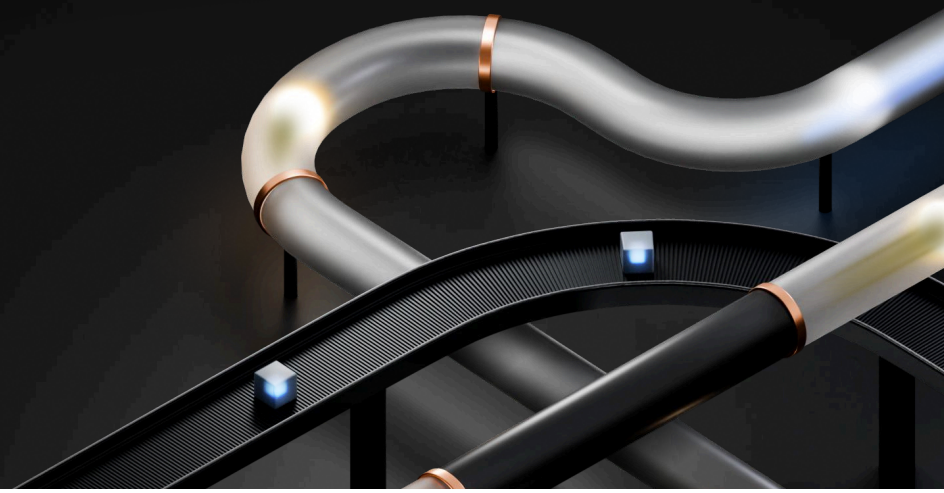
harness

Harness AI Security

Discover, test, and protect AI-native applications



Every organization today is building AI-native applications. These contain new components, such as Large Language Models (LLMs), Model Context Protocol (MCP) primitives, and AI agents that change the attack surface that security must consider. However, security teams often have little visibility into where AI components are being used.

The shift to AI-native applications is also changing the types of threats to protect against. The OWASP Top 10 risks for LLM applications includes new types of risks like prompt injection, data and model poisoning, and system prompt leakage — different from traditional risks like SQL injections, vulnerable components, and misconfigurations.

Harness AI Security provides an, end-to-end solution for securing AI applications across their entire lifecycle. It continuously discovers every AI component in your environment, dynamically tests them for AI-specific risks before deployment, and protects AI applications in production with real-time threat detection and policy enforcement.

Why Harness

Enterprise-Scale Platform
AI Security is built on Harness Web Application & API Protection (WAAP), designed for high-volume, mission-critical environments, and protecting over three trillion API calls annually.

Full Application-Layer Visibility
With Application Security Testing (AST) and WAAP, built on the Harness DevSecOps platform, AI Security provides visibility into your complete AI application and its APIs, services, and data flows.

Flexible Deployment
Deploy AI Security into any application architecture, with 30+ instrumentation methods, including both in-line and out-of-band deployments.

Actionable Signals
AI Security correlates both AI and non-AI telemetry to reduce false positives and surface real risk.

Key Capabilities

AI Discovery

Continuously discovers every LLM, MCP primitive, agent, and AI API in your environment - with real-time inventory, sensitive data flow mapping, and risk scoring for every asset.

AI Testing BETA

Automatically tests AI applications against OWASP LLM Top 10 risks before deployment, using real production traffic to configure tests and integrate directly into CI/CD pipelines.

AI Firewall BETA

Protects running AI applications in real time by detecting prompt injection, inspecting prompts and responses for data leakage, and enforcing custom AI usage policies and guardrails.

harness

Key Features

Harness AI Security delivers everything you need to discover, test, and protect AI-native applications - purpose-built to eliminate blind spots, reduce risk, and enforce policy across your entire AI footprint.

AI Asset Discovery

- Traffic monitoring: Continuously analyze AI API traffic to, from, and within an application environment.
- MCP discovery: Inspect traffic between MCP servers and clients to discover 1st-and 3rd-party MCP components, and detect and threats.
- SSE support: Stitch together and analyze AI streaming responses sent over HTTP+SSE to detect threats.
- AI inventory: See discovered AI assets with relevant details, inc. vendor, model, and sensitive data types.

