AEGYS

FIELD NOTE

Shadow AI is already in your org

How to find the AI tools your team is quietly using, and turn that into an advantage.

A field note from AEGYS
aegys.io · secure@aegys.io

Your team is already using AI. Not "evaluating vendors" or "thinking about a policy," using it, today, on real work. Someone pasted a stack trace into a chatbot to debug it this morning. Someone cleaned up a customer email with one before lunch. That is not a failure of policy. It is a signal that people want the help, and they will find it with or without you.

A ban just turns the lights off

The reflex is to block it. We understand the instinct, but a blanket ban does not remove the risk; it moves it somewhere you cannot see. Usage migrates to personal laptops, personal accounts, and personal phones, where none of your controls reach. You trade a visible, shapeable problem for an invisible one. The goal is not to stop your most capable people from moving fast. It is to give them a fast lane that happens to be the safe one.

What you are actually worried about

Strip away the noise and the real risk is narrow: sensitive data leaving your control. The two failure modes that matter most for a mid-market team are well documented. The first is **sensitive information disclosure**, where confidential data pasted into a tool ends up somewhere you never intended, including, in some configurations, future model training (OWASP lists this as LLM02). The second is the model giving up things it should not, from secrets to its own hidden instructions (OWASP LLM07, system prompt leakage; MITRE ATLAS catalogs the same behavior as LLM data leakage). Notice what is not on that list: science fiction. The common, boring failure is a person pasting the wrong thing into the wrong box.

Find it, then shape it

You cannot protect what you cannot see, so start by mapping reality, not by drafting rules. NIST's AI Risk Management Framework calls this first step **Map**: establish the context before you try to manage it. In practice that is one honest conversation, not a quarter of policy work. Ask the team four questions, with no blame attached:

- Which AI tools have you tried for work?
- What did you use them for?
- What kind of information did you put in?
- What would you automate if you were allowed to?

You will learn more about your real exposure from those answers than from a month of writing standards, and you will surface the use cases worth sanctioning.

Turn it into an advantage

Shadow AI is unmanaged demand. Managed, it is a roadmap. Once you can see what people reach for, three quick moves convert the risk into an asset:

- Bless one or two tools, configured properly (training off, company sign-on, retention understood), and make them the obvious default so nobody reaches for a random site.
- Publish a one-page rule for what data is fine, what needs a check, and what never goes in. People follow rules they can remember.
- Ship one real win on a sanctioned tool, so the safe path is visibly the faster one too.

That is the whole move. You convert a quiet risk into a short list of approved, monitored, genuinely useful tools, and your team keeps the speed it already found.

> **The one idea**
> You do not reduce shadow AI by banning it. You reduce it by making the safe path the easy path.

This is Principle 1 of our Practical AI Playbook, a free guide for security and IT teams. If you would rather have a partner map your exposure and stand up the guardrails with you, that is what a Security and AI Posture Assessment is for. Reach us at secure@aegys.io.

References

MITRE ATLAS, Adversarial Threat Landscape for Artificial-Intelligence Systems. https://atlas.mitre.org

OWASP Top 10 for LLM Applications (2025): LLM02 Sensitive Information Disclosure; LLM07 System Prompt Leakage. https://genai.owasp.org/llm-top-10/

NIST Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1. https://doi.org/10.6028/NIST.AI.100-1