications: define your standard abstraction fields, ingest the lease portfolio, review the output, escalate anomalies.

Where it breaks down: non-standard lease structures. Ground leases. Sale-leaseback arrangements. Leases drafted under non-standard jurisdictional terminology. Leases that have been substantially modified by successive amendments such that the operative economic terms are no longer primarily in the base document. These require attorney review of the underlying documents. Period. The AI will do its best. Its best is not good enough to stake a purchase price on without verification.

Co-tenancy clauses deserve particular attention because they are both common in retail real estate and regularly misread by abstraction tools. A co-tenancy clause allows a tenant to reduce rent or terminate if a named anchor tenant leaves the property. Whether the current deal structure constitutes a "transfer of substantially all assets" that might trigger such a clause is a legal question requiring analysis, not extraction. Flag it. Read the clause. Advise the client.

Regulatory Deals: A Third Layer

Healthcare acquisitions, defense contractor purchases, and deals requiring FCC, FERC, or HSR approval add a regulatory due diligence layer that involves an entirely different document set: government contracts, agency licenses, regulatory filings, and compliance records. AI platforms built for commercial contract abstraction may not be well-configured for government contract review, which involves FAR and DFARS provisions, small-business status certifications, and teaming arrangements with their own assignment restrictions.

Specialized platforms exist for government-contract diligence. For practitioners who handle regulatory deals only occasionally, the more realistic workflow is using a general-purpose AI tool to surface document categories and structure the review checklist, then conducting substantive review with counsel who has specific regulatory expertise.

One particular risk in regulatory due diligence: outdated or lapsed licenses. An AI platform can identify that a document is a regulatory license and abstract the license number, issuing agency, and expiration date. It cannot tell you whether that license has been renewed, whether there are pending enforcement actions, or whether the agency's public records reflect a different status than the document in the data room. That

requires checking the agency's actual records. The FDA's establishment registration database, the FCC's License Management System, CMS's provider enrollment records — these are the authoritative sources. AI gets you to the question. You answer it yourself.

Client Disclosure and Third-Party Platforms

Here is where efficiency gains meet ethical obligations, and where many practitioners have not thought carefully enough.

When your firm uploads a client's data room to a third-party AI platform — Luminance, Harvey, Kira, any of them — you are transmitting confidential client information to a system operated by a vendor. The confidentiality analysis under Model Rule 1.6 requires that you take "reasonable measures to prevent the unauthorized disclosure of, or unauthorized access to, information relating to the representation." ABA Formal Opinion 512 expressly extends this analysis to AI tools: lawyers using generative AI must understand what the tool does with inputted information, including whether it is used for model training, retained on vendor servers, or accessible to vendor personnel.

The due diligence context is particularly sensitive. A data room contains, at minimum, the target's material contracts, financial statements, and employee records. It often contains information about ongoing litigation, regulatory investigations, or pending intellectual property disputes that the target considers among its most confidential information. Uploading that to a vendor platform without a robust data-processing agreement — one that prohibits training on client data, specifies retention and deletion timelines, and addresses breach notification — is not a "reasonable measure." It is a bet that nothing goes wrong.

Before any client matter touches an AI platform, read the vendor's data retention policy. Not the summary page. The policy. Establish whether your firm has an enterprise agreement with the vendor that includes appropriate data protections. If not, get one, or use a different tool. ABA Formal Opinion 512 makes clear that competence under Rule 1.1 now includes understanding the tools you deploy — which means understanding what happens to client information after it enters those tools.

The conflict-of-interest question is less commonly discussed but worth raising. If a major AI platform has processed due diligence for both the acquirer and the target — or for an adverse party in a related transaction — there is at least a theoretical argument that vendor-side model training, if it occurred, creates an information asymmetry. Most major enterprise agreements now include provisions specifically addressing this. Review them. If yours does not, ask.

Client disclosure of AI use in the diligence workflow is increasingly an expectation rather than an optional gesture. Sophisticated buyers — private equity sponsors, strategic acquirers with experienced in-house teams — often ask directly whether AI was used and on what terms. Be prepared for that conversation. Have a clear answer about which platform, what data protections apply, and how attorney oversight was structured. Treating the question as unwelcome is exactly the wrong instinct. Clients whose deal documents are being processed by third-party systems have a legitimate interest in knowing.

The Representation That Goes Into the Purchase Agreement

Everything in the diligence process is preparation for a set of representations in the purchase agreement. The seller represents that the material contracts have been disclosed and are in full force and effect. The buyer represents it has conducted the diligence it needed to conduct. Both sides stake indemnification obligations on the accuracy of those representations.

This is where human review is not optional. Not as a matter of professional preference. As a matter of what the representation requires.

An AI platform can identify that a contract exists and abstract its key terms. It cannot tell you, with the confidence required to place a representation in a signed purchase agreement, that a contract is "in full force and effect" — because that legal conclusion requires judgment about governing law, the conduct of the parties, and whether course-of-dealing or waiver arguments affect the operative terms. It cannot tell you whether a consent requirement is enforceable, whether a counterparty would actually withhold consent, or whether the provision has been waived by the counterparty's conduct over years of dealings.

Before any abstracted term supports a representation in the purchase agreement, an attorney must read the underlying document. Full clause. Full context. The efficiency gain is in getting the attorney to the right document, in the right section, with the right preliminary flags, fast. The judgment remains the attorney's.

This is not a counsel-of-perfection standard. It is a proportionality standard: the depth of attorney review should match the materiality of the representation it supports. A co-tenancy clause in a lease representing 3 percent of the portfolio's revenue warrants junior-associate review with a spot-check by a senior. A consent requirement in an agreement with the target's largest customer — one that could allow that customer to walk away from a relationship representing $18 million in annual revenue — warrants partner eyes on the clause before closing. Those two situations are not the same, and treating them identically is how deals go wrong.

A Practical Workflow

Stage 1 — Intake and Classification (Days 1–2): Load the data room into your AI platform. Generate a document inventory and classification report. Identify gaps against your diligence checklist. Request missing documents from the seller before the review clock runs.

Stage 2 — Contract Abstraction (Days 2–7): Run abstraction on the document set, configured for your standard fields and any deal-specific fields the client has prioritized. Review the abstraction output for completeness and flag obvious errors before distributing to the team.

Stage 3 — Red-Flag Analysis (Days 3–10): Define red-flag criteria with the deal team before running queries. Review AI-identified hits. Verify hits against underlying documents for every issue that could affect deal structure, pricing, or reps-and-warranties coverage.

Stage 4 — Exception Reporting (Days 7–14): Draft the exception report. Each flagged item should include the document reference, the relevant clause quoted in full, the attorney's analysis, and a recommended deal response. Do not let the AI draft the analysis section. It can populate the document reference and extract the clause. The analysis and recommendation are yours.

Stage 5 — Schedule Population (Days 10–20): Use AI-generated drafts as starting points for disclosure schedules. Verify each schedule against the relevant representation before signing. Do not populate a final schedule from an abstracted summary alone.

Stage 6 — Pre-Closing Confirmation: Re-check any abstracted terms that support representations against the underlying documents a final time before signature. Deal documents change after initial review. Amendments get executed. Consents get obtained. The data room at closing may not be identical to the data room you reviewed in week two.

The tool does not close the deal. You do. AI compresses the hours required to find the material provisions, reduces the number of documents that go unread, and helps a lean team cover ground that would otherwise require significantly more headcount. Those are real advantages in a competitive deal environment, and ignoring them is its own form of professional negligence — failing to use available tools to serve your client efficiently.

But the advantage exists only when the workflow is honest about what AI can and cannot do. It can read fast. It cannot read carefully. It can flag provisions. It cannot advise a client whether to close. Every representation that goes into the purchase agreement is backed by an attorney's professional judgment, and that judgment is not something a vendor platform absorbs by ingesting the data room. It is something you bring to the table — which is still, in 2026, the reason you are at it.

CHAPTER VIII

Client Communication: Drafting Summaries, Emails, and Status Updates

Every year, the Clio Legal Trends Report measures what attorneys actually do with their time. The 2023 edition put a number on it that most lawyers recognize instantly: the average attorney bills roughly 2.9 hours out of an eight-hour workday. The other five hours vanish into non-billable work — internal email, admin, and, consuming more of that remainder than most people admit, client communication. Not the strategic kind. The translating kind. The "here is what happened at the hearing today in language you can understand" kind. The "we are still waiting on the court, here is where things stand" email you have now written four times for the same matter.

This is the territory where AI assistance makes an immediate, measurable difference for practicing attorneys. The judgment-intensive work — advising on litigation strategy, negotiating deal terms, assessing credibility — remains entirely yours. But reducing a 22-page settlement agreement to a three-paragraph plain-language summary, or turning bullet-point notes from a deposition into a coherent client update email, is the kind of pattern-intensive drafting that large language models handle capably. If you supervise the output correctly.

That last clause is doing serious work. And this chapter is largely about it.

The Actual Use Cases

Client-facing communication divides into roughly four categories where AI assistance is practically useful: status update emails, matter summary memoranda (for the client file and for new decision-makers on the client side), settlement summaries and recommendation letters, and demand letters where tone and structure are more variable than substance.

Each of these involves two professional obligations that run in parallel. First, the substantive accuracy obligation — you cannot send a client something factually wrong about their own matter. Second, the communication obligation under Model Rule 1.4, which requires keeping a client reasonably informed and explaining things to the extent reasonably necessary for the client to make informed decisions. AI accelerates the drafting side of that obligation. It does not discharge it.

There is also a third obligation lurking underneath both of them: confidentiality. Client communications, by definition, involve client facts — which means any AI tool you use to draft them must be evaluated against your Rule 1.6 duties before you paste anything in. More on that shortly.

What AI Does Surprisingly Well Here

The capability that makes AI genuinely useful for client communication is not cleverness. It is register flexibility. A well-prompted model can take a dense paragraph of legal reasoning — say, the court's discussion of the summary judgment standard applied to a spoliation argument — and restate it at a ninth-grade reading level without losing the operative meaning. It can do this in seconds and produce a draft you revise rather than write from scratch.

That matters because plain-language translation is actually hard. Legal training runs in the opposite direction. Law school teaches you to be precise at the expense of accessible, and that precision hardens over a career. Many experienced litigators find plain-language drafting more cognitively demanding than writing for the bench. AI does not have that problem. It was trained on a vast range of human communication registers, and it can move between them on instruction.

The practical upshot: a lawyer who dictates four bullet points about a deposition into an AI prompt and asks for a 200-word client email in plain English, written in a warm but professional tone, will get a usable draft in under a minute. Revision takes five. The alternative — writing the email directly — takes fifteen to twenty minutes for most people, more when you are tired or interrupted.

That compression compounds. If you are sending eight client updates per week, getting each one from fifteen minutes to six is saving seventy-two minutes. In a five-day week. Every week.

Prompting for Tone, Reading Level, and Format

The quality of what you get back is almost entirely determined by the specificity of what you put in. A vague prompt produces a generic draft. A precise prompt produces something genuinely close to what you would have written yourself.

Three variables determine whether a client communication draft is useful or not: tone, reading level, and format constraints. You need to specify all three explicitly.

Tone is not just "professional." It is the difference between a demand letter that reads as measured and firm versus one that reads as aggressive, or a settlement summary that reads as cautiously optimistic versus one that reads as a strong recommendation to accept. Be specific. "Write this in a calm, factual tone that does not editorialize" is different from "write this to convey that we believe this outcome is positive for the client." The model cannot infer which one you want.

Reading level should be stated numerically or by analogy. "Write at a sixth-grade reading level" or "write for someone with no legal background" produces different results than leaving reading level unspecified, which usually returns text that sounds like a legal professional wrote it because the model is drawing on legal-professional training data. If your client is a sophisticated in-house counsel, you might not need to specify — but if your client is a first-time plaintiff in a personal injury case who has never been to court before, you do.

Format constraints matter especially for emails and summaries. Specify maximum length. Specify whether you want headers or flowing prose. Specify whether bullet points are appropriate for this client relationship. A prompt that says "write a client status update email, no more than 200 words, three short paragraphs, no legal jargon, no bullet points" will produce something more useful than "write a client status update email."

Here is a concrete example. Suppose you are updating a small-business client on a contract dispute. You have just received the other side's response to your demand letter, and it was a partial concession — they acknowledged liability for one of the three disputed invoices but rejected the others. Your prompt might look like this:

"Draft a client email. The client is a small-business owner with no legal background. Tone: factual, reassuring but not overconfident. Reading level: eighth grade maximum. Length: under 250 words. No legal citations, no case names. Summary of facts to include: [three-sentence factual summary you type out]. Update: opposing party acknowledged liability on Invoice #1 ($12,400), rejected Invoices #2 and #3. We plan to respond by [date] with a revised settlement proposal. Do not include any legal strategy. Do not make any promises about outcome."

That last instruction — "do not make any promises about outcome" — is not optional. It is easy to omit and costly when an AI draft slips into projecting certainty about results. You need to either specify it in the prompt or read every draft with that risk specifically in mind before you send it.

Working Through the Common Formats

Status Updates

Status updates are the highest-volume, lowest-risk category. The facts you need to communicate are typically not sensitive — a hearing date, a filing made, a response received. The judgment you are withholding from the communication is also relatively low-stakes.

A useful practice is to maintain a running matter log in your own notes that you can paste as bullet points into a prompt at the end of each week. The model converts the log into a coherent narrative email. You review, adjust tone where necessary, confirm that nothing in the draft constitutes legal advice you did not intend to give, and send.