Dynamic Testing BETA

- Seamless CI/CD integration: Orchestrate testing across every pipeline to test every AI application before deployment.
- OWASP Top 10 protection: Test AI assets against OWASP LLM Top 10 risks and other AI-specific vulnerabilities.
- Automatic test configuration: Simplify configuration and reduce overhead by creating tests using your actual AI traffic.

AI Governance

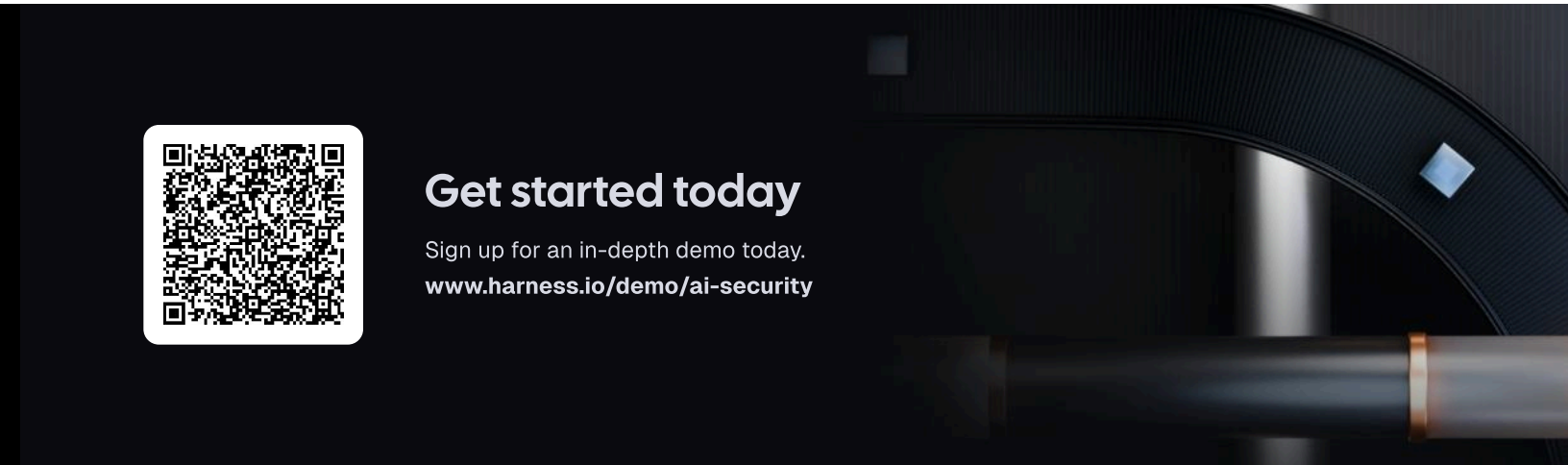
- OOTB policies: Enforce best practices and ensure compliance with organizational policies around securing any discovered AI assets.
- Custom policies: Customize polices to measure compliance against unique organizational standards.
- Integrations: Integrate with existing workflows and streamline AI risk governance by converting findings into tracked tickets on ServiceNow ITSM and JIRA.

AI Protection BETA

- Prompt injection detection: identify prompt injections through direct, indirect, and MCP/tool injection.
- Data leak prevention: Inspect prompts for harmful outputs, data exposure or leakage, risky tool invocations, and code.
- AI Guardrails: Create custom policies to enforce AI usage, including well-formed prompts, token limitations, etc.

AI Security Posture

- Vulnerabilities: Continuously monitor discovered AI assets integrations to identify AI-specific vulnerabilities.
- Sensitive data flows: Identify PII and other sensitive data flowing through AI APIs or MCP endpoints.
- Risk scoring: Proprietary risk score based on authentication, encryption, Internet exposure, sensitive data flow, vulnerability and compliance issues, and more.
- AI security research team: Harness security researchers analyze real-world attack patterns and public breaches to power proprietary risk scoring.



Get started today

Sign up for an in-depth demo today.
www.harness.io/demo/ai-security