CLEMENZA LAW GROUP

NEW YORK, NY · (212) 365-4875 · CLEMENZALAW.COM

THE PRIVATE BRIEF · AI & IP

Reading an AI Vendor's Data Terms: The Six Clauses That Matter

ANTHONY CLEMENZA · MANAGING PARTNER

September 2026

IN BRIEF

- Six clauses decide whether an AI tool is safe for a given kind of information: training, retention and deletion, who else handles it and where, ownership of output, the promise to defend you against infringement claims, and exit and changes to the terms.
- The training clause surprises people most, because it is often written as permission to use your content to develop or improve the services, and because the same vendor's consumer and business versions can carry opposite defaults.
- Even when nothing is negotiable, the reading tells you which tier of your own policy the tool belongs in, and which questions to put to the vendor in writing.

Your operations lead forwards a link with one line: “Can we just sign up for this? Everyone’s already using it.” The link goes to a page of terms longer than your office lease, and the button at the bottom says Accept.

You will not negotiate that page, and you usually do not need to. What you need is to know six things before your team starts pasting: what the vendor may do with what you put in, how long it keeps it, who else sees it, who owns what comes out, who pays if what comes out infringes someone’s rights, and what happens when you leave or the terms change. Read those six and you know whether the tool can see customer information, internal drafts, or nothing but published material.

Before the six: what you have already promised

One question comes first. What have you already agreed to about this information?

Customer contracts, confidentiality agreements and data processing agreements (the contracts that govern how a service provider handles personal information for you) may limit who can process the other side’s information and where. If your profession has its own confidentiality rules, as law does, those apply too. A vendor’s terms can be perfectly reasonable and still be incompatible with a promise you made last year to a customer who has never heard of that vendor. So the vendor read is the second step. The first is a list of what you have already promised.

1. Will they use your information to train their models?

Terms do not always say “we will train our models on your data” in those words. The phrase to look for is some version of using your content “to develop or improve” the services, which is language that can cover training.

The trap is that the same product family can point in opposite directions. Some vendors’ business terms say they will not use customer content to develop or improve their services unless the customer explicitly agrees, and some say plainly that they may not train models on customer content at all. Meanwhile some vendors' consumer versions may use your conversations to train their models unless you switch that off in the settings. Same vendor, opposite defaults, and your team may be on either one.

The questions to answer:

- Is there an explicit statement that your content is not used for training?
- Is that protection the default, or a setting someone has to change, and who at your company owns that setting?
- Does it cover both what you put in and what the tool produces?
- Which version, consumer or business, is your team actually using?

2. How long do they keep it?

These are two questions that terms often blur. The first is ordinary retention. Some vendors' terms say they may keep what you send, and what comes back, for up to 30 days to provide the service and to look for abuse, and then remove it unless the law requires them to keep it. Some also offer what they call zero data retention, but only on request and only for eligible services, which means it is something you ask for, not something you get.

The second is deletion. When you delete something, is the record purged, or does it simply disappear from your view while copies sit in logs and backups? Ask for the retention period as a number. "As long as necessary" is not a number.

Same vendor, opposite defaults, and your team may be on either one.

3. Who else sees it, and where?

Most AI tools run on other companies' infrastructure. Those other companies are called subprocessors: the vendor's own service providers, such as cloud hosting, that touch your data along the way. Some vendors' data processing terms list them in a schedule, promise reasonable notice before a new one gets access to your personal data, and give you a right to object. That is what you want: a list, notice before it changes, and a statement of where processing happens. If your customer contracts promise that data stays in a particular country, this clause either keeps that promise or quietly breaks it.

People deserve their own question. Some vendors disclose that authorized employees, and outside contractors bound by confidentiality, may review stored conversations, for example to investigate abuse. Find out whether a person at the vendor can ever read what your team types, and why. Your own customers may ask you.

4. Who owns what comes out?

Some vendors' business terms say the customer owns the output and the vendor transfers to the customer whatever rights it has in it. Check that yours does, and whether the vendor keeps any license to use your outputs, and for what.