One thing to watch: AI drafts of status updates tend toward false reassurance. "The matter continues to progress well" is filler, not information. If nothing happened this week, say that — your prompt should specify whether the communication is a genuine update or a "no news" check-in, and the tone should match.

Settlement Summaries

Settlement summaries require more care. Here you are asking the model to help you translate a financial and legal outcome — with terms, conditions, release language, and payment structure — into a document a client can read and sign off on. The stakes are higher and the precision requirement is more exacting.

The approach that works: draft the legal analysis yourself, in your normal professional register, and use AI to restate it in plain language. Do not ask the model to summarize the settlement agreement directly from the document. That introduces the risk of the model misstating a term — transposing a number, softening a condition that isn't soft, or omitting a release provision that the client needs to understand before accepting. You own the legal analysis. AI restates what you've already correctly stated.

Run every settlement summary draft through a factual checklist before it goes to the client: dollar figures, payment timing, key conditions, release scope. Every one of those elements should be verified against the underlying document, not just the draft.

Demand Letters

Demand letters occupy interesting ground. The tone is more variable than in most legal writing — some clients want aggressive, some want to preserve a business relationship, some situations call for restrained formality that implies capability without threatening it. AI is reasonably good at adjusting along that continuum if you tell it where on the dial to sit.

The risk in AI-drafted demand letters is the opposite of the settlement summary risk. Where summaries risk omission, demand letters risk overstatement — a model that interprets "firm and aggressive" as "exaggerated threat" can produce a letter that misrepresents what your client is actually entitled to recover or implies legal positions you cannot support. Read the draft as opposing counsel would. Would anything in it embarrass you if challenged?

The Confidentiality Constraint

This chapter is not the place for a full Rule 1.6 analysis — that lives in Chapter 12. But any discussion of AI-assisted client communication has to flag the constraint directly, because the use cases here are exactly the ones where confidentiality violations happen.

Client communication drafts require client facts. You cannot prompt meaningfully without them. And client facts are confidential under Rule 1.6 unless disclosure is authorized or falls within a recognized exception.

The question you must answer before pasting any client information into an AI tool is simple: what happens to the data I'm about to submit? Consumer-tier AI products — the free or personal-subscription versions of tools like ChatGPT or Gemini — typically reserve rights to use input data for model training by default, absent an explicit enterprise agreement or opt-out. Pasting your client's name, opposing party, deal terms, or medical facts into a tool with those data retention terms is, at minimum, a Rule 1.6 risk that requires analysis. In many circumstances it is a violation.

Enterprise agreements with data-processing terms and no-training commitments change the analysis. Most legal-specific platforms (Harvey, Lexis+ AI, Casetext) have executed enterprise terms that include confidentiality protections adequate for law firm use. Check your firm's approved tool list before you use any tool for client communication drafting. If there is no approved list, create one, or ask for one.

The practical rule for this chapter: never paste client-identifying information into a tool you have not vetted. You can often accomplish the same draft by substituting placeholders — "Client A" for the client's name, "the disputed amount" for a specific sum — and then personalizing the output yourself. It takes thirty seconds and removes the exposure.

It Is Still Your Communication

ABA Formal Opinion 512, issued in July 2024, addressed the use of generative AI by lawyers. Its language on supervision is worth reading carefully. The Opinion applies Model Rule 5.3 — which requires lawyers to make reasonable efforts to ensure that nonlawyer assistance conforms to the lawyer's professional obligations — to AI systems. The logic is that when a lawyer uses AI to draft something the lawyer then transmits, the AI's work product is the equivalent of work performed by a supervised nonlawyer. The lawyer is responsible for it.

The Opinion is explicit: "a lawyer using generative AI must ensure the final work product satisfies all professional obligations as if the lawyer had drafted it personally." There is no AI exception to Rule 5.3. There is no "the AI wrote it" defense to a competence complaint or a malpractice claim.

This is not a theoretical concern. Client communications are one of the most common sources of bar grievances. Misstating what a client agreed to in a settlement, creating a reasonable expectation about outcomes that the attorney then fails to deliver, making representations about the law that turn out to be wrong — all of these have generated discipline proceedings. AI makes all of them easier to do accidentally if you are reviewing AI drafts with less rigor than you would apply to your own writing.

The reason human review becomes harder with AI drafts is that they read well. This is the same cognitive trap that made Mata v. Avianca possible — the fabricated citations looked like real citations, read like real citations, and were formatted like real citations, so they were not verified like they would have been if they had looked suspect. An AI-drafted client email sounds like a lawyer wrote it. It flows. It uses the right register. It does not trip your proofreading instincts the way a paralegal's early draft might, because the paralegal's early draft has obvious roughness that signals "needs review."

You need to override that instinct. Read AI-drafted client communications with deliberate skepticism. Not of the grammar — of the substance.

The Competence Obligation Under Rule 1.1

Model Rule 1.1 requires competence, defined to include the legal knowledge, skill, thoroughness, and preparation reasonably necessary for the representation. Comment 8 to Rule 1.1 extends this to keeping abreast of changes in the law and its practice, including the benefits and risks of relevant technology.

ABA Formal Opinion 512 treats that Comment 8 obligation as applying directly to AI tools: a lawyer who uses AI must understand, at a functional level, what the tool does and does not do. You do not need to

understand transformer architecture. You do need to understand that AI models generate plausible text rather than verified facts, that they can produce confident-sounding errors, and that their outputs require review calibrated to the risk level of the communication being drafted.

For client communication, that means understanding two specific failure modes.

The first is factual confabulation — the model stating something as fact that you did not tell it, and that is wrong. This is less common in client communication drafting than in legal research (because you are providing the facts), but it can happen when a model makes an inference you didn't authorize. "The court's ruling was in your favor" is not something you want a model writing unless you explicitly told it that.

The second is tone drift. A model asked to write a "strong" demand letter may write something that sounds like it implies criminal liability where none exists, or that makes settlement demands that don't match the actual exposure. Tone drift is harder to catch on a fast review precisely because it is stylistic rather than factual.

Neither failure mode is exotic. Both are manageable with adequate review. Rule 1.1 requires that review to happen.

A Practical Workflow

The workflow that works in practice is not complicated.

Start with your own notes. Before you touch an AI tool, write — in bullet points or dictated fragments — the substance of what the client communication needs to convey. The facts, the update, the recommendation, the action item. This takes three minutes. It is also the step that forces you to know what you want to say before you delegate the drafting.

Strip identifying information if you are working outside a vetted enterprise tool. Substitute placeholders for client names, dollar figures you don't want logged, and any sensitive factual detail. You can re-insert accurate information into the draft yourself.

Write a precise prompt. Specify tone, reading level, format constraints, and length ceiling. Add explicit instructions against making promises about outcomes or including legal positions you did not supply.

Review the draft for substance, not style. Check every factual claim against your bullet points. Check every number against the record. Look for implied representations you did not authorize. Look for tone drift in both directions — too aggressive, or too reassuring.

Revise in your own voice where the draft doesn't sound like you. A client who has worked with you for three years knows your register. An AI-drafted email that sounds like a different attorney is a small credibility disruption. Worth fixing.

Send it as your own. Not "here's a summary I had AI write for you." You are the communicator. The tool is your drafting assistant, the same as any other. The communication is yours.

That is not just a practical point. It is the ethical one. Model Rules 1.1, 1.4, and 5.3 do not distinguish between an attorney who wrote a letter personally and one who directed a nonlawyer — or an AI — to draft it. The professional obligation attaches to the communication, not the drafting process. Whatever lands in your client's inbox came from you.

AI makes the drafting faster. Your name makes it final.

CHAPTER IX

Deposition Preparation: Outlines, Chronologies, and Impeachment Sets

Few tasks in litigation feel more like punishment than deposition preparation. Not because the work is intellectually uninteresting — understanding a witness thoroughly enough to trap them in a contradiction is genuinely skilled work — but because of the sheer volume of material that has to be processed before the first question can be asked.

Consider a mid-complexity commercial dispute, the kind where two companies are fighting over a $14 million earnout. You have the deponent's prior deposition from related state-court litigation: 340 pages. You have four years of emails between the deponent and counterparties, roughly 1,100 documents the review team has coded as responsive. You have the deponent's prior declarations submitted in support of a summary judgment motion, his LinkedIn posts that somebody helpfully screen-grabbed at the start of the case, and two hours of recorded testimony from an earlier regulatory proceeding. The paralegal who would normally synthesize this into a working outline charges $95 an hour and will need eighteen to twenty-two hours to do it properly. You have a week.

This is exactly the scenario where AI delivers real, measurable value. Not hypothetical, productivity-deck value. Actual reduction in time-to-outline. With the caveat — and it is not a small caveat — that every line the AI produces has to be checked against the underlying record before you take it into a deposition room.

That caveat is the entire lesson of this chapter.

The Anatomy of a Deposition Prep Package

Before discussing what AI can do, it helps to inventory what a deposition prep package actually contains.

A well-prepared attorney walks in with at minimum: a witness outline organized by topic, with anticipated answers noted; an exhibit list cross-referenced to outline sections; a chronology of the witness's relevant acts and communications sourced to specific documents; and, for a hostile or adverse witness, an impeachment set — a sequence of prior inconsistent statements organized to be deployed at the confrontation moment. Building all of that from scratch, for a complex witness, takes time. The paralegal hours are real. The attorney review time on top of the paralegal work is real. And that work is happening simultaneously with briefing deadlines, client calls, and the other three depositions scheduled in the same month.

AI shortens the preparation phase. It does not eliminate the attorney's oversight obligation. Understand that distinction, and you can use these tools effectively. Confuse the two, and you are headed somewhere considerably worse than a bad deposition.

Feeding the Transcript: What AI Does Well

Start with prior testimony. If your deponent has testified before — in related litigation, a government investigation, or a prior version of the same dispute — that transcript is your foundation. It establishes baseline positions the witness is now locked into. AI can process a transcript far faster than a human and surface inconsistencies, topic clusters, and evasive non-answers with reasonable accuracy.

Here is a prompt structure that works:

> *"Below is a deposition transcript. The deponent is [NAME], a former CFO. Do two things: (1) Extract every statement the witness makes about the revenue recognition methodology used in fiscal years 2021 and 2022. For each statement, provide the verbatim quote, page number, and line numbers. Do not summarize — quote exactly. (2) Flag any statement that appears to qualify, hedge, or contradict an earlier statement in the same transcript. Provide verbatim quotes and page-line citations for both statements in every flagged pair."*

Several things about this prompt are deliberate. The verbatim instruction matters enormously. LLMs have a strong tendency to paraphrase. In deposition prep, paraphrase is actively dangerous. If the witness said "we may have used an accelerated recognition schedule," and the AI renders that as "witness confirmed use of accelerated schedule," you now have a false premise for a question. Ask it, the witness corrects you, and you have burned the moment — and telegraphed your theory to opposing counsel.

The page-and-line citation instruction is equally critical. And equally prone to failure. LLMs can and do hallucinate transcript citations. They will return page 247, lines 3 through 8, with apparent confidence, and when you pull the actual transcript, line 3 on page 247 begins a discussion about something else entirely. This is not a flaw that has been solved by 2026. It is a structural feature of how language models process and recall text.

The protocol, therefore: treat every AI-returned citation as provisional. Before any page-and-line reference goes near your outline, you open the transcript and you read the passage. Every one. No exceptions.

If you are using a platform like Casetext — now integrated into Thomson Reuters — or Harvey, which processes uploaded documents with retrieval-augmented generation rather than relying on the model's parametric memory, accuracy improves meaningfully. But "improves meaningfully" is not the same as "is reliable." The verification step does not disappear. It just becomes faster.

Building the Chronology

Document chronologies are the workhorses of deposition preparation and, separately, of trial. They impose temporal order on what otherwise feels like an undifferentiated mass of emails, contracts, board presentations, and forwarded text message threads. A good chronology lets you reconstruct what the witness knew and when they knew it — which is, in most commercial litigation, the central question.

AI accelerates this work substantially. Load the relevant document set and prompt for chronological extraction:

> *"I am uploading 200 documents from a document review export. For each document, identify: (1) the date, (2) the author and recipients, (3) a one-sentence description of the subject matter, and (4) any mention of the following topics: the earn-out calculation methodology, the revenue forecast model, the September 2022 board meeting. Output results as a table sorted by date, oldest first."*

Most enterprise-grade platforms can run this on a multi-hundred-document set in minutes. A paralegal working through the same 200 documents, doing it properly, would take a full workday.

What the AI will miss: context. A document dated October 12, 2022 that says "per our conversation" means something different if you know there was a board meeting on October 10 and a termination notice on October 14. The AI will extract the date and the surface subject matter. It will not reliably infer the significance of the sequence unless you tell it what to look for. Your job is to bring the theory. The AI supplies the sorted list.

Also watch dates. AI tools occasionally misread dates from scanned documents, particularly when formatting is ambiguous — is 01/02/2022 January second or February first? — and the error flows silently through the entire chronology. If a document appears out of sequence in a way that doesn't make sense, check the original. More often than you would expect, the answer is a misread date in the metadata.

Drafting Cross-Examination and Impeachment Sets

This is where AI assistance becomes simultaneously most useful and most dangerous.

