INDUSTRY INSIGHTS

Thought Leadership

Analysis & perspectives from the TermsWatchdog team

TermsWatchdog.com

Does AI Own Your Company's Data?

What the Fine Print Actually Says

A project manager at a mid-sized architecture firm is up against a deadline. A proposal is due the next morning. He drops the RFP into an AI tool, asks for a draft response, and pastes in a few excerpts from a recent healthcare project to strengthen the narrative. It works. The draft is clean, fast, and better than expected.

Two weeks later, someone asks a simple question in a leadership meeting. Where did that data go?

No one in the room has a clear answer. The tool was "approved" in a casual sense. No one read the terms of service closely. No one checked what rights the vendor has to use or retain that information.

This moment is becoming common across professional services firms.

What does "owning your data" actually mean in AI?

When business leaders ask, "does AI own my data," they are usually asking four separate questions at once. Can the vendor use your data to train their models? How long do they store it and where? Who inside or outside their organization can access it? And can they sublicense or share it with other parties, including subcontractors or affiliates?

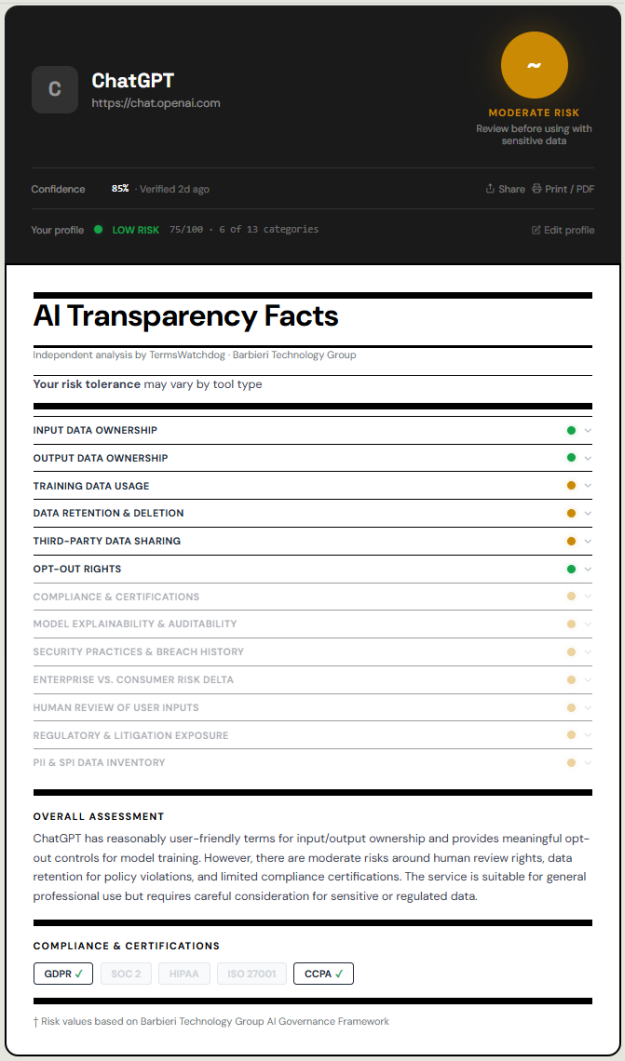
"Ownership" in a strict legal sense often stays with the user. The real issue is usage rights. Most AI terms of service grant the provider some level of license to process, store, and in some cases, learn from your inputs. That is where the risk sits for a business handling client work, proprietary methods, or regulated information.

ChatGPT (OpenAI): Terms of service data and practical risk

OpenAI's public terms and policies distinguish between consumer use and business or enterprise use. In standard consumer usage, OpenAI has historically reserved the right to use conversations to improve models, which includes training. Recent policy updates state that users can opt out of training in certain contexts, and that API and enterprise data is not used for training by default.

In practical terms, if someone on your team is using a free or basic version of ChatGPT, there is a pathway for that data to be retained and potentially used to improve the system. Even when anonymized, that creates exposure for firms working with confidential client information, especially in sectors like healthcare, government, or corporate real estate. There's also the issue around company data on personal accounts and the ramifications of staff attrition.

Enterprise offerings change the posture significantly. OpenAI states that business and API data is not used to train models, and that customers retain control over their data. There are also commitments around data retention windows and access controls. The risk does not disappear, but it becomes something you can manage contractually and operationally.



