DATASHEET

Zaperon

Zaperon AI Governance

Adopt AI with Confidence, Without Exposing Sensitive Enterprise Data



AI is Transforming Enterprises

The enterprise tech landscape is shifting from static SaaS tools to Generative AI (ChatGPT, Gemini, Claude) and autonomous AI agents (AutoGen, CrewAI, LangChain agents, custom Copilots). Employees use AI assistants to accelerate productivity, developers integrate AI APIs into applications, and AI agents increasingly perform tasks autonomously—from analyzing documents to interacting with enterprise systems.

91%

of enterprises are already using AI in different business functions

-McKinsey Report 2026

80%

of enterprise softwares have already embedded AI agents

-Gartner 2026

$2.6 Trillion

Global AI spending in 2026 and is expected to cross $3.5 Trillion in 2027

-Stanford University 2026

AI has also Introduced New Security Risks

While AI is accelerating innovation, it is also creating an entirely new attack surface. The autonomous AI agents operate with privileges, can browse websites, access enterprise applications, execute workflows, submit forms, and interact with sensitive business data with minimal human intervention. These capabilities significantly expand the enterprise attack surface and introduce risks that traditional endpoint, network, and identity security solutions were never designed to address.

Where the New AI Attack Surface Lies

- Prompts & conversations
- File uploads
- Clipboard & copy-paste
- Browser extensions
- API keys & agent identities
- Personal AI accounts

Key AI Security Risks

1. **Shadow AI & Unsanctioned Execution:** Employees are rapidly adopting unauthorized AI applications, browser extensions and AI agents outside the visibility of IT and security teams. Many also access public GenAI services using personal accounts that lack enterprise security and privacy controls, creating significant governance blind spots. According to IBM's Cost of Data Breach 2026 report, 85% of organizations have little or no visibility into employee AI usage, making it difficult to enforce policies or monitor sensitive data exposure.

2. **Non-Human Identity (NHI) Privilege Creep**: AI agents operate using API keys, OAuth tokens, service accounts, and other non-human identities to access enterprise resources. Without proper governance, these identities often accumulate excessive privileges, increasing the risk of unauthorized access, credential compromise, lateral movement, and data exfiltration.

3. **Sensitive Data Leakage**: Employees frequently share source code, intellectual property, customer records, financial information, credentials, and other confidential data with public AI models. This information may be retained, processed, or used by third-party AI providers, creating significant data privacy, regulatory, and compliance risks. According to the 2026 Verizon Data Breach Investigations Report (DBIR), more than 35% of data shared with AI tools contains sensitive enterprise information, with source code emerging as the most frequently exposed data type.

4. **Emerging AI-Native Threats**: AI introduces entirely new attack vectors that traditional security controls cannot detect or prevent. Threats such as prompt injection, AI-assisted phishing, deepfake-based social engineering, malicious AI-generated content, and insecure AI integrations are rapidly becoming common attack techniques. The OWASP Top 10 for LLM Applications identifies Prompt Injection as the most critical AI security risk, where attackers manipulate AI systems to reveal confidential information, misuse connected tools, execute unintended actions, or exfiltrate sensitive data without the user's knowledge.

Why are Traditional Security Tools Blind to AI?

Traditional security tools were built for an era of static files, known malware signatures, and rigid network perimeters. They were never designed to understand natural language interactions, autonomous AI agents, or the dynamic flow of data between users and Large Language Models (LLMs). While these solutions continue to protect enterprise networks and applications, they lack the context and visibility required to govern AI usage.

As organizations rapidly adopt Generative AI and AI agents, legacy security controls leave critical blind spots:

TRADITIONAL DLP CANNOT UNDERSTAND AI CONTEXT

Conventional **Data Loss Prevention (DLP)** solutions are designed to detect structured data such as credit card numbers, Social Security numbers, or file attachments. They struggle to identify unstructured intellectual property, proprietary source code, confidential business strategies, or sensitive information embedded within AI prompts and conversations.