An impeachment sequence is a craft. Done well, it commits the witness to the false or inconsistent position, then confronts them with the prior statement, then closes every exit. It depends entirely on precision: the question has to track the prior statement closely enough to make the confrontation clean, without alerting the witness prematurely that something is coming. You are building a box before the witness knows they are standing in one.

AI can help construct the architecture of that sequence. Given two prior inconsistent statements, a well-structured prompt will generate a line of questioning that moves from neutral background toward the commitment point and then to the confrontation. That architecture is worth having, especially when you are preparing five or six witnesses simultaneously and running on seventy-two hours of available prep time.

Prompt structure for impeachment drafting:

> *"I have two statements from the same witness. Statement A is from a deposition on March 14, 2023 (verbatim text below). Statement B is from a declaration filed September 9, 2023 (verbatim text below). Draft a cross-examination sequence, no more than 25 questions, that: (1) commits the witness to the substance of Statement B without revealing that you have Statement A, (2) establishes that the witness had the same knowledge at the time of Statement A, and (3) confronts the witness*

with Statement A using the exact language of the prior testimony. Questions should be closed-ended. Number them. No commentary between questions."

What comes back will often be structurally sound. The question order will be logical. The commitment sequence will be recognizable. Experienced deposition lawyers will see what the AI is doing and will want to adjust phrasing, add foundation questions, and sharpen the confrontation line. That is normal. Expected, even. The AI gives you scaffolding. You build the actual cross.

What you must verify — and this is not optional — is whether the questions accurately reflect the testimony. Language models sometimes generate questions that misquote, over-read, or subtly distort prior statements. "Isn't it true you told the board the projections were accurate?" is a fundamentally different question from "Isn't it true you told the board the projections had been reviewed?" If the witness said the second thing and you ask the first, they will correct you. Loudly. In front of a court reporter and opposing counsel.

An AI-generated question that misrepresents prior testimony is not merely ineffective. It alerts the witness to the gap in your theory. It hands opposing counsel a redirect opportunity. And in circumstances where the misrepresentation concerns a material fact that subsequently appears in a court filing, it creates professional responsibility exposure under Model Rule 3.3. The connection to Mata v. Avianca is direct. The same failure mode that caused counsel in that case to file fabricated case citations can, in the deposition context, cause counsel to work from a fabricated version of the record. The difference is one of scale. Not of kind.

The Review Obligation That Cannot Be Delegated

ABA Formal Opinion 512, issued in July 2024, addresses generative AI directly. It grounds the analysis in Model Rule 1.1 — competence — and Model Rule 5.3 — the supervisory responsibility lawyers bear over non-lawyer assistance. The opinion makes clear that using AI to assist in legal work does not transfer or reduce the attorney's professional responsibility for the product. The supervision obligation that applies to a paralegal applies equally to an AI system. If the AI produces a deposition outline, the attorney who uses that outline is professionally responsible for its accuracy. Full stop.

That framing matters acutely in the deposition context. A deposition outline is not a contract template where an error in a boilerplate clause might be caught during negotiation. It is a script for a live, adversarial proceeding where errors have immediate and irreversible consequences. The witness answers the question you actually asked. The record is what the court reporter types. There are no redlines, no tracked changes, no second drafts after the fact.

The review protocol is the floor, not the ceiling.

For every question based on prior testimony: pull the transcript, find the passage, read it. Verify that the question accurately reflects what the witness said — not what you or the AI understood them to have said, not the fair inference, the actual words.

For every impeachment citation: verify the page, the line, the verbatim language. If the plan is to hand the witness a transcript page and ask them to confirm their prior words, that page had better be exactly right. An incorrect quote is a gift. The witness will take it.

For every document reference in the outline: verify the exhibit number, the date, the author. AI-generated document references are particularly susceptible to transposition errors — the right document, the wrong Bates range, or the right Bates range, the wrong document. Either one is a problem.

This review should happen after the AI-assisted draft is complete, and it should be done in one uninterrupted pass by the attorney who will be taking the deposition. Not by a junior associate auditing citations in the abstract. By the attorney who will be in the room.

That last point is not about distrust of junior colleagues. It is about how preparation actually works. Reading through the outline line-by-line against the record is how the deposing attorney internalizes the material. It is the preparation. Not a step that precedes preparation — the thing itself. Delegating that review to someone else, AI or human, is delegating the preparation. You can outsource the drafting. You cannot outsource the knowing.

Confidentiality at the Upload Screen

One more obligation before the workflow begins.

Deposition prep materials are saturated with client confidences. Prior testimony may contain privileged work product. Document chronologies reveal your factual theory of the case. Impeachment sets show what you know and when you learned it. Before any of this material enters an AI tool, you need to know — actually know, not assume — what that tool does with uploaded data.

Model Rule 1.6 prohibits disclosure of information relating to the representation without client consent. Uploading a deposition transcript to a consumer AI product that retains training data is a disclosure. ABA Formal Opinion 512 flags this explicitly: lawyers must review provider terms of service and confirm that client data will not be used to train models or retained beyond the session before uploading anything tied to a client matter.

Enterprise arrangements with a data processing agreement and a written no-training commitment are defensible. Consumer-tier products, absent that contractual protection, are not. If your firm has licensed an enterprise tier of the relevant tool, use that. If you are not sure which tier your firm has, find out before you paste anything. That answer should take one email and thirty minutes. The alternative — discovering after the fact that your theory of the case went into a training dataset — takes considerably longer to address.

A Realistic Workflow

Put it together and the sequence looks like this.

You receive the deponent's prior transcript and the relevant document export from review. You upload the transcript to your AI platform under the firm's enterprise account. You run the topic-extraction prompt, asking for verbatim quotes and page-line citations with no summarization. You run the chronology prompt on the document set. Results come back in minutes, not hours.

Then you sit down with the transcript and you check every citation the AI returned. What you find: most of the passage-level references point to the right general area of the transcript, but page-number precision is inconsistent; some quotes have been lightly paraphrased in ways that blunt their impeachment value; a small number — call it three to five out of sixty — reference passages that simply do not say what the AI indicated. Those numbers are not hypothetical. They reflect the realistic accuracy range you should expect from current tools when processing a 300-plus page transcript, even with retrieval augmentation enabled.

You correct every error. You delete the phantom citations. You reorganize the topic clusters based on your theory of the case rather than the AI's neutral extraction order. You add the questions you were already planning — the background questions, the relationship questions, the foundational questions that will matter if this ever reaches a jury. You build the exhibit binder to match.

What the AI gave you was structure and a head start. Four or five hours of paralegal work replaced by forty-five minutes of machine output followed by two hours of attorney review. The math is real. The preparation you produced is also real. Those are different things, and keeping them distinct is the professional discipline that makes AI assistance work without making it a liability.

The deposition room is yours. Prepare accordingly.

CHAPTER X

Litigation Strategy: AI as a Thinking Partner, Not a Decision-Maker

Think about the last time you needed someone to argue the other side. Not a colleague nodding along in a status meeting, not a partner who has too much invested in the case theory to see its holes — someone willing to hit back hard. For most solo practitioners and small-firm litigators, that person doesn't exist. You build your case, you file your case, and you encounter the opposition's arguments for the first time in their brief.

That is the gap AI can partially fill.

Not as an oracle. Not as a replacement for the judgment you've built over years in courtrooms and depositions and settlement conferences. As a structured sparring partner — one that will argue the other side at midnight, flag the case you missed, stress-test the damages theory you've been quietly worried about, and do all of it without billing six hundred dollars an hour.

The distinction matters enormously. An oracle tells you what will happen. A sparring partner helps you think harder about what might. Every serious AI application in litigation strategy belongs in the second category. The moment you start treating AI output as a prediction — "the model says we have a 70 percent chance at summary judgment" — you have left the territory of useful and entered the territory of dangerous.

This chapter gives you the practical architectures. But it starts with the frame, because getting the frame wrong produces worse outcomes than not using AI at all.

The Sparring Partner Frame

ABA Formal Opinion 512, issued in July 2024, addressed generative AI directly under the Model Rules. It didn't ban the tools. It didn't endorse them either. What it did was remind practitioners that competence under Rule 1.1 — requiring "legal knowledge, skill, thoroughness and preparation" — now extends to understanding the benefits and risks of relevant technology. Comment 8 to Rule 1.1, which the ABA formally adopted in 2012 and which requires lawyers to keep abreast of changes in the law and its practice including the benefits and risks of relevant technology, is cited in Opinion 512 as applying with full force to AI systems.

The opinion is worth reading in full. One passage captures the strategic AI question precisely: AI tools may assist lawyers in performing tasks they would otherwise perform themselves, but the lawyer remains responsible for the work product. That sentence does the heavy lifting. AI-assisted litigation strategy is permissible. The attorney's professional judgment is not outsourceable.

The sparring-partner model honors that structure. You use AI to surface arguments, stress-test positions, identify gaps. You evaluate what comes back with your knowledge of the judge, the jurisdiction, the record, and the client. You decide. Every step — the prompting, the evaluation, the decision — runs through your professional judgment. AI is the preparation tool. You are the attorney.

Mata v. Avianca is the cautionary tale about what happens when the frame breaks down. The lawyers in that case treated ChatGPT output — fabricated citations, entirely invented holdings — as if the model had done the legal work. They submitted it without verification. They were sanctioned by Judge P. Kevin Castel of the Southern District of New York, fined a total of $5,000, and referred to disciplinary authorities. The failure wasn't that they used AI. The failure was that they substituted AI output for their own professional judgment and filed it into a federal court.

Litigation strategy is a higher-stakes application than document drafting, because flawed reasoning plays out in courtrooms, not just in redlines. Hold that in mind before every prompt.

Stress-Testing Your Legal Theory

Every case has a load-bearing assumption — a factual inference, a legal interpretation, or a causal chain that the whole theory rests on. Find it early and test it hard, or opposing counsel finds it for you in their summary-judgment brief.

AI is genuinely useful here. Not because it knows your case better than you do — it doesn't — but because it generates systematic challenges faster than you can write them yourself. You give it the theory. You ask it to break the theory. You evaluate whether its attacks are serious.

The key word is systematic. A capable adversary will test your theory from at least five angles: the legal standard (does your claim actually require what you think it requires?), the factual predicate (are the facts you're relying on disputed, thin, or inadmissible?), the damages nexus (do your damages flow causally from the defendant's conduct?), the affirmative defenses (what are the most plausible escape routes?), and the credibility of your key witnesses. AI can generate robust responses on the first three. It has nothing to offer on the fifth, because it has never watched your witness testify. More on that shortly.

Prompt Architecture: Theory Stress-Test

Here is a prompt template that works for this purpose. Adapt it for your jurisdiction and claim type.

> *Role: You are an experienced civil litigator representing the defendant in the following case. Your job is to identify every serious weakness in the plaintiff's theory and develop the most compelling arguments for summary judgment or directed verdict. Case summary: [Insert 3–5 paragraphs describing your client's claims, the key facts, the governing legal standard, and the jurisdiction.] Task: Argue the other side. Identify the top five challenges to the plaintiff's case under [jurisdiction] law, ranked by strength. For each challenge: (1) state the legal argument, (2) identify the specific factual gap or legal weakness it exploits, (3) suggest what evidence or authority supports it. Do not*

> *fabricate citations. If you are uncertain whether a specific case exists, say so explicitly. Format: Numbered list with subheadings.*

Several features of that prompt are deliberate. Assigning the adversarial role shifts the model toward sharper counterarguments than simply asking "what are the weaknesses here?" Providing the full case summary in your own words forces you to articulate your theory clearly, which is itself useful discipline — many litigators have caught a gap in their reasoning while drafting the case summary for the prompt, before the model even responds. And the explicit instruction not to fabricate citations doesn't eliminate hallucination risk, but it reduces the model's tendency to produce confident-sounding case names when it's working from pattern rather than actual retrieval.

You will still verify every legal proposition against Westlaw or Lexis before it appears in any work product. Non-negotiable. The stress-test is for issue identification, not citation gathering.

After you run the prompt, score each challenge the model generates. Some will be obvious — you've already addressed them. Some will be legally weak given your specific jurisdiction. But occasionally one will land. It will identify something you've been avoiding, or surface an affirmative defense you'd deprioritized. That's the value. Not that AI is smarter than you, but that it's faster at generating a full array of attacks, and it doesn't care whether the attack is inconvenient.

Generating the Counterargument

Brief-writing is adversarial by nature, but most attorneys draft their own brief before seriously engaging opposing arguments. The better practice: draft your brief, then force a rigorous encounter with the other side before you file. AI compresses the time that takes.

The counterargument workflow differs slightly from the stress-test. You're no longer testing your overall theory — you're testing specific arguments in a brief that's already substantially drafted. Think of it as a pre-filing adversarial review.

Prompt Architecture: Adversarial Review

> *Role: You are a federal appellate litigator assigned to oppose the following brief. Read the argument section and prepare a point-by-point rebuttal. For each argument made by the movant: (1) identify the strongest counterargument under [jurisdiction] law, (2) identify any logical gaps or unsupported assumptions in the movant's reasoning, (3) suggest what additional facts or authority the nonmovant should surface. Do not fabricate citations. Brief excerpt: [Paste the argument section of your draft, or a structured summary of each argument.] Jurisdiction and procedural posture: [Insert.]*

What you get back is a simulated opposition brief. Not a real one — the model doesn't know what the actual defendant will argue, doesn't have access to facts favorable to the defense that you haven't shared, and will miss jurisdiction-specific nuances a local practitioner would catch immediately. But it will catch thin

syllogisms, unsupported transitions, and spots where you've assumed a legal proposition that requires more support.