1. ChatGPT transparency card from TermsWatchdog.com

The gap most firms miss is internal behavior. Even if leadership has approved an enterprise license, staff may still default to free tools out of habit. That is where leakage occurs.

> *The gap most firms miss is internal behavior.*

Microsoft Copilot: Enterprise AI data rights in context

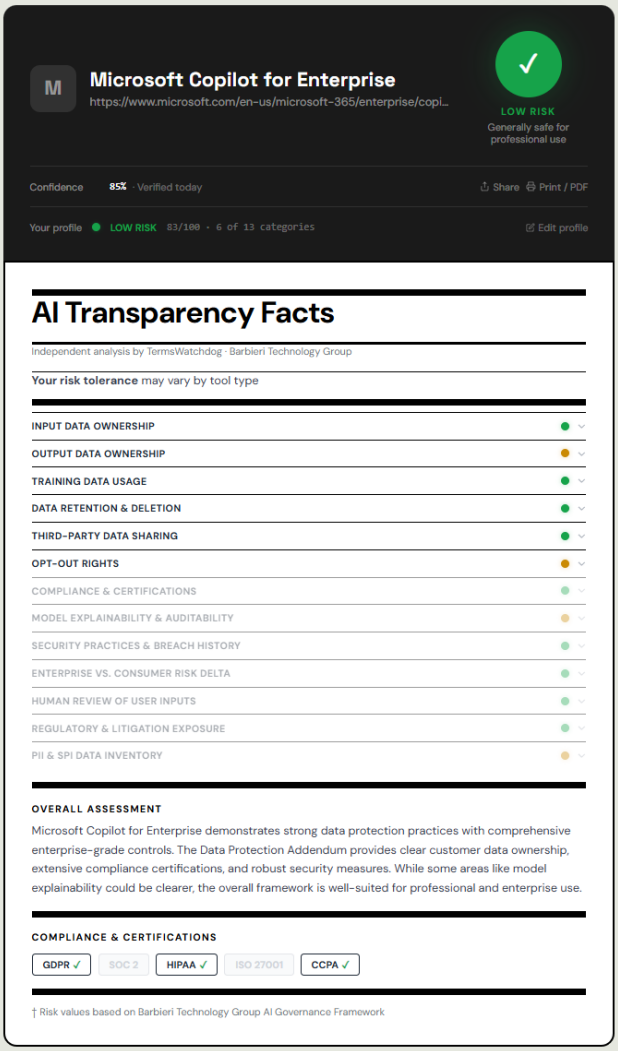
Microsoft positions Copilot as an extension of its existing enterprise stack. The key distinction in its terms is that customer data within Microsoft 365 is governed by existing enterprise agreements, including data protection commitments and tenant isolation.

Microsoft states that prompts and responses in Copilot are not used to train foundation models in a way that would expose one customer's data to another. Data stays within the tenant boundary, and access is governed by existing permissions. That is a meaningful difference compared to standalone AI tools.

From a risk standpoint, Copilot tends to align with enterprise expectations around compliance frameworks such as SOC, ISO, and regional data handling standards. For firms already operating in Microsoft's ecosystem, this reduces friction and perceived risk.

However, the practical issue is overexposure inside the organization. Copilot can surface data from across SharePoint, Teams, and email. If your internal permissions are loose, Copilot can amplify that problem. It does not create new data, it makes existing data easier to access. That can lead to unintended disclosure within the firm.

Enterprise tiers are the default here, since Copilot is generally sold as part of a broader Microsoft agreement. The real control point is governance of your underlying data environment.



