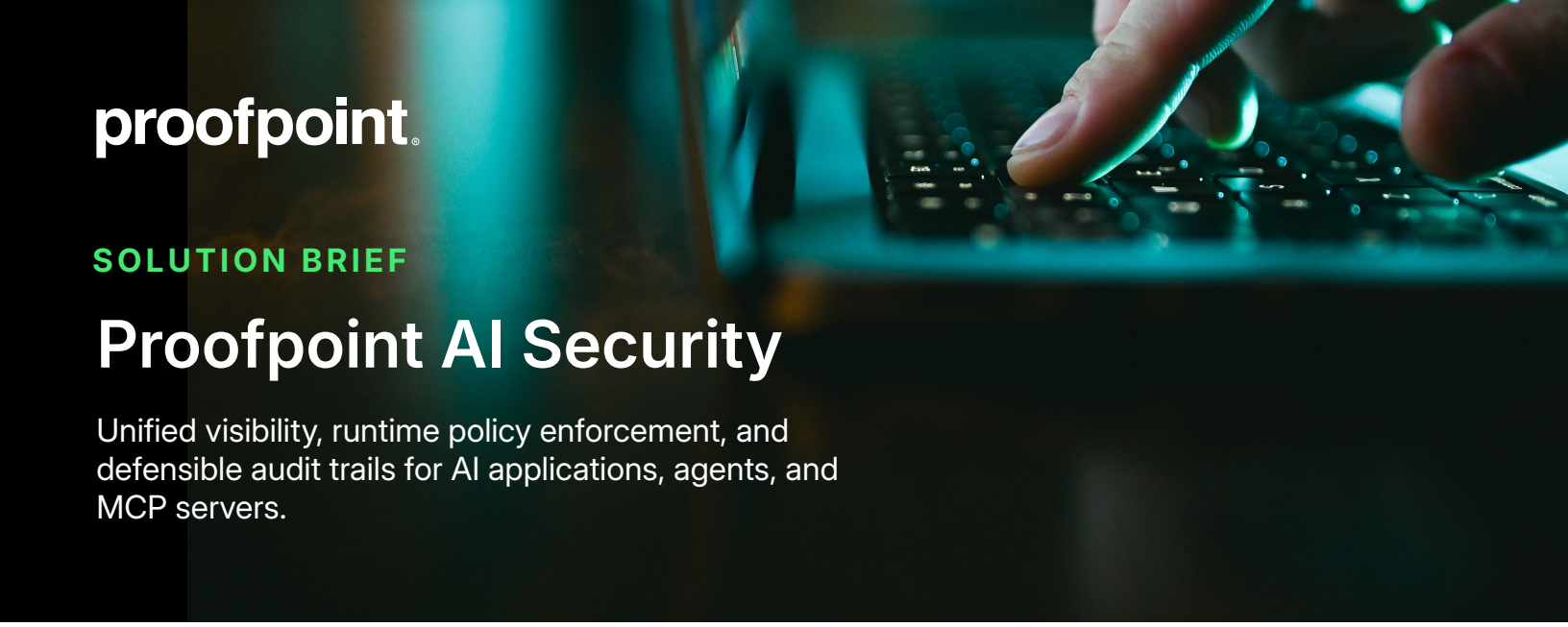

proofpoint.

SOLUTION BRIEF

Proofpoint AI Security

Unified visibility, runtime policy enforcement, and defensible audit trails for AI applications, agents, and MCP servers.

Key benefits

- **Visibility across every type of enterprise AI**, including employee tools, embedded application AI, autonomous agents, and MCP servers
- **Runtime policy enforcement** that blocks, redacts, restricts, or escalates based on what's happening in the interaction
- **Defensible audit trails** for any stakeholder: full transaction reconstruction from user request through agent action to outcome
- **SOC-ready AI telemetry** through native SIEM and SOAR integration that separates security incidents from governance violations

The problem

Many organizations can't see what AI exists in their environments, who is using it, or what data is exposed. These fundamental AI security challenges are getting harder as the technology matures.

AI tools are no longer just available through web browsers. They're now embedded in many software as a service (SaaS) applications. Increasingly, they execute autonomous workflows across enterprise systems and connect to enterprise resources through Model Context Protocol (MCP) servers. Many security teams lack visibility of what AI tools employees use and what those tools can access.

Traditional cloud access security broker (CASB) and secure access service edge (SASE) tools were not designed for these risks. Those tools can block access to known AI service URLs. But they can't examine AI prompts or moderate outputs. They also can't detect when an agent takes actions beyond the scope of the task it was given. This threat, known as semantic privilege escalation, is when every permission check passes but the agent's actions still lead to a security incident.

