The Shadow AI Problem

How to let your team use AI without losing visibility, control, or trust

For business owners, CIOs, CISOs, and leadership teams

Jason Makevich, CISSP

Executive summary

Generative AI is already inside most businesses, whether leadership formally approved it or not.

Employees use AI to draft emails, summarize documents, rewrite proposals, analyze spreadsheets, create meeting notes, research topics, and move faster through everyday work. The productivity benefit is real. So is the risk.

Shadow AI is the use of AI tools, features, apps, browser extensions, agents, or personal accounts that the business has not approved, governed, or secured. It includes employees using personal ChatGPT or Claude accounts for work, as well as less obvious examples like AI browser extensions, meeting assistants, AI features inside SaaS platforms, and agentic tools that can read, write, summarize, or act across business data.

Sensitive data can leave the business through AI prompts, uploads, screenshots, browser extensions, personal accounts, and unmanaged tools, often with little or no visibility. This can affect customer trust, employee privacy, legal exposure, contract obligations, cybersecurity risk, and your ability to explain what happened if something goes wrong.

Banning AI may feel safer, but it often pushes usage into the shadows. If employees believe AI helps them do their jobs and the business does not give them a safe, easy path to use it, many will find their own way through personal accounts, personal devices, screenshots, browser add-ons, or tools no one has reviewed.

The better approach is to make the safe path the easy path: give employees an approved place for AI at work, set clear rules for what data can and cannot be used with AI, bring users into the conversation so they understand the "why," and apply technical guardrails where work actually happens: in AI tools, SaaS apps, and the browser.

A Secure AI Workspace provides an approved front door to AI with policy controls, sensitive data protection, model governance, and audit-ready evidence. Browser-level controls help manage what happens across SaaS apps, web-based work, AI sites, extensions, downloads, copy/paste, screenshots, and personal account usage.

Together, these controls help your organization say "yes" to AI without accepting unmanaged risk.

The easiest path to AI should also be the safest one.

What Shadow AI is

Shadow AI is any AI use that happens outside the visibility, approval, and governance of the business.

It can include:

- Employees using personal ChatGPT, Claude, Gemini, Perplexity, or similar accounts for work
- Staff uploading documents to public AI tools for summaries or rewrites
- AI browser extensions that can read web pages, prompts, or chat history
- AI meeting assistants that join calls, record conversations, and create summaries
- AI features built into SaaS tools that employees activate without review
- AI coding assistants or “vibe coding” tools used without security review
- AI agents or workflow tools that can access apps, files, calendars, email, or business systems
- Free tools that appear useful but have unclear data use, retention, security, or privacy terms

Shadow AI usually starts with good intentions.

People are not trying to harm the business. They are trying to save time, improve their work, respond faster, or keep up with coworkers who are already using AI.

That is what makes the risk difficult. The behavior feels productive, not risky.

An employee may think:

- “I just need help rewriting this email.”
- “I just need a quick summary of this contract.”
- “I just need ChatGPT to clean up this spreadsheet.”
- “I just need this AI extension to make my browser more useful.”
- “I just need a meeting summary so I can focus.”

But if the data includes customer information, employee details, financials, contracts, internal plans, source code, legal context, account numbers, health information, payment data, or other sensitive content, the business may have lost control of information it was responsible for protecting.

Why Shadow AI is growing so quickly

Shadow AI is growing because AI is useful.

Employees are not adopting AI because they want to create security headaches. They are adopting it because it helps them work faster.

Research shows AI adoption is already widespread. Microsoft reported that 75% of knowledge workers were using AI at work, and 78% of AI users were bringing their own AI tools to work. At small and medium-sized companies, Microsoft found that bring-your-own-AI behavior was even more common. [1]

BCG also found that when employees do not have the AI tools they need, more than half say they will find alternatives and use them anyway. BCG described that pattern as a recipe for frustration, security risks, and fragmented efforts. [2]

This explains why policies alone do not solve the problem.

If the organization says “do not use AI,” but employees believe AI helps them work better, some will still use it.

If the organization says “use AI, but do not put anything sensitive into it,” many users will still make mistakes because they do not always recognize sensitivity in context.

If the organization offers an approved AI tool that is too limited, too slow, too hard to access, or not as useful as what employees already use personally, some will go around it.

Shadow AI grows when people do not have an approved path that feels easy, capable, and safe.

Why Shadow AI creates business risk

Shadow AI is not just “unauthorized software.” It is a new way for business data to leave the organization.

The risk is not limited to one tool or one vendor. It can happen through any unmanaged AI workflow where employees submit, upload, paste, summarize, rewrite, or analyze information.