Run this on your summary-judgment briefs, motions in limine, and appellate arguments. On any document where a court scrutinizes your reasoning rather than just your facts, pre-filing adversarial review is worth the twenty minutes it takes. The underlying cognitive phenomenon here is well-documented: researchers call it the illusion of explanatory depth, the consistent finding that people believe they understand their own reasoning better than they actually do. An external challenge — even an artificial one — surfaces the limits faster than re-reading your own work.

Building a Persuasive-Authority Survey

Your case is in the Seventh Circuit. The Seventh Circuit has no binding authority on the precise question you're litigating. You need persuasive authority — decisions from other circuits, well-regarded district court opinions, possibly a state court decision on a connected state-law question.

Traditionally, you'd run a string of searches in Westlaw or Lexis, read through results, and assemble a survey manually. A capable research attorney can do this in three or four hours. AI compresses the ideation phase — identifying which circuits, which doctrinal angles, and which case clusters are most likely to yield useful authority — to minutes. The verification phase remains exactly as long as it always was.

Prompt Architecture: Persuasive Survey

> *Task: I am litigating [describe the legal issue] in [jurisdiction]. There is no binding authority in this circuit. Please identify: (1) the circuit split or circuit consensus on this question, if any; (2) the leading cases from other circuits that would be most persuasive to a federal court in [jurisdiction]; (3) any Supreme Court dicta or secondary authorities — Restatements, law review articles by named scholars — that courts have found influential on this question. For each item, explain why it would be persuasive, not just what it holds. Important: Do not fabricate case citations or holdings. If you are uncertain about a specific case's existence or holding, flag it explicitly.*

Here is what you should not do with the output: cite any case that comes back from this prompt without running it through Westlaw or Lexis first. Every one. Mata v. Avianca happened because attorneys trusted AI-generated citations. The fabricated cases in that filing — Varghese v. China Southern Airlines, Zicherman v. Korean Air Lines as cited, among others — had names that sounded real, holdings that sounded plausible, and reporters that didn't exist. Judges remember.

What you should do with the output: use it as a search roadmap. If the model identifies a Ninth Circuit cluster on a particular theory, run that theory in Westlaw. If it mentions a law review article by a named scholar, verify the article exists before citing it. The AI is pointing toward likely productive terrain. You do the actual digging.

For persuasive-authority work specifically, legal-native platforms with retrieval-augmented generation — Westlaw Precision, Lexis+ AI, Casetext — are meaningfully more reliable than general-purpose models. Those platforms retrieve from actual verified databases and surface source text you can check directly. The hallucination risk isn't eliminated, but it's substantially lower when the model is grounded in a verified corpus rather than generating from statistical pattern. Choose your tool based on the stakes of the research task.

Settlement Range Modeling

Settlement is the end state of most civil litigation. The Federal Judicial Center's data, tracking federal civil cases consistently over the past two decades, puts the trial rate below four percent. The decision of whether to settle — and at what number — is among the most consequential judgment calls in litigation. It is also precisely the decision that AI is least equipped to make for you.

That said, AI can materially improve the analysis that informs your judgment.

Verdict databases — Bloomberg Law's Verdict Analyzer, VerdictSearch, Jury Verdict Research — have existed for years. The contribution AI makes is synthesis. You feed a model the key variables in your case — jurisdiction, claim type, defendant profile, damages categories, liability posture — and ask it to reason through how those variables interact with the pattern data you've gathered. You're not asking the model to predict your verdict. You're asking it to help you build a more rigorous settlement analysis than you might produce under time pressure with a client meeting in two hours.

Prompt Architecture: Settlement Range Analysis

> *Task: I am evaluating settlement in the following case. Please help me structure a systematic settlement analysis. Case summary: [Insert claim type, jurisdiction, key facts, liability posture, damages categories.] Verdict data I have gathered: [Paste or summarize relevant comparable verdicts from your database searches, including jurisdiction, claim type, plaintiff verdict amounts, and notable case-specific factors.] Specific requests: Based on the case summary and verdict data: (1) identify the factors most likely to drive my case above or below the midpoint of the comparable range; (2) identify any factors in my case that lack direct comparables and suggest how courts have handled similar outliers; (3) draft a one-page settlement analysis memo I can share with the client, framed as a range with identified variables, not a prediction. Note: This analysis is confidential attorney-client work product.*

The memo output gives you a structured starting point. Review it carefully. The model doesn't know your client's risk tolerance, your opposing counsel's history of going to trial, the judge's case-management preferences, or the specific weaknesses in your damages expert's methodology. Those variables matter as much as the verdict database.

Before you paste any client facts into an AI tool for settlement modeling, verify your firm's data-handling policy. Opinion 512 is explicit: competence under Rule 1.1 includes understanding the data-retention and

confidentiality terms of the AI tools you use. Consumer-grade AI products with broad training-data retention policies are not appropriate vehicles for live client matter facts. Chapter 12 of this book covers the Rule 1.6 analysis in detail.

What AI Cannot See

The most important limitations of AI in litigation strategy are not about hallucination. Hallucination is a fixable quality-control problem — verify outputs, don't rely on unconfirmed citations, done. The deeper limitations are structural. They won't be solved by the next model version.

AI cannot assess witness credibility. Full stop. It has never watched your key witness testify in a deposition, never seen them struggle under cross, never noticed the tells that suggest their damages narrative is softer than it looks on paper. The entire credibility dimension of litigation strategy — which witnesses to call, which to avoid, how a particular jury pool in your jurisdiction will respond to your client — is invisible to every AI tool currently available.

AI does not know your judge. Every experienced litigator knows that local rules, judicial preferences, and courtroom temperament shape strategy as much as doctrine does. Judge Richard Posner, during his tenure on the Seventh Circuit, was well-documented as departing from doctrine toward economic analysis in ways that surprised practitioners who read only the hornbooks. No AI model has a reliable representation of how the specific judge assigned to your matter handles spoliation sanctions, expert qualification disputes, or late-filed summary-judgment replies. You do. Use it.

AI has no knowledge of what opposing counsel will actually argue. The adversarial prompts described in this chapter generate plausible opposition arguments. Not accurate ones. Skilled litigators choose strategically which arguments to raise and which to save. They have confidential information that shapes their approach. AI generates something like a statistical average of how a case resembling yours tends to be litigated. The attorney across the table is not a statistical average.

And AI cannot model the client relationship. Settlement decisions, strategy pivots, and risk-tolerance questions run through a client who has their own financial pressures, emotional investment in the outcome, and appetite for risk that no model has access to. Those conversations are yours.

Drawing the Line

Here is the rule, stated plainly: AI informs litigation judgment. It does not exercise it.

When you use an AI sparring partner to stress-test your theory, you are preparing for the conversation with your client, the motion to the court, the letter to opposing counsel. You are not delegating the decision about what to argue. When you run an adversarial review of your brief, you are stress-testing your own reasoning. You are not outsourcing the judgment about which arguments to make and how forcefully. When you build a settlement range analysis with AI assistance, you are improving the structured analysis you bring to a client

meeting. You are not providing the client with a machine prediction.

Opinion 512 makes this structure mandatory. It specifically addresses supervisory responsibilities under Rule 5.3, noting that attorneys using AI to assist with legal tasks carry the same obligations they'd have with a junior associate performing the same work. An unsupervised AI output filed or sent without attorney review isn't AI-assisted work. It's abdication. The opinion draws no distinction based on how sophisticated the AI model is or how confident it sounds.

The practical test is simple. If you can explain, in a privilege log or a conversation with a client, exactly what judgment you exercised in evaluating and using the AI's output — what you accepted, what you rejected, and why — you've used AI as a thinking partner. If you can't explain that because you largely accepted the output without independent evaluation, you've used AI as a decision-maker. One is competent practice under Rule 1.1 as the ABA has now explicitly interpreted it. The other isn't.

The capabilities described in this chapter are real and worth taking seriously. Stress-testing theories, running adversarial brief review, surveying persuasive authority, structuring settlement analysis — all of these tasks are faster and often more rigorous with a capable AI tool than without. Solo practitioners and small firms especially should pay attention. The gap between a large firm that can assign four associates to stress-test a summary-judgment theory and a two-lawyer shop that can't has narrowed. Not closed. Narrowed.

The frame is everything. You are the attorney. The AI is prepared to argue. You decide what to do with the argument.

That's not a limitation of the technology. It's the definition of what litigation judgment is.

CHAPTER XI

Compliance, Ethics, and the Model Rules: What the Bar Has Actually Said

The bar's formal guidance on AI arrived the way most important professional changes do — slowly, then all at once. For nearly two years after ChatGPT's launch in November 2022, practitioners were largely improvising. No binding opinions. Conflicting informal guidance. A technology that moved considerably faster than any standing committee could chase. Then 2024.

The ABA issued Formal Opinion 512 in July of that year, the first formal ABA opinion to address generative AI directly. New York's Task Force on Artificial Intelligence released its report the previous April, running to more than 150 pages. California's State Bar had already published practical guidance in November 2023. Florida followed with its own advisory. None of these sources invented new obligations. Every rule they applied already existed. What they did was translate familiar professional duties into a context where the risks look different — because the mechanism producing the work product is invisible to the reader and routinely wrong in ways that sound authoritative.

That translation matters. The operative language from these opinions is reproduced here precisely so you can quote it in firm policy memos, CLE presentations, and engagement letters. Do not paraphrase ethics opinions when you can cite them directly. Courts and bar disciplinary panels can read.

Rule 1.1: Competence Means Understanding What the Tool Actually Does

Comment [8] to Model Rule 1.1 has anchored the technology competence debate since the ABA amended it in 2012. A lawyer must "keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology." Formal Opinion 512 treats this comment as the foundation for everything that follows, and builds on it with specificity the comment never anticipated.

The Opinion makes clear that competence in using generative AI means more than knowing which application to open. A lawyer must understand, at a functional level, how these models generate output — specifically, that they are statistical prediction engines producing plausible text, not retrieval systems that surface verified facts. That distinction is not academic. An attorney who cannot explain why an LLM confidently cites a case that does not exist is not competent to supervise the output of that tool on a client matter.

Opinion 512 concludes that lawyers must understand AI tools sufficiently to exercise independent professional judgment about whether AI-assisted work product is reliable before it reaches a client, a

counterparty, or a court. The floor it sets is meaningful. You are not expected to explain transformer architecture. You do not need to know how to run a fine-tuning job. But you must understand: that these models do not search legal databases the way Westlaw does; that confident tone is entirely decoupled from accuracy; that newer model versions shift hallucination patterns rather than eliminating them; and that any output involving case citations, statutory text, or regulatory provisions must be independently verified against authoritative sources before use.

Practical implication: if you are using an AI tool and cannot describe in a sentence how it retrieves or generates legal citations, you are not yet competent to use it unsupervised on client matters. That is the standard Opinion 512 requires. The California State Bar's November 2023 guidance, Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law, reaches the same position from a slightly different angle — competence requires evaluating not just the AI's output but the AI's process, including whether the tool grounds its legal content in a curated authoritative database or generates it from pattern matching across unvetted training data.

For firms drafting AI-use policies, Rule 1.1 sets the minimum training threshold. Before any attorney uses an AI tool on a client matter, they should be able to articulate how the tool produces its output and identify its known failure modes. "I watched a demo" does not satisfy this standard. Documented training does.

Rule 1.3: Diligence and the Quiet Risk of Feeling Done

Diligence — the Rule 1.3 obligation to "pursue a legal matter on behalf of a client... with reasonable diligence and promptness" — rarely makes headlines in the AI ethics discussion. It lacks the drama of hallucination or confidentiality breach. It deserves more attention than it gets.

The diligence risk from AI is structural. These tools make work feel complete faster than it is complete. A GPT-class model produces a research memo in ninety seconds that reads finished — clean, organized, cited. The psychological pull toward accepting it as finished is real and well-documented in early practitioner surveys. Opinion 512 does not address Rule 1.3 at the same length as competence or confidentiality, but its underlying logic runs through the entire opinion: professional obligations are satisfied not by the production of output but by the production of accurate, reliable output after independent professional review.

The problem compounds in high-volume, deadline-driven practices. Discovery review, due diligence, brief-writing against a filing deadline — these are exactly the contexts where AI is most useful and where the temptation to underreview is strongest. Rule 1.3 requires that you pursue the matter diligently even when AI makes it tempting to move on. In an AI-assisted workflow, diligence means reviewing AI work product with the same critical rigor you would apply to a first-year associate's research memo. Not your own ten-year judgment. Theirs.

Rule 1.4: Communication and the Disclosure Question

Does your client need to know you're using AI? That depends on the jurisdiction, the tool, and specifically what the client has asked. This is one area where bar guidance genuinely varies.

Model Rule 1.4 requires that a lawyer "keep the client reasonably informed about the status of a matter" and "explain a matter to the extent reasonably necessary to permit the client to make informed decisions." Formal Opinion 512 approaches disclosure carefully. It does not mandate blanket disclosure of AI use in every matter. What it does say is this: if a client specifically asks whether AI was used, the lawyer must answer honestly. And if AI use might affect the billing, confidentiality, or accuracy of the work product in ways a reasonable client would consider material, disclosure is required — not optional.

