PUBLICATION

Cybersecurity Awareness Month 2025: Don't Get Haunted by Shadow AI

Authors: Matthew George White, Alexander Frank Koskey, III
October 21, 2025

***"It's the boogeyman… and he's been here all along."* – Laurie Strode, *Halloween* (1978)**

As Cybersecurity Awareness Month unfolds and Halloween looms, one of the most chilling threats haunting today's businesses isn't a hacker in a hoodie – it's shadow artificial intelligence (AI). And it may be lurking in your company right now. What may seem like a harmless shortcut or even a helpful digital assistant can quietly morph into a ghost in your machine, undermining defenses, leaking data, and luring your organization into costly incidents, regulatory nightmares, and reputational damage.

This month, as you check your locks and test your alarms, it's worth doing the same for your AI ecosystem. Let this alert serve as your survival playbook to help you detect, contain, and banish the risks of shadow AI – before it becomes the next potential headline that keeps you up at night.

What Lurks in the Shadows – What Is Shadow AI?

Shadow AI refers to the use of AI tools, models, or services by employees (or third parties, including vendors or service providers) without formal approval, oversight, or integration into the enterprise's governance, security, and data controls. It mirrors the concept of "shadow IT" – users adopting unsanctioned apps or infrastructure – but with a more potent payload, because AI tools can ingest or generate sensitive content, interact with application programming interfaces (APIs), and propagate errors or vulnerabilities at scale.

In practice, shadow AI can take many forms:

- A marketing analyst using a public large language model (LLM) or generative AI plugin to generate campaign text that includes (intentionally or not) customer data
- A finance team member using an AI-based forecasting tool through a personal account, feeding it internal budget spreadsheets
- Engineers using unvetted open-source models or inference servers outside the approved stack
- Individuals using generative image tools with internal design assets or proprietary visuals
- Use of browser extensions or chatbot plugins that connect to internal systems via APIs without security vetting

Because these tools are often externally hosted or integrated via APIs, they may bypass the organization's identity, logging, data loss prevention (DLP), or encryption controls.

Frightening Consequences: Why It's So Dangerous

Shadow AI flourishes in the dark corners of convenience and curiosity. Employees usually aren't trying to break rules – they're trying to get work done faster. But every "just this once" prompt to an unapproved tool chips away at your security perimeter.

These risks can escalate rapidly and include:

1. **Loss of Visibility and Control:** Unapproved AI interactions can escape your company's audit trails. When AI tools are not under corporate governance, you may not know who is using them, on what data, or how models are configured.

2. **Data Leakage and Exposure:** Sensitive or regulated information may inadvertently be exposed to third-party AI services – e.g., proprietary customer lists, personally identifiable information (PII), internal strategy, or source code.

3. **Model Poisoning, Prompt Injection, or Manipulation:** Without controls, an attacker could poison inputs or inject malicious prompts to cause harmful outputs or exfiltration.

4. **Dependence on Unvetted or Unsecure Models:** Employees may choose convenience over security, using models without encryption, proper isolation, or formal updates and patching.

5. **Regulatory and Compliance Risk:** If shadow AI causes data breaches or exposures affecting protected data (e.g., Health Insurance Portability and Accountability Act, Gramm-Leach Bliley Act, and General Data Protection Regulation), the organization may face regulatory enforcement, fines, or class action litigation.

6. **Escalating Breach Costs:** According to IBM's 2025 Cost of a Data Breach Report, organizations with high levels of shadow AI suffered, on average, an extra $670,000 in breach costs compared to those with little or none. Moreover, breaches involving shadow AI had longer detection and containment cycles and a higher incidence of customer PII compromise. To put it another way: you don't just lose control – you pay for it.

The final point is particularly significant as IBM found that 63 percent of organizations reported they lack AI governance policies, and even of those that had policies, only one in three perform regular audits for unsanctioned AI. In fact, IBM reported that 20 percent of surveyed organizations experienced a breach linked to shadow AI – making it a non-trivial vector. In affected cases, 97 percent of those organizations admitted they lacked proper AI access controls.

Tales from the (Data) Crypt: Real-World Scares

To bring the risk from abstract to concrete, here are a few representative (though anonymized) examples and plausible scenarios. These scary stories highlight how harmful shadow AI can be:

- A financial services firm's credit modeling team uses a public AI inference API to augment modeling. One call includes customer social security numbers in the prompt, inadvertently exposing them to the third-party AI provider.

- A marketing director uploads a spreadsheet of customer email addresses and purchase history into a generative AI tool to craft "personalized" emails. The tool's vendor stores and indexes those prompts, unintentionally exposing customer data in its backend.

- A software development team experiments with an open-source LLM on a personal GitHub instance. A misconfiguration leaks internal API tokens and internal code (via logs) to the public.

- A product design team uses an AI image generator, feeding it internal product blueprints. A competitor later finds the images and reverse-engineers aspects of the design.

- An HR administrator uses an AI chatbot to summarize employee records (performance, compensation) without realizing the data is retained by the chatbot's vendor. Later, a vendor's database is compromised, exposing HR data.

Because these use cases might not trigger red flags in traditional security, they may slip under the radar – until something goes wrong.

How to Ward Off the Shadows: Defenses and Safeguards

You don't need an exorcism. Stopping Shadow AI doesn't require halting innovation – it requires smart governance, controls, and cultural alignment. Below is a recommended roadmap.