Data leakage

Sensitive data can leave the organization through normal AI use.

That can include:

- Customer records
- Employee information
- Contracts and pricing
- Sales proposals
- Financial data
- Internal plans
- Product roadmaps
- Source code
- Credentials or security details
- Regulated or protected data
- Confidential client information

One enterprise AI and SaaS security report found that 77% of users paste data into GenAI tools, and that 82% of that activity comes from unmanaged accounts. The same report states that GenAI accounts for 32% of all corporate-to-personal data exfiltration, making it the number one vector for corporate data movement outside sanctioned environments in its telemetry. [3]

Another security report found that 68% of employees use free-tier AI tools through personal accounts, and that 57% input sensitive data. [4]

Not every AI interaction is dangerous. Unmanaged AI simply makes it very easy for sensitive information to move into systems the business does not control.

Loss of visibility and control

When employees use personal AI accounts, the business usually cannot answer basic questions:

- Which AI tools are being used?
- Who is using them?
- What data was pasted or uploaded?
- Was sensitive data involved?
- Was anything retained?
- Can we delete it?
- Was it shared with a third party?
- Did an extension or plugin see the data?
- Did the output create a business decision, customer communication, or legal risk?

This lack of visibility can be just as damaging as the data exposure itself.

If a customer, auditor, insurer, regulator, attorney, or board member asks what happened, the organization may not have evidence. It may only have assumptions.

Provider retention and legal process

Many organizations focus on whether AI providers train on their data. That is important, but it is not the only issue.

Even when training is turned off, prompts and responses may still be retained for operational purposes, abuse monitoring, compliance, legal obligations, or policy enforcement, depending on the provider and plan.

For example, OpenAI's platform documentation says abuse monitoring logs may contain customer content, such as prompts and responses, and are generally retained up to 30 days unless longer retention is required by law or needed to protect services or third parties. [11]

Anthropic's organization data retention documentation states that chats or sessions may be retained as required by law or as necessary to combat usage policy violations. [12]

Turning off training helps. It does not mean sensitive data never leaves your environment, never exists in provider systems, or is never subject to legal or operational processes.

Privacy and compliance exposure

If your organization handles personal information, regulated data, health information, payment data, financial records, or confidential client information, Shadow AI can create serious exposure.

The Office of the Australian Information Commissioner recommends that organizations do not enter personal information, particularly sensitive information, into publicly available generative AI tools because of significant and complex privacy risks. [9]

North Carolina's Department of Information Technology guidance says to never enter personally identifiable or confidential information into publicly available generative AI tools. It also states that entering information into a publicly available generative AI tool is equivalent to releasing it publicly. [10]

Even if your organization is not in a heavily regulated industry, almost every business has sensitive information. That can include customer lists, employee information, contracts, margins, pricing, strategy, vendor agreements, sales pipeline data, and intellectual property.

Cybersecurity risk

Shadow AI also creates cybersecurity risk.

AI tools and AI-adjacent browser extensions can become new collection points for sensitive information. Microsoft Defender reported in March 2026 that malicious AI assistant browser extensions harvested LLM chat histories and browsing data from platforms such as ChatGPT and DeepSeek. Microsoft noted roughly 900,000 installs and activity across more than 20,000 enterprise tenants. [7]

AI risk is not limited to the model itself. Browser extensions, plugins, AI helpers, meeting tools, and workflow agents may request broad permissions, access page content, read prompts and responses, or interact with SaaS applications. If those tools are not reviewed and governed, they can become another path for sensitive data to leave the business.

AI systems also introduce risks such as prompt injection, sensitive information disclosure, supply chain risk, excessive agency, misinformation, and insecure outputs. The OWASP Top 10 for LLM Applications includes prompt injection and sensitive information disclosure as major LLM application concerns. [8]

As AI tools become more agentic, the risk grows. A chatbot that only answers questions is one thing. An AI agent that can read email, access files, interact with SaaS applications, or take actions inside a browser creates a much larger control problem. The more access an AI tool has, the more important it becomes to govern what it can see and do.

Misinformation and decision risk

Shadow AI is not only about data leaving the business. It is also about unreliable outputs entering business decisions.

AI can produce inaccurate, incomplete, outdated, biased, or fabricated information. North Carolina's guidance specifically warns that publicly available generative AI should be evaluated for accuracy and that outputs may include hallucinations, incorrect context, and non-existent sources. [10]

When AI is unmanaged, there may be no standard for review, attribution, approval, or accountability. That can lead to bad customer communications, flawed analysis, incorrect recommendations, and decisions based on unsupported information. The risk is not only that AI may be wrong. It is that people may act on the output because it sounds confident and complete.

