

Safely Enabling Workforce AI

A primer for security leaders on governing how employees and their agents use AI, without slowing down the work.

KEY FINDINGS

64.5%

of activity on personal AI accounts is business work

21.9%

of files uploaded to GenAI tools contained sensitive data

45.6%

of personal AI use happens on enterprise-licensed accounts

Overview

AI at work began as a chat window in the browser. In three years it has become a set of applications on the web and desktop that employees now use as their main interface to work.

These applications can act on their own, working with files on the device and inside the SaaS tools employees used to operate directly.

Organizations that respond by blocking these tools rarely reduce usage. Instead, the usage moves into personal accounts the security team can't see.

This primer describes where employee AI usage creates risk and sets out a three-phase approach to enabling it safely.

Inside This Primer

- 01 How AI became the interface to work
- 02 Why blocking pushes usage out of view
- 03 Where employee AI usage leads to risk
- 04 A phased approach to safe enablement
- 05 Evaluating a solution



Establish Visibility

Know which AI tools are in use, by whom, and what they can reach



Govern in Context

Protect sensitive data and limit agent actions without stopping the work



Reassess and Invest

Put budget behind the AI tools employees rely on most

How AI Became the Interface to Work

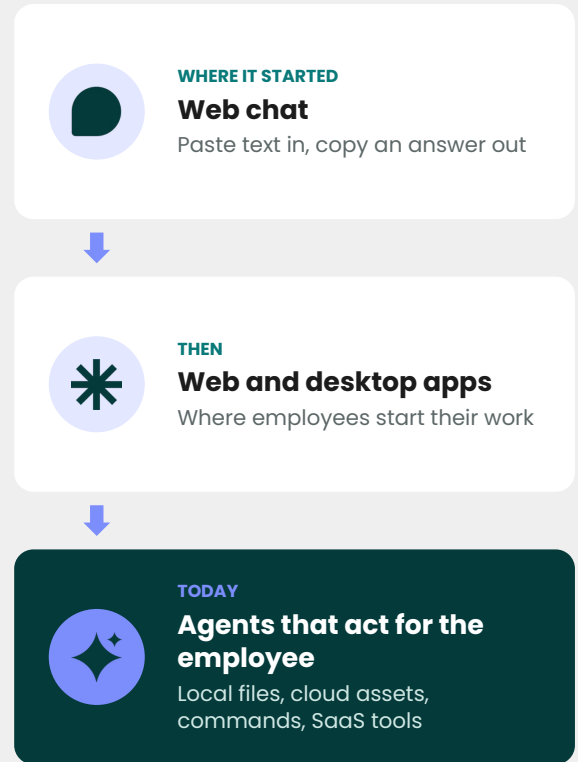
Generative AI entered the workplace as a search engine alternative in the browser. It has since become the surface (whether on web or desktop) where most work starts.

When generative AI first entered the workplace, employees pasted in text and copied out an answer, and the main concern was what they typed.

That is no longer how most people work with AI. AI applications now run on both the web and the desktop, and employees increasingly start their work there rather than in the business applications they used before.

Through plugins and MCP servers, these tools act for the employee. A growing share of these agents run in the cloud, outside the endpoint entirely.

AI is also now built into many of the SaaS applications employees already use, and existing controls often don't classify those applications as AI tools.



WHAT TODAY'S AGENTS CAN REACH



Local files

Read and edit documents, code and data on the device



Cloud assets

Work with files, data and resources in cloud storage and hosted environments



Terminal commands

Run scripts and commands on the device for the employee



SaaS tools

Create, edit, or delete records across tools like Jira, M365, Slack, and Notion.

23

new AI applications found in use at the average organization in a single quarter.

Many are everyday SaaS tools with AI built in, and each one needs reviewing. Source: Harmonic GenAI Data Exposure Report, Q2 2025.

Why Blocking Pushes Usage Out of View

If the most capable AI tools are blocked, employees will find ways to use them anyways, often through free or personal accounts. And those who don't find a workaround will fall behind.

64.5%

of activity on personal AI accounts accessed by employees is business work

26.3%

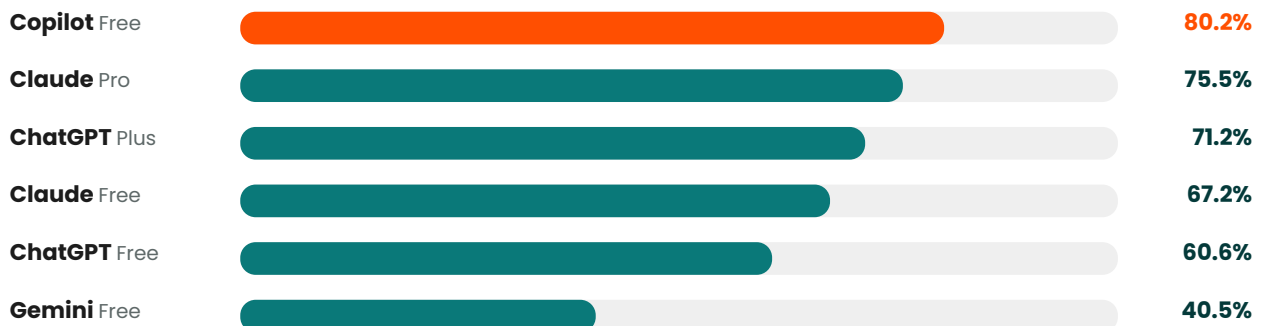
of sensitive prompts and files were submitted through free ChatGPT accounts

45.6%

of personal AI use happens on enterprise-licensed accounts, wasting tokens meant for work

Personal accounts carry business work

Share of activity on each personal or free-tier plan classified as business use



Harmonic Security AI Usage Index 2026. Ambiguous minutes excluded.

