



Matthew Whibley

THE VARTAZARIAN LAW FIRM

Before you buy legal AI, learn to use the AI you already have

PRACTICE WITH GENERAL-PURPOSE SYSTEMS WILL HONE YOUR UNDERSTANDING AND SKILL BEFORE YOU INVEST



There is a growing market for “legal AI” products aimed at lawyers. Some of those products may be useful. Some may eventually become part of a good litigation workflow. But most plaintiff lawyers do not need to start there. Before spending real money on a customized legal-AI platform, lawyers should first learn how to use the general-purpose systems that are already sitting in front of them – ChatGPT, Gemini, Claude, and, in some situations, Grok.

That is not a contrarian point. It is a practical one. If a lawyer does not know how to get reliable work out of a general-purpose AI system, buying a more expensive interface usually just means paying more for the same mistakes. On the other hand, if a lawyer learns a disciplined workflow first, the lawyer can get substantial value right away and will be in a much better position to decide later whether any specialized product is actually worth the money. (State Bar of California, Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law (2023); ABA Formal Op. 512 (2024).)

The simplest way I know to explain this is with a law-library analogy. Imagine you are sent into a very large library and told to come back with the best authorities and the best draft on a narrow issue. If nobody tells you what section to start in, what jurisdiction matters, what procedural posture matters, or which books are already important, you are going to waste time. You may still find something useful, but a lot of your effort will be spent figuring out where to begin. If, by contrast, someone walks in with you, points to the right section, puts the most relevant

books on the table, and tells you what issue you are really deciding, the work gets better and faster.

That is what good AI use looks like. The biggest mistake lawyers make is the cold start. They open a chat, dump in some facts, and immediately ask for a complaint, a motion, a demand letter, or a deposition outline. Then, when the output is generic, incomplete, or wrong, they conclude that the tool is overhyped. Usually, the problem is not the tool. The problem is that the lawyer skipped the part that matters most: deciding what books belong on the table before the drafting starts.

I think the most useful way to teach lawyers to use AI is with a three-step workflow: priming, executing, and verifying. Those three steps are simple enough for a lawyer who is new to AI to use tomorrow, but they are also good enough that lawyers who use AI constantly will recognize why they work.

Step one: The priming phase

Priming is the most important step, and it is the step lawyers most often skip. By priming, I do not mean giving the AI a few facts and hoping it figures out the rest. I mean building the working context before you ask it to draft anything. Put differently, priming is the stage where you decide what books are on the table.

Utilize dedicated project workspaces

For practical legal work, I think lawyers should get into the habit of starting in a dedicated project or matter-specific workspace whenever the platform allows it. In plain English, a regular chat is more like an ongoing conversation with the system generally. Depending on your settings, it may draw on saved preferences, prior history, or other broad context. A project with project-only memory is narrower. OpenAI explains that when you choose project-only memory, previously saved memories are not referenced, chats can reference other conversations within that same project, and chats cannot reference conversations outside the project. (OpenAI Help Center, Projects in ChatGPT (2026).) That is the closest thing to saying, “This case stays in this case.”

That difference matters. If I am working on a motion involving a waiver at a sports facility, I do not want the system reaching into some unrelated employment matter, a vacation-planning chat, or another client file. I want the work to stay inside the matter. A project with project-only memory is the closest analogue to taking one client file into one conference room and shutting the door.

Claude projects and Gemini tools can accomplish similar practical goals, although they do it somewhat differently. Claude projects are built around a dedicated project space. Gemini’s Gems are better understood as reusable instruction containers, and Gemini’s Deep Research is the research engine. (Google,

Get Started with Gems in Gemini Apps (2026).) The practical lesson for the lawyer is the same: Do not mix matters casually. Create a matter-specific container and work inside it.

Define the specific legal issue

Once you are in the right container, the next part of priming is issue framing. This is where lawyers already have an advantage over nonlawyers. We know that legal questions almost never turn on the broad topic alone. The issue is not “waivers.” The issue is whether a particular waiver is enforceable under California law, in a negligence case, at the pleading stage, involving a sports or recreational defendant, with these specific facts, and with these specific bad facts. The issue is not “duty.” The issue is whether a duty exists under California law on this record, under this theory, against this defendant, at this point in the case.