New York's Task Force Report took a harder line. Its recommendations included client notification when AI is used in substantive legal work, on the basis that clients have a legitimate interest in knowing whether their matter involved a technology that carries documented accuracy risks. The Task Force stopped short of requiring disclosure in every context, but the direction of travel in New York is toward more transparency.

California raises a scenario worth internalizing. If a client engagement agreement specifies that identified attorneys will perform the work, using AI tools to handle substantive portions without disclosure may implicate Rule 1.4 independently of any billing question. Read your existing engagement letters with fresh eyes before your firm's AI deployment scales up. If they commit specific attorneys to the work, they may create a disclosure obligation your current policy does not address.

The practical fix is straightforward: draft a paragraph for your standard engagement letter describing your firm's AI use in terms clients can understand. Not a disclaimer — a description. What AI tools are used, what categories of work they assist with, and what human review occurs before AI-assisted product reaches the client. This satisfies Rule 1.4, reduces future disputes, and signals transparency rather than evasion.

Rule 3.3: Candor to the Tribunal and the Sanction That Changed Everything

This is the rule that bites hardest, because the consequence is not a disciplinary proceeding down the road. It is a sanctions order now, in front of the judge in your case.

Rule 3.3(a)(1) prohibits a lawyer from "knowingly mak[ing] a false statement of fact or law to a tribunal." Rule 3.3(a)(3) adds that a lawyer must not "offer evidence that the lawyer knows to be false," and — critically — if a lawyer discovers during a proceeding that material submitted to the tribunal was false, they must take "reasonable remedial measures," including disclosure.

Mata v. Avianca applied this framework publicly. The full facts are examined in Chapter Two; the Rule 3.3 analysis deserves specific attention here. Judge P. Kevin Castel's June 2023 sanctions order in Mata v. Avianca, Inc., No. 22-cv-1461 (S.D.N.Y.), found that attorneys Steven A. Schwartz and Peter LoDuca of Levidow, Levidow & Oberman had submitted a brief citing six cases that did not exist — all ChatGPT hallucinations. The $5,000 sanctions were almost beside the point. The lasting damage was the finding that

counsel had failed to perform the basic verification any reasonably competent attorney would complete before representing to a federal court that binding authority exists.

Formal Opinion 512 addresses Rule 3.3 directly, concluding that lawyers must verify the accuracy of AI-generated legal research before submitting it to a tribunal. This is a professional obligation, not a best practice. The opinion also extends the candor obligation beyond citations: any legal proposition derived from AI output and presented to a court carries the same verification requirement, whether or not it takes the form of a case citation.

Several courts have responded with local rules requiring certification that AI-assisted content in filings has been reviewed and verified for accuracy. Check your local rules before filing. The absence of a specific local rule in your jurisdiction does not create an absence of obligation; it means Rule 3.3 carries the weight alone.

The categorical rule for any AI-assisted brief or motion is this: every citation must be independently verified against Westlaw, Lexis, or an equivalent authoritative source before filing. Not spot-checked. Every one. The ninety seconds saved in drafting is not worth discovering a fabricated citation at 4:45 on a Friday before a Monday deadline.

Rules 5.1 and 5.3: The Supervision Framework

The supervision rules are where Opinion 512 is most useful as a policy-drafting resource, because the opinion provides a familiar conceptual model: treat AI output the way you treat work product from a capable but unvetted non-lawyer assistant.

Rule 5.3 requires lawyers with supervisory authority to take reasonable steps to ensure that a non-lawyer's conduct "is compatible with the professional obligations of the lawyer." Opinion 512 explicitly applies this to generative AI: a lawyer who uses AI to produce work product bears the same supervisory responsibility they would bear if a paralegal or contract researcher produced it. The AI carries no license. It exercises no professional judgment. The attorney does both, and remains responsible for both.

What does reasonable supervision look like in practice? The Opinion identifies several relevant factors: the complexity and stakes of the matter, the known reliability characteristics of the specific tool, and whether the attorney has actually reviewed the output rather than passed it along. A form status update drafted with AI assistance requires less supervisory scrutiny than an AI-drafted motion for summary judgment. Neither requires zero scrutiny. None.

Rule 5.1 adds the management layer. Partners and supervising attorneys must maintain firm-wide systems — not just individual habits — to ensure that AI tools are used consistent with professional obligations across all personnel. A solo practitioner decides her own AI protocols. A managing partner at a forty-attorney firm is responsible for ensuring that the associate two floors down is not submitting unreviewed AI output to clients or courts. Firm-wide AI-use policies are not merely good practice. Under Rule 5.1, they are a professional obligation.

The supervision framework also draws the delegation line clearly. Drafting a first cut of a contract, summarizing deposition testimony, running a preliminary jurisdiction survey — AI can meaningfully assist with these, subject to attorney review. Exercising judgment about which facts are legally material, advising a client on litigation risk, deciding trial strategy — these cannot be delegated under Rules 5.1 and 5.3 because they require independent professional judgment that only the licensed attorney can supply.

What the State Bars Have Added

The ABA's opinion sets a national floor. The states build on it, and some have built considerably higher.

New York. The NYSBA Task Force on Artificial Intelligence spent eighteen months on its April 2024 report before releasing its recommendations. On competence, the report echoes Opinion 512 and adds practical specificity: attorneys should understand a given tool's data sources, training cutoff date, retrieval architecture, and known error rates before deploying it on client matters. On disclosure, the Task Force recommended — though the recommendation has not been codified as a formal ethics rule — that attorneys disclose AI use in substantive matters when clients request it. The Task Force also addressed metadata and data retention directly, noting that attorneys must investigate whether AI platforms retain input data, and for how long, before transmitting any client-related content through them.

California. The State Bar's November 2023 guidance makes an unauthorized-practice connection that Opinion 512 treats as implicit: attorneys must not represent AI-generated work as entirely their own independent analysis when the client or tribunal would reasonably expect exactly that. The guidance is also specific about AI-identified issues. If AI flags a legal question the attorney had not independently spotted, the attorney must investigate that issue; they cannot simply pass through the AI's identification as though it were their own analysis. Discovery by AI does not substitute for attorney judgment about what to do with the discovery.

Florida. The Florida Bar's 2024 guidance largely tracks the ABA framework, with specificity on fee transparency addressed separately in Chapter Thirteen. For Rule 1.1 purposes, the Florida guidance emphasizes the ongoing nature of competence: as AI tools change, attorney understanding must change with them. Competence with GPT-4-era tools in 2023 does not automatically confer competence to deploy agentic AI systems with substantially greater autonomy in 2026. The Florida guidance treats competence as a continuing acquisition rather than a threshold crossed once.

Three Scenarios, Applied

Abstract rule analysis is only as useful as its application. Three scenarios follow, each mapping directly to the guidance surveyed above.

Scenario One. A third-year associate uses a general-purpose LLM to draft a summary of indemnification provisions in a commercial contract and emails it directly to the client without supervisory review. The summary contains one material inaccuracy — a stated liability cap the contract does not actually contain. Rules implicated: 1.1 (associate's competence in reviewing AI output), 1.3 (failure to diligently verify), 5.1

and 5.3 (supervising partner lacked systems to catch this). Under Opinion 512's supervision framework, the firm — not just the associate — has an ethics problem.

Scenario Two. An attorney discovers, during final pre-filing review, that three cases cited in an AI-drafted brief do not exist. She replaces them with verified authority and files. No violation occurred. The supervisory obligation was fulfilled before the filing. This is precisely the scenario Rules 3.3 and 5.3 are designed to produce. Document that the review occurred and that corrections were made before filing. That documentation matters if the methodology is ever challenged.

Scenario Three. A client asks directly whether AI was used in their matter. Under Formal Opinion 512 and the New York Task Force report, the answer must be honest. If AI was used, say so. You may explain the scope — the tools involved, the scope of their use, the review that followed — but the answer cannot be "no" when the answer is "yes." Rule 1.4 requires candor with clients, not just with tribunals.

Building a Compliance Framework That Actually Works

The guidance reviewed here points toward a short practical checklist for any firm drafting or updating an AI-use policy. Start with Rule 1.1: document the training your attorneys and staff receive before using AI tools on client matters, and build annual refreshers into the firm calendar as the tools evolve. Address Rule 1.4 in your engagement letters with language that is accurate about how AI is used — not boilerplate so broad it means nothing. Build supervisory review checkpoints — mandatory, documented — into every AI-assisted workflow that produces client-facing or court-facing output. And review Rule 3.3 compliance requirements with every litigator in the firm before they use AI to draft any submission to any tribunal, not once in a firm-wide meeting but at the matter level, regularly.

The rules have not changed. The conduct required to comply with them has. That is the entire point of every opinion examined in this chapter. The obligations are familiar. The contexts in which attorneys can violate them are new. The bar expects practitioners to have done the work of understanding the difference — and to have the documentation to show it.

CHAPTER XII

Confidentiality and What Not to Paste: Rule 1.6 in the Age of LLMs

Every lawyer who has been sanctioned in the last decade for an email error, a misdirected fax, or a discovery production that included privileged documents will recognize the pattern immediately. The tool is new. The mistake is ancient. You move fast, you paste something in, and six months later you're explaining yourself to bar counsel.

That is the real AI risk for most practicing lawyers. Not hallucinated citations — though those are serious and Chapter 2 covered them at length. The more immediate danger is a partner at a mid-size firm in Chicago who opens Claude or ChatGPT at 11 p.m., pastes in three pages of due diligence findings to get a summary, and never reads the platform's data retention policy. Or an associate who drops a client's settlement demand into a consumer AI chatbot to "clean up the language." Or a paralegal who runs a confidential deposition transcript through a free AI summarization tool because it's faster than anything the firm pays for.

These are not hypotheticals. In April 2023, Samsung engineers used ChatGPT to help debug proprietary semiconductor code. They pasted source code into the chat. They shared confidential meeting notes. The data left Samsung's control and entered OpenAI's servers before anyone in legal or compliance had reviewed what the default data-handling terms actually said. Samsung banned external AI tools within weeks. The legal profession noticed — and then, largely, kept pasting.

This chapter is about stopping that. Specifically, it maps Model Rule 1.6's confidentiality obligations to the actual data architectures of the tools you are likely already using, sorts those tools into three risk tiers, and gives you a practical checklist to run before any client matter touches an AI platform. The analysis applies whether you are a solo practitioner or a partner at an Am Law 100 firm, because Rule 1.6 does not grade on a curve.

What Rule 1.6 Actually Requires

Model Rule 1.6(a) prohibits a lawyer from revealing information relating to the representation of a client without informed consent or a recognized exception. The phrase "relating to the representation" is deliberately expansive. It covers more than privileged communications. It includes deal terms, litigation strategy, client identity in sensitive matters, financial information, and any other fact gathered in the course of the engagement — even facts that might be publicly available elsewhere. Comment 3 to the Rule makes this explicit: the duty applies to all information relating to representation, "whatever its source."

Rule 1.6(c) adds an affirmative obligation: "A lawyer shall make reasonable efforts to prevent the inadvertent or unauthorized disclosure of, or unauthorized access to, information relating to the representation of a client." The "reasonable efforts" standard is not a guarantee of perfect security. It is a negligence-style inquiry calibrated to circumstances. Comment 18 enumerates the relevant factors: the sensitivity of the information, the likelihood of disclosure if safeguards are not used, the cost of available safeguards, how difficult those safeguards are to implement, and whether the safeguards would materially impair the lawyer's ability to serve clients.

This framework was developed in the context of email, cloud storage, and remote access. It applies without modification to AI tools. The question is not whether using an AI platform is categorically prohibited — it is not — but whether the specific tool, under its actual terms of service and data practices, meets the "reasonable efforts" standard given the sensitivity of the information being processed.

That is a factual inquiry. Which means you have to read the policy.

The Disclosure Question Nobody Is Asking

When you paste client information into an AI platform, you are transmitting data to a third-party server. That transmission is a disclosure event for purposes of Rule 1.6, even if the information never appears in a publicly accessible place. Unauthorized access by the provider, downstream breach, training-data exposure, or regulatory production in a foreign jurisdiction — any of these could constitute unauthorized disclosure without a single additional act on your part.

The confidentiality analysis splits into three distinct risks, which most lawyers are conflating into one:

Training exposure. Some consumer AI tools, under their default terms, may use conversation content to improve models. If client facts are in that content, they potentially become part of a training dataset accessible to a future model's outputs. This risk has been reduced for many major platforms after public pressure, but the default terms vary and change without notice.

Data retention. Even if a platform does not train on your inputs, it may retain them for days, months, or indefinitely for operational, debugging, or legal-compliance purposes. Retention creates a breach surface and a potential third-party subpoena target.

Jurisdictional exposure. Where are the servers? If your client is a U.S. government contractor with data handling restrictions, a healthcare entity subject to HIPAA, or a party to litigation with active discovery obligations, the geography of data storage matters. Some consumer AI platforms use globally distributed infrastructure. Their terms of service specify which jurisdiction's law governs.

None of these risks require a hack or a data breach to create an ethics problem. The disclosure may occur the moment you transmit — if the tool's retention policy is incompatible with your confidentiality obligations and you never obtained client consent.

Three Tiers of Risk

The tools available to lawyers in 2026 sort cleanly into three risk categories. The framework is not about brand names — it is about data architecture and contractual protections.

Tier One: Consumer AI (Highest Risk)

Consumer-facing products — the free or low-cost personal accounts you access through a web browser — carry the highest confidentiality risk for legal work. This includes the free tier of ChatGPT, the standard consumer version of Claude.ai, Google Gemini accessed without a Workspace enterprise agreement, and any similar product used under personal consumer terms of service.