2. Microsoft Copilot Ent. transparency card from TermsWatchdog.com

> *The real control point is governance of your underlying data environment.*

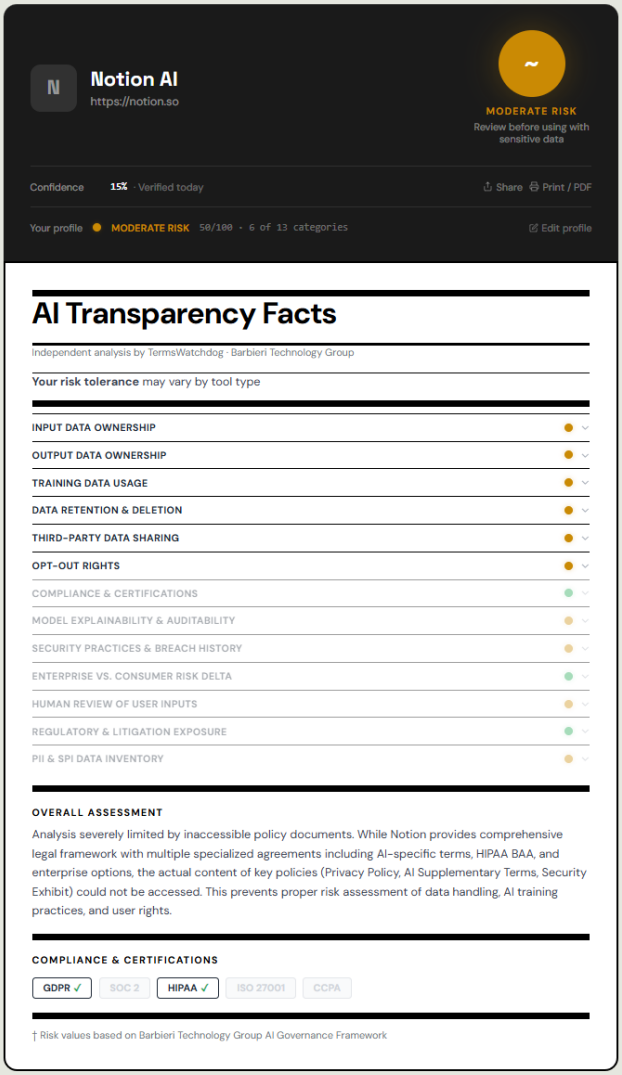
Notion AI: Convenience with broader data usage considerations

Notion's terms indicate that content may be processed through third-party AI providers to deliver its AI features. In earlier versions of its policy, Notion reserved the right to use aggregated and anonymized data to improve its services, which can include AI functionality.

For a business user, this introduces two layers of exposure. First, your data may pass through multiple vendors as part of the AI processing chain. Second, the boundaries of "anonymized" data are not always clear in practice.

The risk is less about outright ownership and more about distribution. A project narrative, meeting notes, or internal strategy document entered into Notion AI could be processed outside your immediate control environment. For firms dealing with sensitive client work, that matters.

Paid and enterprise tiers typically introduce stronger commitments around data handling and security, but the involvement of subprocessors remains a factor. Firms need to understand who those subprocessors are and what jurisdictions they operate in.



3. Notion AI transparency card from TermsWatchdog.com

> *Firms need to understand who those "subprocessors" are and what jurisdictions they operate in.*

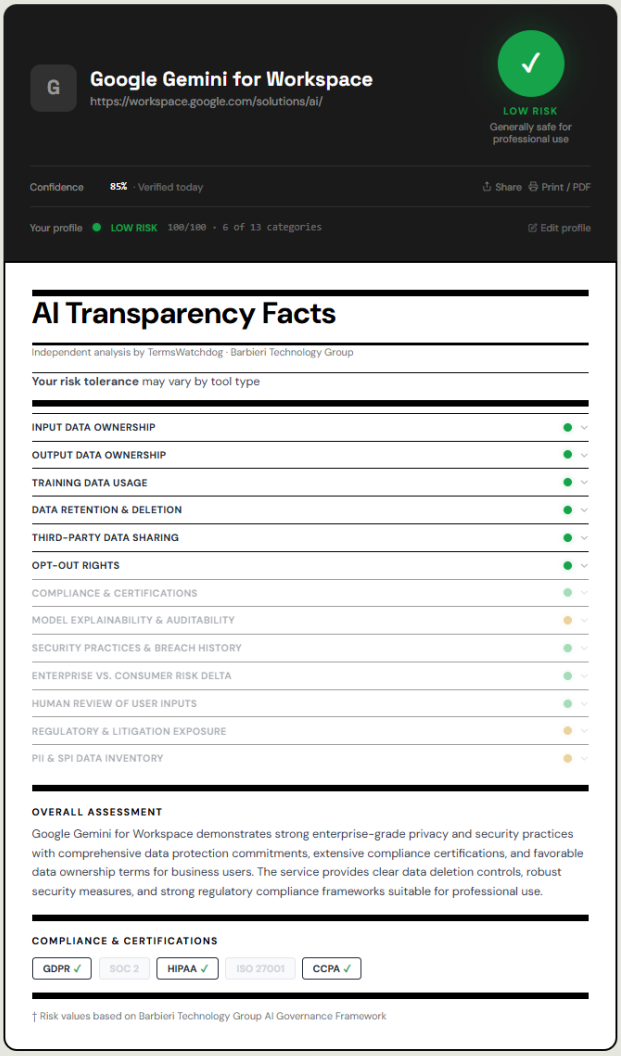
Google Gemini for Workspace: Integrated AI with data handling nuance

