BONFY SOLUTION BRIEF:

Managing Shadow AI Risk

bonfy.ai

Shadow AI creates new, hard-to-govern paths for sensitive data exposure, especially as employees use unsanctioned AI tools, browser-based assistants, and emerging agentic workflows outside approved controls. Bonfy addresses this problem with a unified AI data security platform that combines contextual classification, multi-channel visibility, real-time inspection, and preventative controls across data in motion, at rest, and in use.

The challenge

Shadow AI appears when employees or teams use AI tools before governance, security review, or approved workflows catch up. It enables a major source of exposure as overshared files, AI agents, copilots, and AI-enabled workflows and the sensitive data they contain can flow into prompts, browsers, SaaS tools, and downstream automations that legacy DLP and DSPM tools were not designed to see or control.

The risk is not limited to a single prompt or AI model. The broader problem is when information moves across email, collaboration tools, file shares, SaaS apps, LLMs, browser extensions, MCP-connected services, and autonomous or semi-autonomous agents, creating more pathways for leakage, integrity failures, privacy incidents, and compliance violations.

Why legacy controls fall short

Traditional DLP, SWG, CASB, and DSPM approaches typically focus on static rules, channel-specific controls, or configuration visibility. These tools often lack business context, entity awareness, and cross-channel correlation, which leads to false positives, false negatives, and blind spots around shadow AI, AI assistants, and agent-driven workflows.

This gap is especially important for browser-based shadow AI. Traditional SWG and CASB products may see traffic but cannot understand the content context or identify entity-specific risk with enough precision to support confident action.

Bonfy's approach

Bonfy approaches shadow AI as a data security problem, not just an application discovery problem. Its platform protects unstructured data across email, files, SaaS apps, collaboration tools, AI systems, copilots, and AI agents through one entity-aware architecture designed to deliver visibility, classification, detection, automation, and preventive enforcement grounded in business context.

For shadow AI specifically, Bonfy adds a browser-extension-driven capability for real-time monitoring and protection. This is coverage for shadow AI that can prevent sensitive data release "in the wild," made possible by Bonfy's low-latency architecture that supports near-real-time decisions in browser and API-driven use cases.

How Bonfy reduces risk

Bonfy's solution can be understood in four complementary layers:

CONTROL LAYER	HOW IT HELPS WITH SHADOW AI
Discovery and visibility	Bonfy reveals where sensitive content lives, how it moves, and how humans, systems, and AI tools interact with it, helping organizations identify unsanctioned AI usage and risky data pathways.
Contextual classification	Bonfy's entity-aware engine distinguishes generic content from customer-, consumer-, or relationship-specific data, which improves accuracy and reduces false positives in AI-related detections.
Real-time inspection and prevention	Bonfy can inspect content moving through web traffic, outbound channels, and AI workflows so sensitive information can be flagged or stopped before disclosure.
Unified governance	One policy and automation engine applies controls across channels and workflows, helping organizations avoid the fragmented response that often leaves shadow AI unmanaged.

This approach matters because shadow AI is usually a symptom of broader unmanaged data movement. Bonfy's multi-channel architecture connects risk across browser activity, email, collaboration, file sharing, and AI systems so teams can move from isolated event blocking to coordinated risk management and governance.

Extending from Shadow AI to agentic AI

Shadow AI and AI agents are part of the same broader risk pattern. Agents can access multiple data sources, call external tools such as MCP servers, and take actions across systems, which expands the exposure surface beyond a single human prompt into ongoing reasoning, execution, and outbound communications.

Bonfy addresses that progression with three layers of control: input control over what data is available to the AI system, output control over what the system sends out, and new data-in-use inspection through Bonfy's MCP server during an agent's reasoning process. This is strategically important because it means Bonfy does not only detect risky behavior after the fact; it can help govern data while the AI system is actively using it.

Business value

Shadow AI is a high-value entry point for hybrid organizations that already have some tools but need coverage in emerging AI-driven gaps. Bonfy is an all-in-one foundation for greenfield buyers that want strong coverage across data at rest, in motion, and in use without the complexity and cost of stitching together many point products.

The practical outcome is faster AI adoption with lower risk and less operational friction. Organizations can start with visibility, shape governance from real data, and then progress toward automation and prevention using a single platform rather than separate tools for browsers, collaboration systems, AI apps, and agentic workflows.

Contact Us

For inquiries or additional information, please contact us at info@bonfy.ai

bonfy.ai

/bonfy-ai

bonfy.bsky.social