SECURE WEB GATEWAYS ONLY SEE THE DESTINATION

Legacy **Secure Web Gateways (SWGs)** determine *where* users are connecting, but not *what* they are asking AI models. They cannot distinguish between a harmless prompt and one containing confidential business data, customer information, or proprietary code.

NETWORK SECURITY CANNOT INSPECT AI CONVERSATIONS

Traditional **Firewalls** and network security controls inspect packets and protocols—not the intent or meaning of natural language. They cannot detect prompt injection attacks, malicious instructions hidden in documents, or attempts to manipulate AI agents into exposing sensitive information.

IDENTITY CONTROLS STOP AT AUTHENTICATION

Conventional **Identity & Access Management (IAM)** and **Single Sign-On (SSO)** solutions verify who can access an AI application, but provide little visibility into what happens after login. They cannot monitor prompts, file uploads, AI-generated responses, or the data exchanged during an AI session.

ENDPOINT SECURITY MISSES SHADOW AI

Enterprises are increasingly developing custom AI agents and AI-powered applications that connect to internal systems, access enterprise data, and automate business workflows. Unlike conventional AI chatbots, these agents may operate through APIs, embedded interfaces, browser extensions, or custom applications, making them difficult for traditional endpoint security and DLP solutions to identify. This can create blind spots that enable Shadow AI, unauthorized data sharing, and ungoverned AI activity.

The Need for AI-Native Governance

As AI adoption accelerates, these blind spots create significant security, compliance, and operational risks. Organizations need an **AI-native governance layer** that extends beyond traditional security controls to provide continuous discovery of AI applications and AI agents, real-time monitoring of AI interactions, protection against sensitive data leakage, and policy enforcement based on user identity, device posture, and data sensitivity.

Zaperon AI Governance bridges these gaps by delivering comprehensive visibility, intelligent policy enforcement, and identity-centric controls that enable organizations to embrace AI securely—without compromising productivity or innovation.

Zaperon Approach for AI Era

As organizations rapidly adopt GenAI, copilots, AI agents, and AI-powered applications, traditional security controls are no longer sufficient. Zaperon Identity-centric SSE governs enterprise AI usage across four critical dimensions: **Identity, Device, AI Application, and Data**.

Zaperon offers **Unified AI Governance** where every AI interaction is evaluated against these four pillars before access is granted or enterprise data is shared. Zaperon continuously verifies **who is accessing AI, from which device, which AI application is being used, and what data is being shared**—enabling organizations to securely adopt AI without compromising security, compliance, or productivity.

1 IDENTITY:

Who is using AI?

- Verify the user's identity
- Check user authorization based on role, policies, and privileges

2 DEVICE:

Is the device trusted?

- Verify device identity and authorization against user
- Continuously validate 15+ security posture parameters - IP, antivirus, OS, location, disk encryption and device password.

3 AI APPLICATION:

Which AI application is being accessed?

- Discover GenAI chatbots, embedded and custom AI agents in real-time
- Enforce tenant isolation to allow corporate workspaces & block personal accounts

4 DATA PROTECTION:

What enterprise data is being shared?

- Inspect prompts, clipboard activity, and file uploads to AI apps
- Prevent leakage of PII, PCI and other sensitive business data from AI interactions, including those involving custom AI agents and AI-powered applications



Core Zaperon AI Governance Features

SHADOW AI DISCOVERY

Discover and identify unauthorized or unsanctioned AI applications being accessed across the organization. Gain visibility into GenAI chatbots, copilots, custom AI agents, browser-based AI tools, and embedded AI services, along with user and application activity. Classify AI applications based on risk and bring shadow AI usage under enterprise governance.

IDENTITY-AWARE AI ACCESS CONTROL

Enforce granular AI access policies based on user identity, role, group, device posture, location, application, and risk. Allow organizations to control who can access specific AI applications, models, tenants, or AI capabilities. Prevent unauthorized users and unmanaged identities from accessing enterprise AI resources.