Permissions can be valid while intent is manipulated. And intent can be recognized while behavior drifts. Securing AI means evaluating and aligning access, intent, and behavior at the same time.

The Proofpoint approach: AI integrity

Most AI security vendors focus on a single dimension: intent detection, access control, or behavioral monitoring. However, these approaches bring the following problems:

- **Access alone** tells you what an agent can reach, but not whether it should be reaching it for a given task.
- **Intent alone** tells you what a user or agent meant to do, but not whether downstream actions stayed within scope.
- **Behavior alone** tells you what happened, but not whether it was authorized or aligned with the original request.

By contrast, Proofpoint AI Security observes, evaluates, and verifies that access, intent, and behavior align. When they do, AI operates with integrity. When they diverge, organizations see security incidents—even if individual permission checks all pass. This key difference is what separates AI integrity from AI security.

Secure every layer of your AI

Secure how humans use AI

Proofpoint discovers shadow AI across on-device apps, cloud services, and MCP servers. This gives security teams visibility of what AI tools employees are using and where sensitive data might be exposed. Proprietary detection models analyze the content and behavior of every prompt and response.

The detection models also score each interaction for risk based on data sensitivity, vendor posture, and usage frequency. When interactions violate policy, configurable acceptable use controls block or flag them per user population. This approach guides safe adoption rather than driving AI use underground.

Secure how agents use AI
Proofpoint delivers full visibility and runtime security for agents without requiring code changes, in both visibility mode and inline enforcement. Intent-based access control (IBAC) evaluates whether agent actions align with their original intended purpose. This catches semantic privilege escalation before it results in improper actions. The entire execution chain is captured in security-annotated forensics that trace every step, tool call, and data interaction back to the originating request.

Secure how AI interacts with enterprise systems
The MCP gateway is a single control point for all MCP traffic. No large language model (LLM) can connect to enterprise data sources without passing through it. A registry of 800-plus secure MCP servers with supply chain vetting can secure any new server from source in less than 15 minutes. All connections are authenticated and inspected for sensitive data, prompt injection attempts, and policy violations.

Business outcomes

Accelerated AI adoption
Security teams can confidently adopt AI because governance is built into the deployment path. AI programs move from pilot to production in weeks, not months.

Reduced risk exposure
The attack surface shrinks to what's known, governed, and actively monitored. This applies across every form of enterprise AI, from chatbots to autonomous agents.

Defensible audits without manual reconstruction
Every AI interaction produces the evidence a defensible audit requires. Compliance-readiness becomes a continuous state, not a quarterly scramble.

Security operations that scale
Native integration with security information and event management (SIEM) and security orchestration, automation, and response (SOAR) platforms delivers AI-specific telemetry with full context. Security teams can extend coverage to AI without hiring a dedicated AI security team.

Capabilities of Proofpoint AI Security

CAPABILITY	WHAT IT DOES	PROOFPOINT DIFFERENTIATOR
AI usage control	Discovery of AI usage across SaaS, IDE extensions, desktop apps, CLI tools, browser plugins, and local models. Contextual risk scoring by data sensitivity, vendor posture, and usage frequency.	AI usage governance at the semantic layer, not basic URL filtering. Risk evaluation of each interaction, not just whether a tool is sanctioned.
AI security posture management	Comprehensive inventory of AI applications, agents, tools, integrations, external services, and MCP servers. Continuous posture assessment across the full AI lifecycle.	Discovery extends to agent toolchains and MCP server inventory—mapping the full execution graph, not just chat interfaces.
AI runtime defense	Runtime inspection of AI interactions with inline evaluation of prompts, outputs, and model behavior. 18 built-in detectors across modalities.	Semantic-layer evaluation distinguishes legitimate behavior from unusual or manipulated activity—reducing false positives that erode security operations center (SOC) trust.
Agentic AI ecosystem security	Addresses semantic privilege escalation through intent-based access control (IBAC). Captures user intent and propagates it through the full agent workflow.	IBAC validates alignment at each action. Behavioral anomaly detection, agent manifests for policy-as-code, and full transaction forensics across agent chains.
AI supply chain security	MCP gateway enforces authentication and content inspection for all tool access. Discovery and registry of all MCP servers in use.	Runtime enforcement governs what data crosses MCP boundaries and what actions are permitted through protocol connections.

REQUEST A DEMO

Contact your Proofpoint account team to schedule a demo of Proofpoint AI Security, or visit proofpoint.com to learn more.