The risk is structural. Consumer terms of service are written to serve the platform's product interests, not yours. They typically reserve broad rights to retain, review, and potentially learn from user inputs. Opt-out mechanisms exist for some platforms, but they are not always enabled by default, they do not retroactively cover prior sessions, and they depend on you reading the fine print before you paste.

More fundamentally: consumer AI products do not offer data processing agreements. They do not accept HIPAA Business Associate Agreements. They do not commit to data deletion timelines in enforceable contracts. You have no legal remedy if their handling of your client's information creates downstream harm, beyond whatever their terms of service permit — which is usually very little.

The rule is simple. Consumer AI and client-specific information do not mix. Not for a quick summary. Not for a late-night research shortcut. Not "just this once." The Samsung engineers were not malicious. They were efficient. That is exactly the failure mode.

You can use consumer AI for fully hypothetical work — drafting template language with no client facts, learning how to prompt, running through research frameworks on fictional scenarios. The moment the information is specific to a real client matter, you need a different tool.

Tier Two: Enterprise API Arrangements with Data Processing Agreements (Manageable)

The second tier covers commercial AI platforms accessed through enterprise licenses or API agreements that include enforceable data protections. This means OpenAI's enterprise and API tiers, Anthropic's commercial API with a privacy addendum, Microsoft Azure OpenAI Service under an enterprise subscription, Google Workspace with Gemini under enterprise terms, and legal-native platforms like Harvey, Lexis+ AI, and Westlaw Precision.

What distinguishes this tier is the contract. A proper enterprise data processing agreement (DPA) will typically commit the vendor to: not using your inputs to train models, limiting data retention to defined periods, maintaining data isolation from other customers, providing deletion on request, agreeing to applicable regulatory regimes like HIPAA or GDPR where relevant, and accepting data processing liability

under applicable law.

ABA Formal Opinion 512, issued in July 2023, addresses this distinction directly. The opinion holds that lawyers must "review the [AI tool's] terms of service and privacy policies to understand how information submitted to the tool will be used." It further notes that professional-grade tools with appropriate contractual protections can be compatible with Rule 1.6 confidentiality obligations — but only if those protections are actually in place, verified, and understood. The opinion is explicit that the reasonableness of using any AI tool with client data is a fact-specific inquiry that requires actual inquiry.

Enterprise arrangements are "manageable" rather than "low risk" because they depend on execution. The DPA has to actually exist. Someone at your firm has to have read it. Your firm's AI policy has to specify which tools fall into this category. And the sensitivity of the specific information still matters: a run-of-the-mill commercial lease review has a different risk profile than drafting a whistleblower complaint or processing a client's medical records in a personal injury matter.

Legal-native platforms deserve a specific note here. Harvey, Lexis+ AI, and Westlaw Precision are designed for law firms, and their commercial terms reflect that. They typically include enterprise-grade DPAs as standard, prohibit training on customer data, and provide audit trails. They are not immune from breach, but the contractual architecture is built for the profession. That matters. It is not a reason to stop thinking — it is a reason to rank them ahead of general-purpose consumer tools when both would serve the task.

Tier Three: On-Premise or Firm-Hosted Deployments (Lowest Risk)

Some large firms and institutional legal departments are deploying AI models on infrastructure they own or control. This includes running open-source models like Meta's Llama series on local servers, configuring private instances of commercial models through Microsoft Azure's virtual network isolation, or building proprietary tools on top of vendor APIs with zero data egress to third-party systems.

This approach eliminates most of the external disclosure risk. Client data processed on a firm-controlled server does not leave the firm's environment. There is no third-party data retention. There is no vendor's training pipeline. The risk profile looks more like the firm's internal document management system than a SaaS platform.

The tradeoff is resource intensity. Hosting capable LLMs requires significant compute infrastructure, IT expertise, and ongoing security maintenance. As of 2026, this is realistic for firms above roughly 200 attorneys and for sophisticated in-house legal departments with strong technology infrastructure. Solo practitioners and small firms are not the target audience for on-premise deployment — which is why the consumer-versus-enterprise distinction at Tiers One and Two matters so much for most of the bar.

What ABA Formal Opinion 512 Actually Requires You to Do

Opinion 512 is the ABA's primary guidance on attorney use of generative AI. It is worth reading in full, not in summary. But the core confidentiality obligation it establishes, applied to the three-tier framework above,

reduces to a checklist of specific inquiries.

The opinion holds that before using any AI tool with information relating to a client representation, the attorney must: understand how the tool stores and handles inputs, determine whether the tool's data practices are compatible with Rule 1.6, obtain client informed consent if the planned use exceeds what reasonable clients would expect, and ensure that any third-party AI vendor who handles client data is subject to confidentiality obligations comparable to those binding the attorney.

That last point echoes Rule 5.3's supervision obligation — you are responsible for the conduct of non-lawyers, which includes AI platforms processing your data, in the same way you are responsible for a paralegal or contract reviewer.

The opinion also addresses client consent. Rule 1.6(a) allows disclosure with client "informed consent." Where a client has not generally consented to AI-assisted work, you may need to obtain specific consent before running their information through even an enterprise-grade platform. Some engagement letters now address this explicitly. If yours do not, that is a gap worth closing. Standard language addressing AI tool use, the data protections in place, and the client's right to opt out can be incorporated into retainer agreements without friction.

What the State Bars Have Added

Several state bars have issued guidance that adds obligations on top of or alongside Opinion 512. The NYSBA's AI Task Force report, released in January 2024, addressed confidentiality at length and recommended that New York attorneys treat AI tools that lack enforceable data isolation as equivalent to unsecured cloud storage — subject to the same pre-use due diligence required by the state's cloud-computing ethics guidance. California's State Bar released practical AI guidance in late 2023 that emphasized the breadth of California Business and Professions Code section 6068(e), which imposes confidentiality obligations that are at least as demanding as Rule 1.6 and in some respects broader.

Florida has addressed cloud and AI tool use through advisory opinions that treat data retention and training use as disclosure events triggering confidentiality analysis. The uniform message across these jurisdictions is the same: the obligation runs to the practice, not to the platform. "I didn't read the terms of service" is not a defense.

Reading a Data Policy Before You Need To

Most lawyers will not read a twenty-page vendor privacy policy and DPA before pasting something in. That is a real constraint, and pretending otherwise produces useless guidance. What you need is a five-minute due diligence process — the questions that surface the risks that matter.

Before your firm approves any AI tool for use with client information, run through these:

Does the vendor use inputs to train models? This is the first question and the most important. If the default answer is yes, the tool belongs in Tier One regardless of branding.

What is the data retention period? How long do they keep conversation or processing logs? Is deletion on request available and enforceable by contract?

Is there a signed DPA available? Not a linked privacy policy — a signed, negotiated data processing agreement that creates enforceable obligations. Ask your vendor. If they cannot produce one, treat the tool as Tier One.

Who has access to your data on the vendor's side? Some vendors reserve rights for human review of conversations for safety, trust, or quality purposes. Know this before you assume automated processing only.

Where are servers located? Relevant for clients with regulatory data restrictions and for international matters.

What security certifications does the vendor hold? SOC 2 Type II and ISO 27001 are the baseline. Their absence is a flag.

What is the breach notification commitment? Under what timeline, and to whom?

This is the intake checklist. Run it once per tool, document the results, reflect those results in your firm's AI policy, and update the analysis when the vendor updates its terms — which they will, sometimes without prominent notice.

The Engagement Letter Fix

The simplest risk management step most lawyers are not taking is updating their engagement letters. A single paragraph, reviewed by your malpractice insurer and your ethics counsel, that discloses the firm's general use of AI tools, identifies the data protection category of those tools, and obtains the client's consent to that use changes the Rule 1.6(a) analysis entirely.

Informed consent does not require a client to understand the difference between a DPA and a consumer terms of service. It requires that you explain the nature of the tool, what data may be processed, the protections in place, and the client's option to object. Clients who understand the explanation and proceed have consented. Clients who object can be served without AI-tool processing of their specific information.

This is not hypothetical exposure management. Several malpractice carriers began asking in 2024 and 2025 whether firms had AI policies and client disclosure procedures in place. The absence of such procedures is increasingly a factor in coverage analysis for technology-related claims.

The Discipline the Rule Actually Requires

There is no AI confidentiality problem that is technically unsolvable. Tier Three deployments exist. Enterprise DPAs exist. Client consent provisions are a paragraph in an engagement letter. The problem is behavioral, not technical: lawyers move fast, tools are convenient, and the consequences of a paste error are

invisible until they are not.

Rule 1.6(c) requires reasonable efforts. Reasonableness is contextual. A solo practitioner drafting a low-stakes commercial contract for a small business client operates in a different risk environment than a partner handling an SEC whistleblower matter or a healthcare system's litigation defense. The tool that is acceptable for the first matter may be categorically off-limits for the second. The analysis is not one-size-fits-all, which means the analysis has to actually happen — not once, at the firm level, but for each matter, each tool, each category of information.

The Samsung engineers were not reckless. They were trying to do their jobs faster, and the tool was right there. Every lawyer reading this has had an identical moment. The difference between a Samsung-style incident and clean practice is not superior moral character. It is a decision, made in advance, about which information goes where — and a firm commitment to honor that decision at 11 p.m. when the deadline is real and the shortcut is tempting.

The bar will not be sympathetic to convenience as a defense. The client will not be sympathetic either. And unlike a hallucinated citation — which is correctable through verification — a confidentiality breach cannot be unpaused. The information is out. The question then is not whether to fix it, but what to tell the client, how to report it, and what the consequences are for the representation.

That is a conversation no attorney wants to have. The intake checklist is how you avoid it.

CHAPTER XIII

Billing and AI: The Candid Conversation

The math arrives before the coffee is finished. A partner at a mid-size litigation firm assigns a second-year associate a motion to compel. Last year that task took twelve billable hours, minimum. This year, with the firm's AI drafting platform doing structural scaffolding and initial argument framing, the associate finishes in four. She corrected three confident-sounding legal errors the tool inserted. She added deposition-specific facts no model could have known. She formatted the citations properly, because the AI got two of them slightly wrong. Four hours. Solid work.

The timesheet stares back.

That is the central problem of AI billing. Not hypothetical. Right now, replicated across thousands of practices at every billing cycle, usually without explicit firm guidance, bar instruction, or client communication. The billable hour as a fee structure was designed for a world where effort and output tracked together — where twelve hours of drafting produced twelve hours of value. AI has broken that assumption. The profession has not decided what to replace it with.

This chapter is not comfortable reading. There is no clean answer here. What you will find is what the Model Rules actually require, what the bars have actually said, what clients are actually demanding, and what the realistic options look like for a practice that intends to use AI both ethically and sustainably.

The Model Rule That Governs This

Model Rule 1.5 requires that fees be reasonable. That is the whole analysis. Everything else — the time-entry debates, the write-down conversations, the alternative-fee arguments — flows from one word in the rule.

Reasonableness under Rule 1.5(a) is assessed against a list of factors: the time and labor required, the novelty and difficulty of the questions involved, the skill required, the customary fee in the community, the experience and reputation of the lawyer, and the results obtained. Notice what that list does not say. It does not say you bill whatever hours you logged. It says the fee must be reasonable given multiple factors, of which time is only one.

That framing matters enormously now. If a task that required twelve hours of skilled work last year now requires four hours of skilled work plus a competent AI tool, the relevant question is not "what is the customary rate times four?" The relevant question is whether the resulting fee is reasonable given the value delivered, the complexity of the matter, and the skill involved in supervising and correcting the AI output.

Sophisticated bar ethics committees have been careful not to say that AI efficiency must always translate into lower client bills. That would overreach. What they have said, consistently, is that billing twelve hours when

four were worked — because twelve is what the task used to take — is ethically indefensible. The distinction matters. Value-based billing has always been permissible. Fabricating worked time has never been.

What ABA Formal Opinion 512 Says About Money

ABA Formal Opinion 512, issued in July 2023, was the American Bar Association's first formal guidance on generative AI tools. Its primary focus was competence, confidentiality, and supervision. But the opinion engages billing directly, and its position on fees deserves careful reading.

The opinion draws on principles the ABA established in earlier technology guidance, particularly around what may be passed through to clients as costs versus what must be absorbed as overhead. A general enterprise AI subscription — the firm's platform license, used across all matters — is overhead, the same category as Westlaw access and word processing software. You cannot line-item it on a client invoice the way some firms tried to charge per-page copying rates in the 1990s. Those charges were rejected by clients and, in some jurisdictions, by courts. The AI equivalent will face the same resistance.

Where Opinion 512 breaks new ground is in reinforcing what the profession has long accepted in principle but rarely had to apply at scale: the client is owed the benefit of structural efficiency improvements. Not every efficiency gain — no one argues that a faster typist's time savings must be rebated. But efficiency gains that are large, consistent, and traceable to a technology deployment the client did not negotiate into the engagement cannot simply be absorbed into an unchanged fee structure without raising Rule 1.5 questions.

The opinion stops short of mandating specific disclosure language. It creates no checkbox obligation. What it creates is an ethical environment in which billing twelve hours for a four-hour task, on the grounds that twelve was the old baseline, has no defensible anchor.