Governance should cover both what information goes into AI and how AI-generated output is reviewed before it influences customers, operations, financial decisions, or security actions.

Why policy alone is not enough

AI policies are necessary. They set expectations, define acceptable use, and give leadership a basis for governance.

But a policy by itself does not inspect prompts, block uploads, review browser extensions, detect personal account usage, or prove what did and did not leave the organization.

Policies also depend on users recognizing risk in the moment.

Many employees do not realize that a harmless-looking prompt can contain sensitive data. A prompt may include a client name, account number, private employee issue, contract detail, project codename, confidential business plan, or screenshot with hidden context. The employee may see only the task they are trying to complete, not the full data exposure created by the prompt.

And users are not compliance programs. They are people under time pressure.

If the policy says one thing but the easiest path says another, the easiest path usually wins. This is why governance has to be built into the workflow, not left only to memory, judgment, or annual training.

Strong AI governance combines:

- Clear policy
- User education
- User involvement
- Approved tools
- Technical controls
- Visibility and evidence
- Ongoing review

The policy tells people what the business expects.

The tools make the right behavior easier.

The controls reduce the chance of mistakes.

The logs and reporting help the organization prove what happened.

Together, these elements turn AI policy from a document people may forget into a working system the business can actually rely on.

The easiest path to AI should be the safest one

If the safest path is slow, confusing, limited, or frustrating, users may avoid it. If the unsafe path is fast, familiar, and powerful, users may choose it even when they know they should not.

A good AI program must balance innovation and governance.

Employees should not feel that governance means “no.” They should feel that governance gives them a better way to use AI at work. The approved path should help them move faster while also protecting the business, its customers, and its data.

That means the approved AI experience should be:

- Easy to access
- Familiar to use
- Fast enough for daily work
- Capable enough to replace personal tools
- Flexible enough for different roles
- Governed enough to protect sensitive data
- Visible enough for leadership and security teams
- Clear enough that users understand why it exists

This is where a Secure AI Workspace becomes valuable.

A Secure AI Workspace gives employees an approved place to use AI, often with access to multiple leading models through one interface. It can apply policy before sensitive information reaches model providers, provide logging and reporting, and support a more consistent user experience than a patchwork of personal accounts.

It also gives the organization a better adoption story. Instead of asking users to give up AI, the business can give them a safer and more useful place to use it. That shift is important. People are more likely to follow the rules when the approved option feels like an upgrade, not a restriction.

The best programs make governance feel natural. Users know where to go, leaders know what is being used, and the business can keep improving policies based on real behavior instead of guessing.

Safe AI should feel easier than Shadow AI.

Bring users into the conversation early

One of the best ways to reduce Shadow AI is to invite users into the AI conversation.

Employees closest to the work often see the best AI opportunities first. They know which tasks are repetitive, which workflows create frustration, where customers get delayed, and where AI could improve quality or speed.

If leadership builds AI governance only from the top down, the organization may miss some of the most useful ideas. It may also create rules that feel disconnected from real work.

When employees only hear:

- “Do not use that tool.”
- “Do not paste that data.”
- “That site is blocked.”
- “That app is not allowed.”

they may see governance as an obstacle.

But if they understand the “why,” they are more likely to follow the approved path.

They need to understand that the concern is not anti-AI. The concern is protecting customers, employees, company data, regulated information, client trust, and the business itself while still allowing the organization to benefit from AI.

Good approaches include:

- AI innovation workshops
- Department-level AI champions
- Internal AI use case submissions
- Clear processes for requesting new AI tools
- Approved use case libraries
- Training that explains risks in plain language
- Regular conversations about what is working and what is not
- Safe channels for reporting risky workflows or unclear policy questions

User buy-in is not just a soft benefit. It is part of the risk strategy.

If users are not invited into the AI conversation, some will create their own path.

Innovation should happen inside the organization, not around it.

What good Shadow AI governance looks like

Strong Shadow AI governance does not require solving everything at once.

It starts by creating a workable operating model:

1. Define what Shadow AI means for your business

Do not assume everyone understands the term.

Define it simply:

Shadow AI is any AI tool, account, extension, app, agent, or feature used for business purposes without approval, visibility, or governance.

Then provide examples employees recognize.