If you do not tell the system that, it will often give you a broad law-school answer instead of a lawyer’s answer. So, the priming instruction should usually identify at least these things:

The jurisdiction.

The procedural posture.

The exact legal issue.

The good facts.

The bad facts.

The type of authority you want.

What you want first – research, not drafting.

That last point matters. At the priming stage, I usually do not want a draft yet. I want the law. I want the framework. I want the best cases for me, the best cases against me, the factual distinctions that matter, and the weaknesses I need to address before anything gets drafted.

Apply deep research for priming

This is where the stronger research modes matter. OpenAI’s Deep Research, Gemini’s Deep Research, and the heavier research modes in other systems are useful because they are designed to plan, search, reason, and synthesize before they answer. OpenAI describes Deep Research as a tool that can work with uploaded files, search the public web or specific sites, use enabled apps, and produce a documented report. (OpenAI Help Center, Deep Research in ChatGPT (2026).) Google describes Gemini Deep Research in similar terms: It can search the web and, if you choose, draw from sources like

Gmail, Drive, and uploaded materials to produce a multi-step research report. (Google, Use Deep Research in Gemini Apps (2026).)

That is exactly what I want for priming. I do not want the system rushing to write. I want it researching first. A plain prompt is usually enough. Something like this: “California law only. Published authority preferred. I am evaluating whether a pre-injury waiver bars a negligence claim arising from a sports or recreational facility. First give me the governing framework, the best authorities enforcing waivers, the best authorities limiting or refusing to enforce them, the facts that seem to drive the distinction, and any obvious weak points in the plaintiff’s position. Do not draft yet.”

That is not fancy. It does not need to be. It works because it does what a good supervising lawyer would do with a junior lawyer: Define the issue, identify the jurisdiction, give the legal lane, and say what work product should come first.

Avoid overloading with excessive data

One of the mistakes lawyers make when they start using AI is overloading the system with everything. Every record. Every email. Every transcript. Every idea. That feels thorough, but it is often the opposite of useful. Anthropic’s documentation makes the point clearly: The context window is the model’s working memory, and more context is not automatically better: As the token count grows, accuracy and recall can degrade. Anthropic refers to that phenomenon as “context rot.” (Anthropic, Context Windows (2026).) Its engineering materials make the related point that what matters is curating the optimal set of information at inference time, not flooding the model with everything you have. (Anthropic, Effective Context Engineering for AI Agents (2025).)

Again, the law-library analogy helps. If you cover the table with every book in the library, you have not helped the researcher. You have made the research harder. The point of priming is not to dump the whole warehouse into the system. The point is to put the right books on the table.

Step two: The execution phase

When I say “executing,” I mean drafting the actual document you want. This is the stage where you ask for the complaint, the motion, the deposition outline, the PMK notice, the separate statement, the discovery requests, the demand letter, or the medical summary. The important point is that execution comes after priming, not before it.

By the time you reach this stage, the system should already know the legal issue, the controlling jurisdiction, the relevant factual distinctions, the key authorities, and the problem areas. That changes the quality of the draft dramatically. Instead of drafting from a cold start, the system is drafting inside a preserved authority set. That is the phrase I think matters most: Drafting inside a preserved authority set.

That is what lawyers actually want. We do not want the AI free-associating across the universe of legal topics. We want it writing from the authorities and factual distinctions that were already identified in the research stage.

Priming improves drafting quality

The answer is not mystical. The model is working from a narrower, better-framed, better-supported context. Anthropic’s materials distinguish between the model’s training and its current context. Anthropic describes context engineering as the process of curating and maintaining the optimal set of tokens available to the model during inference. The context window is the text the model can reference when generating a response – its working memory – not the broader data it was trained on. (Anthropic, Context Windows (2026).)

That is why priming works. You are not “teaching the model new law” in the training sense. You are narrowing the working context so the model is more likely to reason from the right materials and less likely to wander.