State Bars Have Been More Specific

California went further. The State Bar's November 2023 Practical Guidance on Generative AI addressed billing with unusual directness: lawyers using AI tools that reduce the time required to complete a task should not bill clients as if the task took the original, unaided time. The guidance frames this under Rule 1.5 and the California equivalent prohibition on clearly excessive fees.

The New York State Bar Association's 2024 AI Task Force report focused more on disclosure than billing mechanics, but the direction was consistent. Clients should generally know when AI was used in a significant way on their matter, particularly when the fee reflects that use. The NYSBA drew a sensible line: not every AI-assisted interaction — autocomplete, grammar check, basic formatting — rises to the level requiring client communication. But substantial use that materially affects the nature of the work product or the time required is a different matter.

Florida's guidance, consistent with the emerging consensus across jurisdictions, emphasizes that the attorney's professional product is the attorney's responsibility regardless of AI involvement, and that the fee must reflect

actual work performed rather than what would have been performed without AI.

Read together, these opinions do not require you to disclose AI use in every client communication. They do require that if AI materially changes the economics of your work, you cannot simply ignore that change when the invoice goes out.

The Time Entry Problem Is Real

Here is what actually happens inside most firms right now. An associate uses an AI drafting tool. The task takes three hours. The customary task, without AI, would have taken eight. The associate enters 3.0 hours with a description: "Draft motion for summary judgment." The billing partner reviews it, reduces it by a token half hour for good measure, and the client receives an invoice showing 2.5 hours. Nobody discussed the AI involvement. Nobody flagged it. The matter file contains no record of which tools were used.

That is legally defensible only if 2.5 hours at the stated rate is a reasonable fee for the work performed. At $400 per hour for a second-year associate at a regional firm, $1,000 for a supervised AI-assisted motion draft may be entirely reasonable. At $650 per hour for an associate at an AmLaw 100 firm, the analysis is more complicated — and more exposed.

The deeper problem is that most firms have not built time-entry guidance for this scenario. Associates who worked three hours but know the "old" version took eight are left to guess. Some bill actual time. Some bill the "task rate" — eight hours, on the theory that they are billing for value, not effort — and justify it internally as value-based billing without ever having disclosed that to the client. Some split the difference and feel vaguely uncomfortable about both options. None of them should be making this decision unilaterally.

Your firm needs written AI time-entry policy. Not a philosophy statement. A policy. Specifically: what description language is required when AI drafting or AI research tools materially contributed to a work product? Does the firm require disclosure in the time entry itself, in the billing narrative, or only upon client inquiry? What is the authorization threshold for billing AI-generated work at full associate rates without write-down review? Who signs off? These are not rhetorical questions. They are operational decisions that your managing partner, general counsel, or ethics partner should have made already.

If your firm does not have that policy, you are leaving individual timekeepers — mostly associates who need billable hours for their performance reviews — to make ethics decisions without guidance. That is not a tenable position.

Clients Who Know What AI Costs

Do not mistake client silence for client acceptance. The Association of Corporate Counsel, representing in-house lawyers at corporations and institutions across the country, has been vocal about outside firm AI practices since 2023. ACC members know what these tools cost. They know what they can do. Many are deploying the same tools internally and watching their own teams' task times drop.

A sophisticated general counsel at a large company is not going to believe that a document review her team's AI platform could process in three days required six weeks of associate time. She will ask. She will negotiate. She will win, because she is your client, she controls the relationship renewal, and she has every right under Rule 1.5 to challenge unreasonable fees.

The ACC and its chapters began circulating model outside counsel guidelines language on AI billing in 2024. The specifics vary, but the direction is uniform: disclose when AI was used, adjust fees to reflect AI-generated efficiency gains, and do not charge research fees for information that AI retrieved in seconds. Some guidelines go further, requiring firms to certify that time entries reflect actual attorney and staff time, with AI-generated work either excluded or separately noted.

This is not abstract pressure from a trade association. Firms with substantial institutional client bases — large corporations, private equity sponsors, insurance carriers, government entities — are already receiving outside counsel guidelines that address AI billing. If yours does not yet contain that language, it will. The question is whether you want to be ahead of that conversation or behind it.

The Write-Down Conversation

Write-downs have always existed. Associates over-bill. Tasks run longer than expected. Billing partners cut invoices before they go out. Nothing new.

What is new is systematic pressure to write down AI-assisted work at the time-entry stage rather than the billing review stage, and client-side auditing that identifies specific tasks for retroactive adjustment. Some clients are demanding pre-emptive write-downs in engagement letters. Others are auditing invoices after receipt, flagging tasks that AI assistance would obviously have compressed, and deducting the difference unilaterally before paying.

The deduction methodology is aggressive. If a client believes a forty-page brief should have taken an AI-assisted associate ten hours instead of twenty-two, the client deducts twelve hours of associate time from the invoice. The correct number might be fourteen or eight, but the client's leverage is not the number — it is the relationship. You need that client relationship more than you need those twelve hours. The write-down happens.

The write-down pressure is most acute in four areas: first-draft contract work, where AI is widely known to produce serviceable starting drafts in minutes; basic legal research on settled questions of law, where AI survey tools are a known commodity; document review in standard litigation, where AI-assisted review platforms have been in commercial use since the mid-2010s; and routine client correspondence — emails, status updates, demand letters — where AI drafting is essentially universal.

If your invoice shows eight hours of associate time on any of these tasks and the client employs a general counsel who attended any legal technology conference in the past two years, expect a conversation.

The intelligent response is to have that conversation before the invoice arrives. Not apologetically — proactively. Tell clients how your firm uses AI, how you supervise AI output, and how you have calibrated your billing to reflect AI-generated efficiency. That conversation is immeasurably easier before the first invoice than after.

Can You Bill for the AI Tool Itself?

Firms are experimenting with three approaches to recovering AI platform costs. The first is overhead absorption: treat AI subscriptions the way you treat Westlaw and Word, build them into operating expenses, and price accordingly. The second is rate adjustment: modestly increase hourly rates to account for AI platform costs while billing fewer hours. The third is disbursement line-items: charge AI processing as a direct pass-through, similar to expert witness fees or document production costs.

The disbursement approach is the most legally questionable and the most commercially toxic. Clients have been pushing back on technology surcharges since the 1990s, when some firms tried to bill "computer research time" as a separate charge on top of database subscription costs already embedded in rates. Courts rejected those charges. Clients rejected them. The profession moved on.

Billing a per-matter AI processing fee is the same error in new packaging. Unless your engagement letter specifically discloses and explains the charge — and the client has affirmatively agreed to it — a line-item AI fee is both ethically vulnerable under Rule 1.5 and commercially likely to generate disputes that damage the relationship and invite scrutiny of every other line on the invoice.

The more defensible approaches are the first two. Absorb AI costs as overhead and reduce time billed accordingly, or modestly adjust rates with a corresponding reduction in hours. Both are transparent. Both can be explained. Neither creates the adversarial dynamic of a surprise technology surcharge on a client who thought they were getting standard legal services.

Alternative Fee Structures Actually Solve This

Here is the clean answer. It has been available for thirty years without the profession fully embracing it.

Fixed fees, success fees, subscription arrangements, and value-based pricing all eliminate the billable-hour perverse incentive in one step. If you charge a flat $7,500 to handle a commercial lease review, it does not matter whether your AI platform reduces the internal time from eighteen hours to four. Your margin improves. Your client gets certainty and a known cost. Nobody has a billing ethics problem, because there are no hours to dispute.

The legal industry's resistance to alternative fee arrangements is well-documented and not entirely rational. Clients — particularly institutional clients — consistently express preference for fee certainty. Clio's annual Legal Trends data has shown for multiple years running that fixed-fee arrangements correlate with higher client satisfaction scores and faster invoice payment than hourly billing. The resistance is internal: partners

whose compensation ties to collected billings are wary of capping their upside, and associates whose performance reviews track billable hours have no incentive to be efficient.

AI is the forcing function that may finally tip the balance. As AI compresses task times across more and more practice areas, the economics of hourly billing become structurally uncomfortable. A fixed-fee contract practice where AI drafts first cuts and attorneys supervise, correct, and finalize is a high-margin, scalable operation. The same work billed hourly, with AI efficiency either hidden (ethically problematic) or disclosed (margin-compressing), creates constant friction with clients and constant internal anxiety about time entries.

Not every practice area fits fixed-fee models cleanly. Complex litigation with uncertain scope, criminal defense, and regulatory matters with unpredictable government timelines are genuinely difficult to cap. But routine transactional work — standard commercial leases, employment agreements, NDAs, entity formations — fits well. Standard litigation tasks with defined scope fit. Client-service subscription arrangements for small and mid-size businesses, which provide access to certain services for a monthly retainer, are growing precisely because AI makes them economically viable for the firm.

If your practice has not had a serious internal conversation about alternative fee structures in the AI era, that conversation is overdue. Not because of ethics pressure alone. Because the math increasingly points that direction.

The Disclosure Question in Plain Terms

Two positions circulate in the profession on disclosure, and both are too simple.

The first: you must proactively disclose AI use to clients in every significant matter. The second: AI is just another tool, like a form bank or a grammar checker, and you no more disclose it than you disclose your word processor.

The emerging bar consensus — expressed in ABA Formal Opinion 512, the California Practical Guidance, and the NYSBA Task Force report — points to a middle ground. Disclosure is not required for every AI interaction. It is required when AI use materially affects the nature, cost, or quality of the work product, and when a reasonable client in the circumstances would want to know.

A brief drafted substantially by AI, reviewed and corrected by an attorney, and billed at attorney drafting rates: disclose. A contract redline generated by an AI platform and reviewed line by line before partner sign-off: probably disclose. An AI-assisted grammar pass on a memo the attorney personally researched and wrote: no reasonable client expects notification of that.

When in doubt, disclose. The cost of unnecessary disclosure is a brief client conversation that most clients will receive positively — it demonstrates transparency and sophistication. The cost of insufficient disclosure, if a client later discovers that AI heavily contributed to work product they paid attorney rates for, is a fee dispute, a bar complaint, or a malpractice allegation. That is not a close call.

What Mata v. Avianca Costs in Billing Terms

The 2023 SDNY sanctions in Mata v. Avianca are usually discussed in terms of professional embarrassment and competence failure. Judge Kevin Castel's order — requiring the filing attorneys to notify the judges on each fictitiously cited case, complete legal research courses, and pay financial sanctions — reads as a cautionary tale about verifying AI output.

The billing dimension gets less attention. It should not.

The client in that matter paid for legal research. The invoices presumably reflected research hours. The research produced citations to cases — Varghese v. China Southern Airlines, Shaboon v. EgyptAir, Petersen v. Iran Air, and others — that do not exist in any reporter. The client paid for work that was not just substandard but affirmatively false, generated by a tool the attorneys used without adequate supervision and then certified as accurate to a federal court.

That is not a Rule 1.5 billing problem in the technical sense. It is a worse problem: charging for work that caused direct harm. The billing lesson is that AI-assisted work billed to a client carries an implicit representation that the attorney performed the verification that professional standards require. An attorney cannot ethically charge for research they did not verify. Hours are not "worked" in any billable sense if the work product is fabricated output that the attorney filed without checking.

This is the connection between the hallucination problem and the billing problem. They are not separate ethical concerns. Every AI-assisted task that generates a client invoice is a task the attorney is implicitly representing as competently supervised. The verification obligation is not merely a malpractice issue. It is the billing justification.

A Practical Framework

Before you enter another time entry on an AI-assisted matter, your firm needs answers to four questions.

First: does your engagement letter address AI use and its potential effect on fees? Standard boilerplate engagement language predates generative AI by decades. It does not cover this. Update it — not with alarming disclosures that will unsettle clients, but with a clear, professional statement of how your firm uses AI tools, how you supervise them, and how your fee structure reflects that use.

Second: do you have written time-entry guidance specifying how AI-assisted tasks should be described, and what write-down thresholds apply? If associates are making these decisions individually without policy guidance, you have a compliance risk on every AI-assisted invoice.

Third: have you reviewed your active client relationships to identify which outside counsel guidelines now address AI billing? That review takes an afternoon. The consequences of missing a client guideline and billing in violation of it can take months to resolve.

Fourth: are your standard fee structures — hourly, fixed, or hybrid — calibrated to actual task times in an AI-assisted workflow, not the task times of 2021? If you are still pricing fixed fees based on old hour estimates, you are underpricing. If you are still billing hourly at old hour estimates, you have an ethics problem. Neither outcome is acceptable.

None of this requires a bar investigation or a client complaint to motivate action. The rules are clear enough. A reasonable fee for work actually performed. Disclosure when a reasonable client would want to know. No charging for time not spent. These are not new principles. AI has made compliance more complex — and made the consequences of ignoring them more visible, more quantifiable, and more likely to surface in a billing dispute before the year is out.

CHAPTER XIV

The Next Five Years: What to Watch, What to Ignore, and How to Stay Competent

The forecasters who predicted that large language models would replace lawyers by 2025 were wrong. The forecasters who predicted nothing would change were also wrong. Where we actually land, as of 2026, is messier and more interesting than either camp anticipated: AI has become a serious professional tool, not a replacement for professional judgment, and the next five years will sharpen that distinction rather than dissolve it.

This chapter does not offer a crystal ball. It offers a working map of what is genuinely likely to happen between now and 2031, what vendors will promise but fail to deliver, and what you need to do right now to stay competent as the landscape keeps moving.

The Signal: Four Developments Worth Tracking

Autonomous Agent Workflows

The most significant near-term shift is not better chat. It's agents.