INTERACTION MONITORING & COMPLIANCE

Monitor interactions between users and AI applications to provide visibility into AI usage and policy compliance. Track prompts, responses, uploads, downloads, copy/paste activity, and other AI interactions. Detect policy violations and maintain detailed audit trails for security, governance, and compliance requirements.

DATA PROTECTION

Protect sensitive enterprise information from being exposed to AI applications. Inspect prompts, uploads, clipboard activity, and other data flows to detect sensitive information such as PII, PCI, credentials, confidential business data, and intellectual property. Apply configurable policies to allow, block, redact, or mask sensitive data before it reaches AI services.

PROMPT INJECTION PREVENTION

Detect and prevent malicious or manipulative instructions designed to influence AI models or bypass established security controls. Identify suspicious prompt patterns and enforce policies to reduce risks associated with prompt injection, malicious instructions, and attempts to manipulate AI behavior.

SECRETS DETECTION

Identify sensitive secrets and credentials before they are submitted to AI applications. Detect API keys, access tokens, passwords, authentication credentials, private keys, and other confidential secrets using configurable detection policies. Block, mask, or alert on attempts to expose secrets through AI interactions.

DOS DETECTION

Detect abnormal or excessive AI activity that may indicate automated abuse, resource exhaustion, or denial-of-service attempts. Monitor AI request patterns, frequency, and usage behavior to identify anomalous activity and enforce controls to protect AI services and enterprise resources.

AI TOKEN USAGE MONITORING

Monitor AI token consumption across users, applications, teams, and AI models. Identify unusual or excessive token usage, establish usage thresholds, and provide visibility into AI consumption and associated costs. Help organizations optimize AI usage while preventing misuse and uncontrolled AI expenditure.

Key Use Cases for Modern Enterprises

Practical controls for secure AI adoption across employees, developers, security teams and enterprise applications.

USE CASE	HOW ZAPERON HELPS
Prevent sensitive data leakage	Stop employees from submitting confidential organizational information to public AI tools. Policies can inspect content and file types and block sensitive data before it leaves the protected endpoint.
Govern AI agents	Discover AI agents running or being used on endpoints and introduce an approval workflow so only administrator-approved agents can be used for business activities.
Protect PII and credentials	Detect PII, API keys, passwords and tokens inside AI prompts and Log, Warn or Block the activity according to organizational policy.
Defend against prompt injection	Identify prompts attempting to override system instructions or jailbreak an AI model, helping reduce the risk of malicious or unintended AI behavior.
Prevent AI resource abuse	Detect prompts designed to consume excessive tokens or resources and enforce policy controls to reduce potential Denial-of-Service or resource-exhaustion scenarios.
AI usage visibility & audit	Maintain visibility into user AI activity and policy events. Security teams can use logs to investigate incidents, understand AI adoption and support governance and compliance processes.
Secure AI adoption without blocking productivity	Use graduated controls—Log, Warn and Block—to allow legitimate AI usage while applying stronger enforcement to risky prompts, sensitive data and unauthorized AI tools.

About Zaperon

Zaperon is an Identity-centric Security Service Edge (SSE) platform designed to secure enterprise access to applications, data, AI services, and cloud infrastructure from any user or device, regardless of location or network. By extending Zero Trust principles to enterprise AI, Zaperon enables organizations to adopt Generative AI securely through continuous identity verification, device trust, contextual access controls, and intelligent data protection. As AI Governance is natively integrated into the Zaperon SSE platform, organizations gain centralized visibility, consistent policy enforcement, and unified control over AI usage without deploying standalone security solutions. By minimizing data exposure, preventing unauthorized AI interactions, and providing real-time governance across users, devices, and AI applications, Zaperon empowers enterprises to accelerate AI innovation while keeping sensitive information protected, compliant, and under complete organizational control.

+1 202 600 5234
+91 8800 170387
sales@zaperon.com
www.zaperon.com

© 2026 Zaperon, Inc.
All Rights Reserved.

