

Before you sign: where will *our AI* run?

Twelve questions to ask any AI provider before you sign. They work for a five-person office or a large company, and for any kind of AI provider, including us.

How to use this list

- Send the questions in writing. Keep the answers with the contract.
- A good answer is specific. “Industry standard” and “we take security seriously” are not answers.
- “We don’t know yet” is fine early on. It should become a written answer before you sign.
- Bring in the people who own operations, finance and your data, plus your lawyer or compliance lead if you have one.

01 Where will the AI model run? ☐

Why it matters. “Cloud,” “private” and “on-premise” are not the same thing, and a “private” setup can be in your building or in a host’s data center. The answer tells you how many outside parties will handle your data.

A good answer. Names the setup plainly: the provider’s shared cloud, a dedicated server a host runs for you, or hardware you control.

02 Where is our data stored and processed, including backups? ☐

Why it matters. Data can be processed in one place, stored in another and backed up in a third. Location can change your risk and which rules apply.

A good answer. Lists each place by country or region, including backups and disaster recovery copies, and says whether that can change without notice.

03 Who can see our data? ☐

Why it matters. Access is often wider than buyers expect. It can include the provider’s support staff, engineers, contractors and automated review systems.

A good answer. Names the roles that can access your data, when, and how that access is approved and recorded. It also explains how you control access on your side.

04 What gets logged, and who can read the logs? ☐

Why it matters. Logs show who asked what and when. But logs of prompts and outputs can hold the same sensitive data as the documents themselves.

A good answer. Says what is logged, where the logs live, how long they are kept, and whether you can get a copy for your own records.

05 How long do you keep our data, and how is it deleted? ☐

Why it matters. Prompts, files, outputs and logs can each have a different retention period. “Deleted” can still mean “kept in backups for a while.”

A good answer. A retention period for each type of data, how deletion works, and written confirmation that data is returned or destroyed when the contract ends.

06 Will our inputs or outputs train anyone’s model? Can that change? ☐

Why it matters. Training on your data can carry it into a model other customers use. Terms of service can also change after you sign.

A good answer. A clear “no” in the contract, for inputs and outputs, with a promise to give notice before any change to that term.

07 Which other companies touch our data? ☐

Why it matters. Many AI services are built on other services: a model provider, a host, a support tool. Each is another party with its own terms, often called a subprocessor.

A good answer. A current list of every outside company that can receive your data, what each one does, and how you will be told when the list changes.

08 Who owns the outputs and what you build for us? ☐

Why it matters. If the provider builds workflows, prompts, connectors or a tuned model for you, ownership decides what you keep and can reuse.

A good answer. Says in writing who owns the outputs, the custom work and any trained model, and what the provider keeps as its own licensed software.

09 If we leave, what comes with us? ☐

Why it matters. Switching gets expensive when your data, settings and workflows only work inside one provider’s system. This is vendor lock-in.

A good answer. How you export your data, prompts, settings and workflows, in what format, how long you have after the contract ends, and any cost to leave.

10 Where does a person check the output, and how will we test it? ☐

Why it matters. AI can give confident answers that are wrong. Models also get updated, and an update can change the results.

A good answer. Shows where a person reviews the work before it reaches a customer or a record, how you can test on your own sample data, and how you hear about model changes first.

11 How fast will you tell us about a security incident? ☐

Why it matters. If something goes wrong, the hours after matter. Contracts can be vague about when and how you hear about it.

A good answer. A time limit in writing, a named contact, and what the notice will include: what happened, what data was involved and what is being done.

12 What will it cost to run a year from now? ☐

Why it matters. The first month is often the cheapest. Usage fees, seats, hosting, hardware, support and maintenance can add up later.

A good answer. Every ongoing cost and what makes it go up, such as more users, documents or usage. It also says who maintains the system after launch.

Where these questions come from

HHS guidance on HIPAA and cloud computing. For health data, HHS says a service level agreement can address system availability, backup and data recovery, how data is returned after the service ends, security responsibility, and limits on use, retention and disclosure. On data stored outside the U.S., it notes that the risks “may vary greatly depending on its geographic location.” It applies to HIPAA covered entities and their business associates. We use it here only as a model of good questions.

hhs.gov/hipaa/for-professionals/special-topics/health-information-technology/cloud-computing/index.html

OMB Memorandum M-25-22 (April 2025). Tells federal agencies to put certain terms in their AI contracts, including a permanent ban, unless the agency explicitly consents, on using non-public agency data and outputs “to further train publicly or commercially available AI algorithms,” and lock-in protections such as “knowledge transfer, clear data and model portability practices, clear licensing terms, and pricing transparency.” **M-25-22 binds federal agencies only. It does not apply to private companies.** It is still a useful model for what to ask.

whitehouse.gov/wp-content/uploads/2025/02/M-25-22-Driving-Efficient-Acquisition-of-Artificial-Intelligence-in-Government.pdf

If you handle sensitive data

Health data. If you are a HIPAA covered entity or business associate, HHS says a cloud provider that stores or processes electronic protected health information (ePHI) for you is a business associate, even if the data is encrypted and the provider has no key. You need a business associate agreement (BAA) with it before it handles that data.

Client files at a law firm. ABA Formal Opinion 512 (July 2024) says lawyers must evaluate the risk of disclosure before they put client information into a generative AI tool. For self-learning tools, it says the client's informed consent is required first. These questions help with that review.

This checklist is general information, not legal advice. For HIPAA, bar rules or other legal questions, talk to your own attorney or compliance advisor.

You do not need perfect answers to all twelve. You need answers in writing, and no surprises after you sign. To talk through where your AI should run, start with a free consultation at microve.ai/contact.html.