An AI agent, in the technical sense, is a system that can take multi-step actions toward a goal — browsing documents, calling external APIs, writing and editing files, looping back to check its own work — without a human approving each individual step. As of 2026, early versions of this already exist. Anthropic's computer use capability, OpenAI's Operator, and several legal-native platforms are building what they call agentic workflows that can, for example, pull a contract from a data room, extract defined terms, cross-reference them against a checklist, and populate a due-diligence summary with minimal human intervention at the intermediate steps.

This is genuinely different from prompting a chatbot.

The productivity gains, in controlled settings, are real. A 2024 pilot at a large Am Law 100 firm — details not for attribution, but the results were discussed at ILTA's annual conference — found that an agent-assisted due-diligence workflow reduced first-pass contract review time by roughly 60 percent on a 400-document data room. That's not a rounding error. For transactional practices that bill those hours, or absorb the cost under fixed-fee arrangements, 60 percent is a competitive threat and an opportunity simultaneously.

But agents introduce a new category of risk that single-prompt interactions don't: compounding error. If an agent misclassifies a document at step two of a fifteen-step workflow, every downstream output inherits that error. You won't see a hallucination as a discrete wrong answer in a chat window. You'll see a due-diligence

schedule with systematically wrong information that looks clean because it was formatted by the same automated process that produced the error.

Your obligation under Model Rule 5.3, as ABA Formal Opinion 512 makes clear, is supervision of nonlawyer assistance. That rule does not distinguish between a contract paralegal and a software agent. The supervision requirement attaches to the output and the process alike. Before your firm deploys any agentic workflow on a client matter, someone with legal judgment needs to audit the intermediate outputs — not just the final deliverable. A spot-check at the end is not supervision. It's hope.

Expect this to be the dominant conversation in legal AI between 2027 and 2029.

Multimodal Contract Analysis

Text-only AI is already receding. The models available in 2026 — including the major general-purpose systems and several legal-native platforms — can process images, PDFs with embedded graphics, handwritten annotations, and tables with meaningful accuracy. That changes what legal AI can actually review.

Think about what a typical M&A data room actually contains: scanned leases with handwritten margin notes. Signature pages photographed on someone's phone. Excel exhibits embedded as images inside a Word document. Environmental reports with maps and site photographs. A text-only model simply couldn't engage with most of that material. A multimodal model, given a sufficiently high-resolution input, can.

The practical application closest to production-ready, as of 2026, is lease abstraction. Commercial real estate attorneys and their clients have been early adopters here. A multimodal system can read a scanned lease — not a clean digital PDF, but an actual scan — extract key economic terms, identify non-standard provisions, and flag handwritten addenda for human review. Platforms like Kira Systems, now part of Litera, have been doing machine-learning-based contract analysis for years. The LLM-generation versions are faster to configure and more flexible with novel clause language.

By 2028, expect multimodal analysis to be standard on any platform serious about M&A or real estate diligence. The caveat — and it is a significant one — is that OCR and image-interpretation errors do not announce themselves. A model that misreads "3%" as "8%" in a rent-escalation clause will present the wrong number with the same confidence it uses to present a correct one. Page-level human spot-checking of numerical outputs from multimodal review is not optional. It is the job.

Court-Specific AI Filing Rules

Courts have been faster to regulate AI than most practitioners expected. By early 2025, more than thirty federal district courts had issued standing orders addressing AI-assisted filing in some form. The approaches varied. The Northern District of California, the Eastern District of Texas, and the Fifth Circuit required attorneys to certify that AI-generated content had been reviewed for accuracy and that all cited authorities actually exist. Others required disclosure whenever generative AI was used in drafting a document filed with

the court. A few imposed blanket prohibitions on AI-assisted research citation pending further guidance.

This fragmentation is a practical problem. If you practice in multiple jurisdictions — or if you're a partner supervising associates working across circuits — you need a jurisdiction-by-jurisdiction tracking system, not a one-size-fits-all firm policy.

Between 2026 and 2031, expect consolidation. The Judicial Conference of the United States has been working on model AI disclosure guidance since late 2024. A uniform federal rule addressing AI-assisted filings is likely before 2029, though the precise form is genuinely uncertain. Several state supreme courts have already issued guidance — the Michigan Supreme Court's 2024 administrative order and Colorado's parallel guidance are the most detailed — and state-level standardization will follow.

The lesson from Mata v. Avianca is relevant here at a systemic level. Judge P. Kevin Castel's June 2023 sanctions order — five thousand dollars each to attorneys Steven Schwartz and Peter LoDuca, plus remedial obligations — was not primarily about AI. It was about the failure to verify what was filed. The courts that issued standing orders in 2024 and 2025 were, in many cases, explicitly citing Mata. Every filing rule issued since that case carries the same core requirement: the lawyer is responsible for what goes to the court, regardless of how it was generated. No rule changes that. None will.

Know your local rules. Know them before you file. This is basic professional hygiene, and the AI overlay doesn't make it harder — it just makes the consequences of skipping it more visible.

The Regulatory Environment: EU, U.S., and the Gap Between Them

The EU AI Act came into force in August 2024. Its phase-in schedule runs through 2027, but the relevant classification for attorneys is already clear: AI systems used "in the administration of justice and democratic processes" are classified as high-risk under Annex III of the Act. High-risk classification means mandatory conformity assessments, transparency requirements, human oversight obligations, and logging requirements that processors must maintain.

This matters to American attorneys in two ways. First, if your clients are EU-based or your matter involves EU-jurisdiction AI systems, you need to understand what your counterparties and clients are legally required to do with their AI tools. Second, the EU AI Act is the most detailed statutory framework that currently exists for regulating AI in legal contexts. U.S. legislators and regulators are watching it closely.

On the American side, the picture is more fragmented. No comprehensive federal AI statute existed as of mid-2026. The FTC has been active on AI-related deception claims under Section 5 of the FTC Act. Several state legislatures — Colorado's SB 205, Illinois's AEDT law, New York City's Local Law 144 — have enacted sector-specific AI regulations, primarily in employment and consumer contexts. None directly govern the use of AI in legal practice. The ABA's own guidance, including Formal Opinion 512, operates at the level of professional ethics rather than statutory regulation.

What is coming, with high probability, is some form of federal AI disclosure or transparency rule — likely sector-specific rather than horizontal — that touches legal services as a regulated profession. The timeline is

genuinely uncertain. It will not arrive before 2027 at the earliest. What will arrive sooner, and is already arriving, is state-level bar regulation incorporating AI-specific competence standards into CLE requirements and disciplinary frameworks.

California's State Bar issued a framework for AI use in legal practice in late 2023; the Florida Bar followed with detailed guidance in 2024; the New York State Bar Association's Task Force on AI published recommendations in April 2024 that form the basis for pending formal opinions. The trajectory is clear: bar authorities are moving from general statements about competence toward specific rules about disclosure, supervision, and verification. Stay ahead of your jurisdiction's guidance. It will become binding faster than you expect.

The Noise: What Is Not Coming on Schedule

Vendor marketing in legal AI has a consistent failure mode: announcing capabilities that exist in a laboratory setting and implying they are production-ready. You have been here before — with e-discovery's predictive coding promises in 2012, and with blockchain-for-contracts in 2018.

Here is what is not arriving by 2031, at least not in the form currently being sold.

AI that reliably delivers current, binding legal authority. The retrieval-augmented generation systems deployed by Westlaw Precision and Lexis+ AI are meaningfully better than raw LLM hallucination. They are not a substitute for verifying that a case is good law, that a statute has not been amended, and that the cited proposition actually appears in the cited source. Every vendor demo you see in the next three years will use a clean, well-indexed example. Your actual matters will not always be clean. Verify.

Fully autonomous research memos. Agents can draft research memos. They cannot exercise the judgment to know when a circuit split is unresolved, when a case that looks on-point is actually distinguishable on facts that matter, or when a state court is signaling a coming shift in doctrine. That judgment requires understanding of legal context that no model in current development reliably provides. A memo drafted entirely by an agent and submitted to a partner or a client without attorney review is a professional responsibility problem waiting to happen.

Real-time AI strategy during depositions or oral argument. Several startups are pitching real-time AI assistance — live transcription plus instant analysis plus suggested follow-up questions — for depositions. The technology works, in the sense that it produces output. The liability question of who is responsible when the AI suggests a question that inadvertently discloses work product, or misidentifies a contradiction in testimony, has not been worked through. By 2031, some version of deposition-support AI will be defensible and deployed. It is not there yet.

AI judgment on credibility, venue strategy, or jury selection. These are high-stakes decisions with minimal ground-truth data, serious sampling biases in historical databases, and outcomes affected by local culture that no model has been trained on adequately. Jury consultants with regional trial experience will still outperform any AI profiling tool for at least the next five years. Possibly longer.

A Personal Competence Framework

ABA Formal Opinion 512, issued in August 2024, states that attorneys have a duty of competence under Model Rule 1.1 that includes understanding "the benefits and risks" of relevant technology. The opinion does not specify what that understanding looks like in practice. That's your job to figure out. Here is a starting point.

The Skills Worth Building Now

Prompt literacy. Not prompt engineering in the software-developer sense — you don't need to write system prompts or fine-tune models. What you need is the ability to give an AI tool a clear, well-scoped instruction that includes the jurisdiction, the governing law, the relevant facts, and the desired output format. A vague prompt produces a vague output. "Draft an indemnification clause" produces something generic. "Draft a mutual indemnification clause governed by New York law, covering third-party IP infringement claims, with a carve-out for gross negligence, in the style of a SaaS subscription agreement" produces something you can actually use as a starting draft. That skill is teachable and worth practicing deliberately.

Verification habits. Every citation from an AI tool gets checked against the primary source before it goes anywhere. Not most citations. Every one. This needs to become as automatic as running a conflicts check. If you internalize nothing else from this book, internalize this: the absence of a hallucination in ten queries does not mean the eleventh query is safe. The model does not know when it's wrong.

Understanding retrieval architecture. You don't need to know how transformers work. You do need to know the difference between a model answering from its training data — high hallucination risk for legal citations — and a model answering from a retrieved, grounded document set, which carries lower risk but not zero. When a vendor tells you their system is "RAG-based," ask what corpus the retrieval is drawing from, how recently it was updated, and what the system does when it can't find a supporting document. Does it say so, or does it generate an answer anyway? Those questions distinguish products that are defensible from products that are not.

Data hygiene. Chapter 12 covers Rule 1.6 in detail. The competence point here is simpler: before you paste anything into an AI tool, ask what happens to that text after you hit send. Consumer AI tools — the free tier of any major LLM — typically retain inputs for model training unless you opt out. Enterprise arrangements with data-processing agreements are different. On-premise deployments are different again. One thirty-second habit — knowing which category your current tool falls into before you use it on a client matter — prevents the majority of inadvertent disclosure problems.

The Habits Worth Building Now

Set a quarterly review of the AI guidance issued by your state bar. Most bars publish opinions and advisory letters through their websites; signing up for email updates takes five minutes and keeps you from being surprised by a new disclosure requirement. The California State Bar, the New York State Bar Association, the

Florida Bar, and the ABA's Standing Committee on Ethics and Professional Responsibility are the four sources most likely to produce opinions affecting your practice, wherever you are licensed.

Run one AI tool deliberately, on a real but low-stakes matter, every quarter. Not to demonstrate you're keeping up. To keep the skill current. AI tools update faster than any other technology legal practitioners use — the version of any given platform you learned six months ago may have meaningfully different capabilities and risks today. The only way to know is to use it.

Read the court docket in Mata v. Avianca occasionally. The full record — the original motion, Judge Castel's show-cause order, the response brief, the sanctions opinion — is public on PACER. Rereading the sanctions opinion once a year takes twenty minutes. It is the clearest available account of what professional responsibility looks like when an attorney trusts AI output without verification. The names of the attorneys involved, the dollar amounts of the sanctions, the remedial CLE requirements — these details make abstract ethics obligations concrete. Concrete is how rules become habits.

The CLE Investments Worth Making

California's State Bar guidance from late 2023 explicitly connects the duty of competence under Rule 1.1 to the need for technology-specific education. Several state bars — California, Florida, New York — have begun offering or approving CLE credit for AI-specific programs. By 2028, expect AI competence CLEs to be a standard category in most jurisdictions.

The question is which programs are actually worth your time.

Useful AI CLE covers: how LLMs generate text and why that makes citation unreliable; specific tool evaluation frameworks; data retention and confidentiality analysis for common platforms; emerging court rules by jurisdiction; and ethics opinion walkthrough across Model Rules 1.1, 5.3, 1.6, and 3.3. Useless AI CLE is a vendor demo disguised as education, or a general overview of "AI's impact on the legal profession" that could have been written in 2019.

Ask before you register: does this program teach me to evaluate a specific tool's risks, or does it just describe what AI is?

The attorneys who will be well-positioned in 2031 are not the ones who adopted every new tool first. They are not the ones who refused to use any of them. They are the ones who built enough technical literacy to distinguish signal from noise, enough verification discipline to catch what AI gets wrong, and enough engagement with bar guidance to stay ahead of the regulatory curve.

That is, ultimately, the same thing that has always distinguished careful lawyers from careless ones: the ability to use available tools well while maintaining independent judgment about their limits.

AI is a powerful tool. It has real limits. Knowing both — with specificity, not in the abstract — is what competence looks like from here.