Google's Workspace AI tools, including Gemini, operate within the broader Google Cloud and Workspace terms. Google states that customer data is not used to train models outside of the customer's domain without permission, and that enterprise data protections apply.

Like Microsoft, Google leans on its existing enterprise compliance posture. Data remains associated with the customer account, and access is governed by administrative controls. This aligns with expectations for enterprise AI data rights.

The nuance is in how data flows through Google's infrastructure. Processing may occur across regions, depending on configuration, and may involve internal systems that are opaque to most users. For firms with strict data residency requirements, this needs to be examined closely.

Enterprise plans offer more control over data location and retention. As with Copilot, the main operational risk is internal misuse rather than external leakage, assuming the enterprise controls are properly configured.



4. Google Gemini AI transparency card from TermsWatchdog.com

The hidden risks most companies miss

Even firms that take a disciplined approach to procurement tend to miss how AI tools are actually used day to day. A designer pastes a section of a client brief into a prompt to "clean up the language." A project manager asks for a quick summary of a contract. A junior staff member experiments with a new tool they found online. None of these actions feel significant in isolation. Collectively, they create a steady stream of data leaving the firm's controlled environment.

None of these actions feel significant in isolation. Collectively, they create a steady stream of data leaving the firm's controlled environment.

Subprocessors are another blind spot. Many AI tools rely on a chain of vendors to deliver their functionality. Your data may pass through multiple systems, each with its own policies and risks. Most firms do not map this chain before adopting a tool.

Data residency is often overlooked until it becomes a problem. If your firm works on government projects, healthcare facilities, or sensitive infrastructure, where data is stored and processed matters. Many AI tools do not provide clear or enforceable guarantees about geographic boundaries unless you are on a specific enterprise plan.

Finally, there is the issue of auditability. If a client asks how their data was handled in the course of a project, most firms cannot provide a clear answer when AI tools are involved. *That is shifting quickly from a theoretical concern to a contractual requirement*.

What to do about it

The first step is procedural. AI tools should go through the same level of review as any other software procurement. That means reading the terms of service, understanding data usage rights, and documenting the decision. This is not a one-time exercise. These policies change frequently.

The second step is structural. Where possible, use enterprise tiers that provide clear commitments around data usage, retention, and access. Free or consumer-grade tools have a role, but not for handling client or proprietary information.

The third step is behavioral. Firms need a clear internal AI usage policy that defines what can and cannot be entered into these tools. This includes examples that resonate with staff, such as

project narratives, client names, or contract language. Training matters here. ***Policies without context do not change behavior.***

The fourth step is verification. Third-party tools that analyze AI vendor terms can provide an additional layer of insight, especially for firms without in-house legal or compliance teams. This is where platforms like TermsWatchdog.com become relevant. They translate dense legal language into practical risk categories that business leaders can act on.

In my experience leading technology strategy across large AEC organizations, vendor selection and contract negotiation often determine the success or failure of a digital initiative. The same is now true for AI. The difference is that the risk surface extends into everyday actions by staff, not just formal system implementations.

Closing

AI procurement is moving out of the IT department and into the realm of legal, compliance, and executive leadership. The question is no longer whether a tool works. It is whether its terms align with how your firm operates, how your clients expect their data to be handled, and how much risk you are willing to carry.

Firms that treat AI as just another productivity tool will continue to expose themselves in small ways that add up over time. Firms that treat it as a contractual and operational decision will be in a stronger position as clients begin to ask harder questions.

The fine print has always mattered. AI is simply making that more visible.

- *Craig Barbieri, Fractional CTO and AI technology strategist for the AEC, visit at www.Barbieri.biz*