AI Security: Your competitive advantage in the AI race

A CxO guide to accelerate AI adoption while managing risk

Introduction

As organizations rush to harness the transformative power of AI, adoption often outpaces security. For every sanctioned initiative, there are countless instances of shadow AI by individuals and teams. In fact, 85% of IT decision-makers say employees are adopting AI tools even before their IT teams can assess them.[1] And 93% of employees admit to putting information into AI tools without approval.[1]

Meanwhile, AI-native attack vectors are surging, up 47% over the past year,[2] as traditional security tools struggle to keep up. New forms of data exposure and compliance gaps challenge existing governance practices. Executive leaders face a critical question: How do you manage the risks introduced by AI without blocking the innovation it enables?

The requirements are clear. Organizations need to create an environment where:

- All AI usage is known
- All AI communications can be secured
- All AI policies can be enforced
- All AI models can be protected from abuse

Cloudflare AI Security Suite gives organizations the confidence to move faster with AI by removing the uncertainty around risk. Our unified platform secures your entire AI lifecycle by discovering AI usage, applying zero trust access, and protecting web and API endpoints with proactive threat defense. With integrated data governance, teams can innovate freely without compromising security.

85%
of IT decision-makers say employees are adopting AI tools even before their IT teams can assess them.[1]

The AI security challenge

The risks of AI adoption are fueling a fast-growing market for AI security solutions. For instance, cloud-native application protection platforms (CNAPPs) have focused on AI development workflows. Early offerings such as these are an essential line of defense, but they’re not enough. Companies need to secure the entire AI lifecycle, including protecting AI systems once they’re running in production.

The full scope of the AI security challenge is becoming all too clear. Today, developers are building AI features, employees are using external AI tools, and customers are interacting with AI-powered applications. Securing all these different environments manually is complex, and doing it consistently is even harder.

Preventing workforce misuse of AI

As the organizational workforce adopts both sanctioned and unsanctioned AI tools, sensitive data and operations are at risk. Security teams have to deal with:

- **Lack of visibility into AI usage:** Leaders often lack a complete picture of what AI tools their employees are using, where sensitive data is being processed, and how these AI systems connect to business applications. This blind spot creates significant risk exposure.
- **Data security and compliance risks:** AI changes how data flows through your organization. Personal information, proprietary data, and customer records can end up in AI systems in ways that trigger compliance violations or expose competitive intelligence.
- **Access controls for agentic AI workflows:** It's not just human access to AI tools that must be managed. Agentic AI access to MCP servers and other critical systems must also be managed. This requires new approaches to identity management and access controls.

Protecting public-facing AI applications and models

Whether developed internally or consumed via a third party, AI systems are becoming an essential pillar of the customer and user experience, and need to be protected as such. The OWASP Top 10 for LLM Applications highlights threats that traditional tools cannot address:

- **Unbound consumption attacks:** Similar to traditional DoS attacks, unbound consumption attacks seek to overwhelm LLMs with resource-intensive requests. Pay-per-use cloud pricing models can send the financial impact of such attacks soaring, while legitimate users face degraded service quality.
- **Model poisoning:** Attackers implant backdoors, biases, or vulnerabilities in LLMs by injecting corrupted data into public datasets or repositories used by developers for model training. A model poisoned in this way behaves normally until specific triggers activate harmful behaviors.
- **Prompt injection attacks:** A popular tactic for data exfiltration through AI interfaces, prompt injection manipulates LLM inputs by embedding malicious instructions within user prompts or external content, causing models to ignore original instructions and execute attacker commands instead.
- **Jailbreaking:** Carefully crafted prompts such as role-playing scenarios, instruction override commands, and multi-turn strategies can bypass LLM safety guardrails to generate prohibited content or extract sensitive information.

To allow rapid AI innovation without putting the organization at risk, we recommend leaders adopt an approach to AI security that spans the full spectrum of requirements:

- Protecting workforce use of GenAI
- Protecting AI-powered apps and workloads
- Building AI-powered apps that are secure by design

Securing AI development and training workflows

AI projects involve massive datasets, costly resources, and iterative experimentation — all of which create new attack surfaces and vulnerabilities. Security teams must address:

- **Training data security and integrity:** Compromised training data can introduce biases, backdoors, or vulnerabilities that persist in production models. Organizations must secure access to training datasets, prevent unauthorized modifications, and ensure data provenance throughout the model lifecycle.
- **Credentials and secrets management:** AI development workflows require access to systems including cloud storage for datasets, compute clusters for training, model registries, and third-party services. Inadequately secured secrets or API keys can expose sensitive credentials, enabling unauthorized access to proprietary models, training data, or production systems.
- **Development environment access controls:** AI engineers often need elevated privileges to experiment with models and access sensitive data. Without proper access controls, this can lead to insider threats, accidental data exposure, or unauthorized model extraction.