Free and personal tiers commonly train on user inputs by default. Any business context stored in a personal account also leaves with the employee when they move on.

Plan tier is an unreliable indicator of risk as well. Enterprise accounts carry personal activity, and personal accounts carry business work, so risk has to be assessed from the content of each interaction.

WHAT GETS EXPOSED Top categories among 579,113 sensitive exposures in 22.4M enterprise AI prompts, 2025

26.5%

Source code

22.3%

Legal documents

16.6%

Financial projections and pipeline data

12.6%

M&A material

Coverage has to follow the employee across every tool they use.

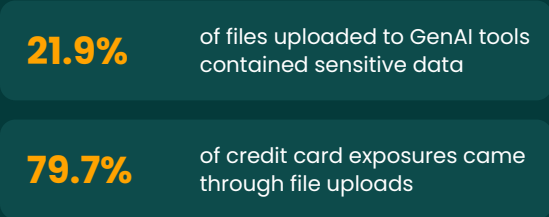
Where Employee AI Usage Leads to Risk

Harmonic research shows that most incidents fall into two categories.



Sensitive data exposure

Sensitive information, like IP, PII, legal discourse, and financial data leave as unstructured text and files that traditional data controls can't interpret.



Unauthorized or destructive actions

Agents working with employee credentials can take actions that can't be reversed, including:

- Deleting a database table
- Overwriting a code branch
- Removing an environment
- Pushing an unapproved config change

THREE FACTORS CONTRIBUTE TO BOTH



Employee error

Employees are focused on getting work done. AI increases the consequences of a routine mistake.



Unpredictable models

Models can take actions no one requested. More capable models are less likely to make mistakes but have a larger blast radius.



Prompt injection

Agents that read web pages, tickets or documents can be manipulated by instructions hidden in that content.

EXISTING TOOLS EACH SEE PART OF THE PICTURE

Network controls

SEES

Where web traffic is flowing to

MISSES

What's within the traffic or occurring in the app

Endpoint tools

SEES

Malicious software on the device

MISSES

Legitimate AI used in risky ways, and cloud agents

AI vendor consoles

SEES

Corporate accounts in their own product

MISSES

Personal accounts and every other tool

Legacy tools remain valuable, but none were designed to govern how employees and agents use AI.

A Phased Approach to Safe Enablement

Start with visibility, add controls that understand context, then use what you learn to decide where to invest.



PHASE 01

Establish Visibility

Know what's in use and what it can reach.

- Cover browser, desktop, CLI and agent connectors, including personal accounts
- Map activity to people and teams through your identity provider
- Inventory MCP servers and connectors, and what each can read, change or delete
- Keep an AI register: approved, monitored or unsanctioned
- Run data policies in monitor mode before you enforce them



PHASE 02

Apply Context-Aware Governance

Protect data and limit agent actions without stopping work.

- Tell employees what's changing before you change it
- Warn first and explain why; save blocking for the highest risk
- Tailor policies by data type, app, department and role
- Let agents read freely, gate changes, and block destructive actions
- Tighten native settings: sandboxes, tenant limits, allowlists



PHASE 03

Continuously Reassess and Invest

Put budget behind the tools that deliver.

- Find the workflows that deliver measurable value
- Spot personal-account demand and idle enterprise seats
- Move that demand onto enterprise accounts that don't train on your data
- Share findings with the AI committee, IT and business leaders
- Review new features before a broad rollout

What Phase 03 finds feeds back into Phase 01.

Evaluating a Solution

Look for two things: visibility into how AI is actually used, and guardrails that understand context.

VISIBILITY

Coverage



Sees all AI usage, sanctioned or not, across browser, desktop, command line, embedded AI and agents, including personal accounts.

Depth



Shows how each tool is used at the prompt and task level: who, for what purpose, and where risky behavior occurs.

GUARDRAILS

Prompt layer



Reads the full prompt or file, understands context and intent, and warns or blocks in real time without adding delay.

Action layer



Governs what agents can do with connected systems, blocks destructive actions and data transfers, and redirects to a safe path.

QUESTIONS TO ASK VENDORS

1

How do they categorize unstructured information in context? Are they able to provide custom, task-level usage insights?

2

Can they provide inline guardrails over the information agents share in tool calls and the actions they take in the web and on the desktop?

3

What apps do they actually cover at the prompt level? And can they distinguish between personal and enterprise accounts?

With this approach, security teams can support broad AI adoption while keeping sensitive data and critical systems protected.



About the Research

Figures come from Harmonic Security research: the AI Usage Index 2026 (1.9 million classified AI-session minutes, May 2026) and the GenAI Data Exposure Report (1 million prompts and 20,000 files across 300+ AI applications, Q2 2025).