2. Define the data that should never go into public AI tools

Examples may include:

- Customer data
- Employee records
- Payment or banking information
- Health or benefits information
- Legal strategy
- Confidential client information
- Non-public financials
- Contracts and pricing
- Source code
- Credentials
- Security findings
- Product strategy
- Anything regulated, protected, or contractually restricted

If you wouldn't post it publicly, don't paste it into public AI models.

3. Provide an approved AI path

Do not simply tell users what not to do.

Give them a better option.

A Secure AI Workspace can provide one approved place to use AI, with access to leading models and guardrails that reduce the risk of sensitive data exposure.

4. Put controls where work happens

Modern work happens in browsers, SaaS apps, collaboration tools, and AI interfaces.

That means control needs to happen at the point of action:

- Prompt entry
- File upload
- Copy and paste
- Download
- Print
- Screenshot
- App access
- Personal account usage
- Browser extension activity
- SaaS sharing
- AI output review

This is why browser-level controls are becoming more important. Standard consumer browsers were not designed as business data governance platforms. They were built for general web use, ecommerce, and consumer browsing. But today, much of the workday happens inside the browser.

A secure browser built for work can help apply policy and visibility where users interact with SaaS apps, AI sites, extensions, and sensitive data. Some organizations evaluate purpose-built secure browsers because they bring business controls into the place where cloud work actually happens.

5. Measure, report, and improve

AI usage will change quickly. New tools will appear. Existing tools will add AI features. Employees will find new workflows. Business priorities will shift.

Good governance is not a one-time rollout. It requires review and adjustment.

Your organization should be able to answer:

- Which AI tools are being used?
- Which use cases are approved?
- Where are users trying to send sensitive data?
- Which policies are creating too much friction?
- Which teams need more training?
- Which controls need adjustment?
- What can leadership confidently say to customers, auditors, insurers, or regulators?

The two control planes that help most

Most organizations can greatly benefit from two complementary control planes for Shadow AI.

Control plane 1: A Secure AI Workspace

A Secure AI Workspace provides the approved place to use AI.

It helps with:

- Giving users a safe and easy path to AI
- Reducing reliance on personal accounts
- Supporting multiple models in one governed experience
- Applying policy to prompts and uploads
- Detecting or protecting sensitive data before it reaches model providers
- Providing logs and reports showing how AI is used
- Supporting reusable workflows, prompts, and approved AI use cases

This is the positive control. It answers the question:

“Where should employees go when they want to use AI for work?”

Control plane 2: Browser-level security and visibility

A secure work browser or browser-level governance helps address what happens across SaaS, web apps, extensions, personal accounts, and unmanaged AI activity.

It helps with:

- Reducing risky copy/paste and uploads
- Managing browser extensions and AI add-ons
- Applying policies to SaaS and web workflows
- Controlling access from unmanaged devices or BYOD
- Separating business identities from personal identities
- Providing visibility into work that happens in the browser
- Reducing the chance that employees bypass approved AI paths

This is the guardrail control. It answers the question:

“What happens if users try to work around the approved path?”

Neither control is perfect by itself.

A Secure AI Workspace gives users the approved AI experience.

Browser-level controls help reduce unmanaged activity outside that approved experience.

Together, they make it easier to enable AI while reducing the risk of Shadow AI.

Why your IT or security provider should be involved

Most businesses should not try to solve Shadow AI alone.

This is not just a tool-selection decision. It touches security, identity, browser activity, SaaS usage, data protection, training, policy, compliance, contracts, and day-to-day workflows.

Your IT or security provider can help you:

- Discover how AI is already being used
- Identify sensitive data exposure paths
- Review browser extensions and AI add-ons
- Define practical AI usage rules
- Deploy an approved AI workspace
- Apply browser-level controls where needed
- Configure policies without disrupting work
- Create reporting for leadership
- Train users in plain language
- Review and improve the program over time

The best provider-led programs start with helping the organization use AI safely, productively, and visibly.

A good first step

You do not need to solve every AI risk immediately.

A good first step is to ask your IT or security provider for a Shadow AI review.

That review should answer:

- What AI tools are already being used in the business?
- Are employees using personal AI accounts for work?
- Are AI browser extensions or add-ons installed?
- What sensitive data could be exposed through prompts, uploads, screenshots, or SaaS workflows?
- What AI use cases should be approved?
- What data should never be entered into public AI tools?
- What approved AI workspace would users actually choose?
- What browser-level controls would reduce workarounds without creating unnecessary friction?
- What reporting would leadership need to feel confident?

From there, build a phased plan.