1. **Establish an AI Governance Framework**

- Define roles and responsibilities: Establish an AI governance committee (with legal, security, privacy, and business representation).
- Create clear policies: Articulate what AI tools are approved, what uses are allowed, what data may be fed into the tools, who may access the tools, and what training is required before use.
- Mandate AI risk assessments: Require privacy, security, and ethical reviews before deploying new AI tools or models – even small ones.
- Audit and review: Schedule periodic reviews to detect unsanctioned models or integrations.

2. **Inventory and Discover Shadow AI Activity**

- Network traffic monitoring: Look for anomalous outbound calls to AI inference services (e.g., APIs to model hosts), and inspect API usage patterns, including unknown models and new endpoints.
- Endpoint agents and Endpoint Detection and Response (EDR): Detect local AI tool usage, browser extensions, or unusual library usage.
- User surveys and training: Ask teams what AI tools they use and why; often you'll uncover hidden usage by engaging users directly.

3. **Apply Data Access, Classification, and Control**

- Data classification: Tag data by sensitivity and restrict which classes (e.g., PII, source code, internal strategies) may even be used with AI tools.
- Data sanitization/anonymization: Require that only pseudonymized or masked data is used in AI prompts where possible.
- Tokenization or encryption: Use tokenized or encrypted datasets so that even if data is exposed, it is not in its raw, sensitive form.
- Least privilege/role-based access: Control which systems or individuals can invoke AI tools.
- Prompt-level controls: Limit what prompts may do (for instance, disallow file upload features or limit token lengths).

4. **Secure AI Deployments and Model Infrastructure**

- Access controls and multi-factor authentication (MFA): Enforce MFA and zero-trust access for AI systems.
- Segmentation and isolation: Ensure that AI systems, inference engines, and model training environments are isolated from core data stores.
- Monitoring, logging, and provenance: Maintain full audit logs of inputs, outputs, model versions, and user identity, and maintain lineage tracking of data flowing through models.

- Adversarial testing/red teaming: Simulate prompt injection, model poisoning, or malicious inputs to test resilience.
- Vulnerability management: Patch and secure frameworks, libraries, and dependencies used by AI systems.

5. **Integrate Shadow AI Management into Incident Response**

- Expand your incident response (IR) playbooks to reflect AI-related events (e.g., compromised models, exposed training data, and ingestion of PII).
- Design breach scenarios around AI tools (e.g., what if an attacker hijacks an inference endpoint?).
- Include forensic readiness: Capture prompt logs, model snapshots, API logs, and memory dumps.
- Conduct tabletop drills specifically for AI/shadow AI breach scenarios.

6. **Build a Strong Culture and Training**

- Educate employees – particularly data scientists, analysts, marketing, and business users – on the risks of unsanctioned AI use.
- Incentivize compliance, not punishment – reward teams for flagging or refusing to use unapproved AI.
- Integrate shadow AI risks into regular cybersecurity awareness programs.

7. **Leverage AI in Defense**

- Use AI-powered threat detection, anomaly detection, and log analysis to help uncover misuses or suspicious usage patterns.
- Automate policy enforcement (e.g., scripts that flag prompt patterns or block disallowed API calls).
- Use watermarking or fingerprinting techniques to track AI-generated content or to detect unauthorized model use.

8. **Don't Forget Vendor Management**

- Assess third-party security and privacy controls, including data retention, model training, and subcontracting practices.
- Include AI-specific contract clauses addressing confidentiality, data use (including training data), and intellectual property ownership of model outputs.
- Review data residency and regulatory alignment, especially for vendors using cross-border model hosting.
- Establish offboarding procedures to ensure data is deleted or returned if the relationship ends.

As you can see from this roadmap, addressing shadow AI isn't a task for one department or one leader – it requires a coordinated effort across the enterprise. Executives must set strategy and governance, IT must implement technical safeguards, and every employee plays a role in using AI responsibly. By working together, organizations can harness the power of AI without letting its darker side take hold – ensuring that innovation remains a treat, not a trick.

No More Nightmares: Your Shadow AI Survival Guide

So where do you start? If shadow AI has been creeping around your organization, it's time to turn on the lights. The good news? You don't need a silver bullet – just a solid plan. Below is a quick checklist to keep your AI ecosystem safe, secure, and firmly under your control:

7. Create and enforce an AI governance framework and policy.

8. Discover and inventory existing unsanctioned AI usages.
9. Classify and restrict sensitive data usage in AI workflows.
10. Secure AI systems with zero trust, isolation, authentication, and logging.
11. Update your IR plans and conduct AI-specific tabletop drills.
12. Educate and align your workforce on safe AI use.
13. Leverage AI capabilities in your defensive arsenal.

While embracing AI is essential for staying competitive, neglecting oversight exposes your organization to incremental but material risk – risk that, according to IBM's 2025 Cost of a Data Breach, already adds hundreds of thousands in costs and delays breach response.

If your organization is grappling with how to use AI, govern AI, detect shadow AI, or bake security into your AI innovation pipeline, we are happy to help you navigate the legal, compliance, and cybersecurity challenges that arise. For more information or if you have any questions, please contact Matt White, Alex Koskey, or any member of Baker Donelson's Data Protection, Privacy, and Cybersecurity Team or its Artificial Intelligence Team.