What a law firm can hand an AI agent, and what it must never hand over

The four questions a firm has to answer before an agent touches a client file — privilege and confidentiality, which data can be sanitized, what to keep an eye on in the case law, and how to collect documents from clients without a hundred emails.

GUIDE · 2026-11-08 · 13 MIN READ

Who this is for

A managing partner, office manager or operations lead at a firm of two to fifty lawyers who is being told — by a vendor, by a client, or by their own frustration — that an AI agent should be doing more of the work. It is not a sales argument for agents. It is the boundary: what a firm can safely automate, what it must keep behind a wall, and the four decisions that have to be made before any of it starts.

This is not legal advice and it is not an ethics opinion. Every state's rules differ, and the answer for a criminal defense practice is not the answer for a real estate practice. What follows is the shape of the question, so the conversation with your malpractice carrier and your bar's ethics hotline starts in the right place.

The one rule everything else follows from

A client's information is confidential, and in most matters it is privileged. Handing it to a third-party tool is not automatically a breach — lawyers use stenographers, experts, e-discovery vendors and outside copy services — but it is a **disclosure to a third party** that has to be deliberate.

The ABA's Model Rule 1.6 governs confidentiality; Model Rule 1.1 Comment 8 says a lawyer should keep abreast of the benefits and risks of relevant technology. Most states have adopted something close to both. Neither tells you which tool is acceptable. That judgement is yours, and it is the exposure if it goes wrong.

So the practical question is never "is this tool safe." It is: **what data class is going into it, and does the arrangement survive the confidentiality obligation for that class.**

The four data classes, and what each allows

Sort the work before you sort the tools.

Class	Examples	What it allows
Public record	Filed pleadings, published opinions, court dockets, statutes	Anything. No confidentiality interest exists.

Firm operational	Your own templates, checklists, billing rates, matter-opening forms	Most tools, subject to who can see the account. No client information involved.
Client matter, sanitized	Facts with names, identifiers and privileged strategy removed	Usually workable with a vendor that does not retain or train on inputs, plus a written policy
Client matter, raw	The actual file: names, privileged communications, strategy, PII	The highest bar. Full document, defined purpose, no training, no retention, and a human accountable for each use

The mistake almost every firm makes is skipping the middle class. They either refuse everything, or they paste a client's entire intake narrative into a chat window because the tool was convenient. Sanitization is the skill that makes automation possible at all: a fact pattern with the names removed is often just as useful for drafting, summarising and analysis.

What a firm should ask a vendor, in writing

Six questions, and a vendor who cannot answer them in writing is the answer:

1. **Do you train on our inputs?** The only acceptable answer for client matter is no.
2. **How long do you retain it?** "Until we delete it" is not a retention policy.
3. **Where is it processed and stored**, and does that cross a border your clients would object to?
4. **Who at your company can read it**, and under what conditions?
5. **What happens to our data if we leave**, and can we export it?
6. **Will you sign a confidentiality addendum** that tracks our obligation to the client?

There is a longer version of this, with the specific regimes — HIPAA, §7216 for tax preparers, GLBA, PCI and the breach-notification clock — in the **agent access rulebook** in this library. Read that before a deployment, not after.

Where the real hours are, if the boundary is respected

Once the data question is settled, three workflows are worth automating in a firm of almost any size, and none of them requires an agent to touch privileged material.

1. Document collection from clients

The back-and-forth is the tax. A real estate closing needs the contract, the title commitment, the survey, the entity documents, the lender package. A tax controversy needs the notices, the returns, the correspondence. In most firms this lives in an email thread where the same document is requested three times and received twice.

A structured collection system does four things an inbox cannot:

- **One list, visible to the client**, showing what has been received and what is outstanding.
- **Reminders that fire on the calendar**, not on someone remembering to chase.

- **A file that lands in the right matter folder** with the date and the sender attached, so the chain of custody is obvious later.
- **A record of what was requested and when**, which is the document you want when a client says they were never asked.

This is the highest-return automation in a small firm and the least sensitive: the agent is moving files and sending reminders, not reading strategy.

2. Case law and rule change watch

Not a research tool — a **watch**. The difference is that research answers a question you already have, and a watch tells you about something you did not know to ask.

A watch is a standing list of sources (your jurisdiction's appellate opinions, a practice-area reporter, the rules committee, the bar's ethics opinions) checked on a schedule, with a short written summary of anything material and a link to the source. The output is not "here are 40 cases." It is "here are the two that affect your practice area this week."

This is also the only item on this list that a firm can start with **entirely public data**, which is why it is the safest thing to automate first.

3. Your own presence, rather than a directory's

Most firms' online presence is a profile on someone else's directory — a page whose layout they cannot change, whose leads they cannot see, and whose ranking they do not control. The firm pays for the privilege in referral fees or lost attribution.

A firm-owned page can do what a directory cannot:

- **Show the work by matter type**, in the language a prospective client actually searches — not the firm's internal practice-group names.
- **Capture the inquiry with the facts you need**, so the first call is a qualification rather than an intake.
- **Tell you which page produced which inquiry**, which is the difference between marketing and guessing.

Two constraints, and they are not optional: advertising rules vary by state and several regulate anything that could be read as a prediction of outcome; and any statement about prior results generally needs the accompanying disclaimer. Ethical marketing is narrower than marketing, and the narrow version is still worth more than a directory profile.

The audit

Nine questions. Every no is work that is either being done by a person who should not be doing it, or not being done at all.

Boundary

1. Can you name, for each tool the firm uses, which of the four data classes it is allowed to see?
2. Do you have a written answer from each vendor on training and retention?
3. Does anyone in the firm know what "sanitized" means in practice, or is it assumed?

Collection

1. Can you see, for any open matter, what has been requested and what is outstanding, in one place?
2. Does a missing document generate a reminder without a human remembering?

Intelligence

1. Is there a standing list of sources someone checks on a schedule, or does case law reach you by accident?
2. When something changes in your practice area, do you find out the week it happens?

Presence

1. If a prospective client searches your practice area in your city, does your own page compete with the directories?
1. Can you tell which page produced your last five inquiries?

What to do first

Do not start with the agent. Start with the classification:

1. **Write the four-class table for your own firm**, with actual tool names in the cells. One page.
2. **Send the six vendor questions** to every tool currently holding client information.
3. **Start the case law watch** — public data, no confidentiality question, immediate value.
4. **Then the document collection**, which is where the hours are.

The firms that get this wrong skip to four, hand an agent a client file, and discover the boundary at the worst possible moment.

IF YOU WANT THIS BUILT

We do this for a living: the audit, the build, and the person who answers when something breaks. The next step is a short call and it costs nothing.

prospyr305.com/book-with-franklin | (561) 990-2428 | scan for this page

Prospyr 305 · a subsidiary of Prospyr Inc · West Palm Beach, Florida. Free to read, print and share. General information, not advice for your specific situation.