A strong starting plan usually includes:

1. **Establish simple AI rules**
   Make clear what is approved, what is not, and why.
2. **Give users a safe AI option**
   Deploy a Secure AI Workspace that is easy enough and useful enough to replace personal tools.
3. **Bring users into the process**
   Collect use cases, train teams, and explain the risks in plain English.
4. **Add browser-level guardrails**
   Control risky interactions where SaaS, web apps, extensions, and AI usage meet.
5. **Review and tune regularly**
   Use evidence to improve policies, reduce friction, and expand AI safely.

Conclusion

Shadow AI is already inside many businesses.

Employees are using AI because it helps them move faster. That is not the problem.

Unmanaged AI use can move sensitive data into places the business cannot see, control, govern, or explain.

Blocking AI entirely is not enough. Policies alone are not enough. Training alone is not enough.

The right approach is to make safe AI the easiest path.

That means giving users approved AI tools they actually want to use, bringing them into the conversation, applying guardrails where work happens, and working with your IT or security provider to manage AI governance over time.

The business goal is not to slow people down.

The business goal is to let people innovate safely.

Learn more about Secure AI Workspaces at: https://port1.io/SecureAI

Learn more about browser-level control for modern work at: https://port1.io/island

References

1. Microsoft Work Trend Index, “AI at Work Is Here. Now Comes the Hard Part.” Microsoft reported that 75% of knowledge workers use AI at work and 78% of AI users bring their own AI tools to work. https://www.microsoft.com/en-us/worklab/work-trend-index/ai-at-work-is-here-now-comes-the-hard-part
2. BCG, “AI at Work: Momentum Builds, but Gaps Remain.” BCG reported that more than half of employees would find alternatives if they do not have the AI tools they need. https://www.bcg.com/publications/2025/ai-at-work-momentum-builds-but-gaps-remain
3. LayerX, “Enterprise AI and SaaS Data Security Report 2025.” LayerX reported high levels of GenAI copy/paste activity, unmanaged account usage, and GenAI as a major corporate-to-personal data movement vector. https://go.layerxsecurity.com/hubfs/LayerX_Enterprise_AI_and_SaaS_Data_Security_Report.pdf
4. Menlo Security, “2025 Report Uncovers 68% Surge in Shadow Generative AI Usage.” Menlo reported high personal/free-tier GenAI usage and sensitive data entry into AI tools. https://www.menlosecurity.com/press-releases/menlo-securitys-2025-report-uncovers-68-surge-in-shadow-generative-ai-usage-in-the-modern-enterprise
5. Netskope, “Cloud and Threat Report: 2026.” Netskope highlights growing data policy violations involving unauthorized cloud services, personal applications, and GenAI platforms. https://www.netskope.com/resources/cloud-and-threat-reports/cloud-and-threat-report-2026
6. IBM, “Cost of a Data Breach Report 2025.” IBM highlights an AI oversight gap, including organizations lacking proper AI access controls and AI governance policies. https://www.ibm.com/reports/data-breach
7. Microsoft Security Blog, “Malicious AI Assistant Extensions Harvest LLM Chat Histories.” Microsoft reported malicious AI assistant browser extensions harvesting LLM chat histories and browsing data. https://www.microsoft.com/en-us/security/blog/2026/03/05/malicious-ai-assistant-extensions-harvest-llm-chat-histories/
8. OWASP, “Top 10 for LLM Applications 2025.” OWASP identifies prompt injection, sensitive information disclosure, and other LLM application risks. https://owasp.org/www-project-top-10-for-large-language-model-applications/assets/PDF/OWASP-Top-10-for-LLMs-v2025.pdf
9. OAIC, “Guidance on privacy and the use of commercially available AI products.” The OAIC recommends organizations do not enter personal information, particularly sensitive information, into publicly available generative AI tools. https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/guidance-on-privacy-and-the-use-of-commercially-available-ai-products
10. North Carolina Department of Information Technology, “Use of Publicly Available Generative AI.” NCDIT says not to enter PII or confidential information into publicly available GenAI tools and states that entering information into such tools is equivalent to releasing it publicly. https://it.nc.gov/resources/artificial-intelligence/use-publicly-available-generative-ai
11. OpenAI Platform Documentation, “Data controls in the OpenAI platform.” OpenAI notes abuse monitoring logs may contain prompts and responses and are generally retained up to 30 days unless longer retention is legally required or otherwise needed. https://developers.openai.com/api/docs/guides/your-data
12. Anthropic Privacy Center, “How long do you store my organization’s data?” Anthropic notes chats or sessions may be retained as required by law or to combat usage policy violations. https://privacy.claude.com/en/articles/7996866-how-long-do-you-store-my-organization-s-data