Switching modes during the draft

After the research is done, I often do switch out of the heavier research mode for the drafting stage. But I think this needs to be explained clearly. The benefit of that switch is not that the mode switch itself somehow creates legal rigor. The rigor comes from the work that was

already done in priming. The reason the drafting stage still works is that you are staying in the same project, the same thread, or the same preserved workspace where the research, files, and earlier analysis remain available.

So, when lawyers talk about “switching to a faster mode,” I think the clearer way to say it is this: Once the legal groundwork has been laid, you may not need the heavier research engine to draft every paragraph. The draft stays grounded because the legal frame is already there.

Practical applications of the execution

For plaintiffs’ lawyers, this stage is immediately useful. Once the priming is done, these systems can help draft:

- complaints after the viable theories are narrowed;
- motions in limine after the precise evidentiary issue is framed;
- opposition sections after the best and worst authorities are identified;
- deposition outlines after the liability theory is settled;
- PMK notices after the corporate knowledge issues are identified;
- discovery requests after the proof gaps are defined;
- medical summaries after the chronology has already been understood.

The common mistake is asking the AI to decide the legal theory and draft at the same time. Those are different jobs. When you separate them, the drafting stage becomes much more reliable.

Step three: The verification phase

Verification is not optional. It is not an afterthought. It is part of the workflow. The ABA’s Formal Opinion 512 says lawyers using generative AI must fully consider duties that include competence, confidentiality, communication, supervision, meritorious advocacy, candor to the tribunal, and reasonable fees. The State Bar of California’s practical guidance says AI-generated outputs can be a starting point, but they must be carefully scrutinized, and that lawyers must critically review, validate, and correct both the input and the output. A lawyer’s professional judgment remains the lawyer’s responsibility. (State Bar of California, Practical Guidance for the Use of Generative Arti-

ficial Intelligence in the Practice of Law (2023).) That is exactly right.

Conduct cross-checks in fresh sessions

One useful habit is to take the authority list or the draft section from one session and drop it into a fresh session – sometimes even a different model – and ask that second system to attack it. Not because the second system can certify the law. It cannot. But it can often catch obvious problems quickly. It can tell you that a case looks fake, that the proposition seems overstated, that the quote appears suspicious, or that the authority does not seem to match the point being made. That can save time. I think of that second system as a first-pass cross-examiner, not as the final verifier.

The lawyer’s role in verification

The lawyer still has to do the real verification work:

- Confirm the authority exists.
- Confirm the citation is accurate.
- Confirm the quoted language is real.
- Confirm the case actually supports the proposition attached to it.
- Confirm the case is the right jurisdiction, the right posture, and the right publication status.
- Shepardize or KeyCite it.

That is not a defect in AI. That is lawyering. The mistake is expecting the system to eliminate the need for that work. It will not. What it can do, if used properly, is make the earlier stages of issue framing, authority gathering, and drafting materially faster.

When to purchase specialized AI

Sometimes the answer will be yes. If a specialized product materially improves document integration, source traceability, cite-checking, workflow control, or collaboration, it may absolutely be worth paying for. But that should be an informed purchase, not a panic purchase. Lawyers should first learn the basic method:

- put the matter in its own project;
- keep the case inside the case;
- prime the issue before asking for a draft;
- use the heavier research tools for research;
- draft only after the authorities are on the table;
- verify the work aggressively.

If you can do that well with a general-purpose system, then you will know exactly what you want a paid legal-AI tool to add. If you cannot do that well, buying a more expensive product will not solve the underlying problem.

Conclusion

I do not think most plaintiffs’ lawyers need to begin their AI journey by purchasing customized legal software. I think they should begin by learning a better workflow. The right analogy is still the law library. AI works best when the right books are already on the table. Priming is how you choose those books. Executing is drafting the actual document from that set of authorities and facts. Verifying is what makes the work usable in the real world. It is practical. It is available now. It does not require a sales demo. And, for many lawyers, it will produce a very large percentage of the value they are looking for before they ever spend money on a specialized platform. 💎

Matthew Whibley is a partner at The Vartazarian Law Firm. Previously, he toured the world in a Grammy-nominated punk rock band and then graduated Southwestern Law School at the top of his class.