Two further points are not contract questions at all. First, some vendors' terms say plainly that output may not be unique, and that other users may receive similar content. Second, copyright. The U.S. Copyright Office's registration guidance says that when an AI tool determines the expressive elements of what it produces, that material is not the product of human authorship, is not protected by copyright, and must be disclaimed when you apply to register the work. A person's selection, coordination and arrangement of material may still be protected, even though the AI-generated pieces standing alone are not. Owning the output as against your vendor is not the same as holding exclusive rights in it against the world.

5. Who pays if the output infringes?

Some vendors now promise to defend their customers against claims by third parties that the service or its output infringes their intellectual property. Read the conditions before you rely on the promise. Some make it depend on your not having disabled or interfered with the product's content filters or safety systems, on your not using the output in a way you knew or should have known would infringe, and on your having the rights to what you put in. Some exclude claims arising from output you modified.

Then ask two more questions. Does the promise sit inside the vendor's general limit on liability, which in some terms is the fees you paid in the previous twelve months, or outside it? Some vendors' terms carve the intellectual property promise out of that limit entirely. Where it sits inside a limit tied to a year of a modest subscription, it is not much of a backstop. And what does it cover: only claims about the output your team generated, or also claims about the material the model was trained on? Some vendors' terms expressly include the training data. Read yours.

6. How do you leave, and how do the terms change?

Leaving: what can you take out, in what format, and what happens to what remains? Some vendors' terms commit to deleting your content within 30 days after the agreement ends, unless the law requires them to keep it or you have agreed otherwise in writing.

Changing: some vendors' terms let them update the terms on notice, and treat your continued use after the update takes effect as acceptance. The training clause you approved in March is not necessarily the one in force in September. Put a date in the calendar to read the terms again.

A worked example

An illustration, a composite rather than a client: a twelve-person consulting company evaluates a drafting assistant from a newer vendor. The terms say the vendor may use customer data "to develop and improve the Services," with an opt-out in the account settings. Data is kept "as long as necessary to provide the Services." Output belongs to the customer. There is an intellectual property promise, but it sits inside a general liability limit equal to fees paid in the prior twelve months. Subprocessors are listed on a web page the vendor may change at its discretion.

The read, in five minutes:

- **Clause 1 fails as written and passes if configured.** Someone has to turn the setting off, and that someone has to be named.
- **Clause 2 is unusable.** "As long as necessary" is not a retention period. Ask for the number.
- **Clause 3 is a live risk** if the company's own customer contracts limit who handles their data, because the list can change without notice.
- **Clause 4 is fine.**
- **Clause 5 is real but small,** capped at a year of a modest subscription.
- **The verdict:** approved for internal drafts, not for anything that identifies a customer, until the opt-out is confirmed in writing and the retention period is stated as a number.

That is a decision a policy can be written around.

"As long as necessary" is not a number.

Reading the data terms, the training-rights clause and the exit provisions before a company signs, and negotiating the ones that matter, is one of the engagements described on the firm's AI counsel page.

Where to start

Pick the AI tool your team uses most. This week, find its terms and its data or privacy page for the version your team is actually on, and answer the six questions above on a single page. Where the terms do not answer a question, write that down too: an unanswered question is itself the answer for now.

This is general information about what to look for in AI vendor terms, not legal advice about a specific agreement. Vendors change their terms, so read the current version before you rely on it.

WRITTEN BY

Anthony Clemenza

MANAGING PARTNER

Anthony is the Managing Partner of Clemenza Law Group and serves as general counsel to a wide range of companies and philanthropic organizations. Nearly every essay here began as a real question from a real client: the contract that landed at 4 p.m., the board that meets on Thursday. He writes them the way he would answer you across the table, plainly and with care.

CLEMENZA LAW GROUP

NEW YORK, NY · (212) 365-4875 · CLEMENZALAW.COM

The current version of this essay is kept at clemenzalaw.com/insights/ai-vendor-data-terms-six-clauses