Unifying security with Cloudflare

Successful AI adoption focuses on enabling productivity, not restricting it. When teams can use AI confidently, knowing the proper safeguards are in place, they innovate faster and take on more ambitious projects.

Cloudflare AI Security Suite creates a safe environment for AI innovation. By unifying secure access service edge (SASE) and web application security capabilities, CxOs can connect and protect two fundamental domains:

- External, public-facing AI-enabled applications
- Internal, private AI systems and workloads

Focusing on the entire AI lifecycle, Cloudflare AI Security Suite addresses security needs from discovery and risk management to data protection, securing access for users, and protecting AI-enabled applications and development workflows. To provide an essential layer of real-time production security, the Cloudflare global network sits inline to inspect and filter every AI interaction, protecting data across all users and applications.

Instead of just detecting problems after they happen, Cloudflare prevents issues and blocks threats before they can reach AI models. Security teams gain the visibility they need to stay ahead of emerging threats.

Core capabilities of Cloudflare

Cloudflare AI Security Suite brings comprehensive monitoring, real-time protection, and proactive risk management under a single umbrella for a holistic approach to AI security.

Comprehensive AI discovery and visibility

Effective AI security depends on a complete, real-time inventory of both sanctioned and unsanctioned AI resources and usage. The foundation of Cloudflare AI Security Suite is continuous monitoring and automated discovery to identify all AI models, assistants, agents, and shadow AI deployments across every type of environment — public, private, or internal.

Proactive AI risk management

Cloudflare AI Security Suite helps organizations prevent attacks by detecting and mitigating AI-specific vulnerabilities, misconfigurations, and attack paths, including those in the OWASP Top 10 for LLMs. Application confidence scoring helps prioritize remediation so teams address the most significant risks first.

Application security for AI-powered apps

To help SecOps stay up to date on the latest AI threat vectors, Cloudflare AI Security Suite includes proactive threat detection and mitigation for AI-specific vulnerabilities, misconfigurations, and attack paths within AI pipelines, including protection against prompt injection, data poisoning, and model abuse.

Specialized AI firewalls discover and label generative or agentic AI and API endpoints, detect attempts to exfiltrate PII, and block malicious prompts before they impact AI model performance or poison the model with toxic content or misinformation.

Zero trust access for GenAI and agentic AI workflows

Zero trust principles like least privilege apply to the workforce and AI agents alike. Cloudflare AI Security Suite can enforce zero trust network access policies (ZTNA) for both human-to-AI and AI-to-AI interactions. Centralized logging and controls for MCP servers ensure that agentic AI only accesses what it's been authorized for — including just-in-time workflows.

AI-aware data protection

Effective AI security requires a data loss prevention capability — one that leverages multiple language models to understand the content of a prompt as well as the intent behind it. Cloudflare AI Security Suite incorporates data loss prevention (DLP) capabilities throughout training, prompts, and responses to prevent PII exposure, data leakage, and unauthorized access within AI models and pipelines. Deployed inline, API-centric runtime security acts as a fast, simple first layer of defense to complement the shift-left approach enabled by a CNAPP.

Data localization

LLMs and AI applications fall under the same regulations as any other type of data environment. Cloudflare AI Security Suite helps organizations ensure that AI workloads respect geographic and jurisdictional boundaries with policies that keep training data and inference requests within approved regions.

Security by design for AI development

As with all types of software, security for AI apps should be built in from the beginning of the development cycle, not bolted on afterwards. Cloudflare AI Security Suite provides developers with tools and frameworks to build AI-powered apps that are secure by design.

The business impact of Cloudflare AI Security Suite

The rise of AI isn't just an evolution — it's a fundamental disruption on the order of industrialization or computerization. That makes fast and effective adoption a matter not just of driving growth, but of organizational viability. Slow or unsafe adoption is now an existential threat to organizations across entire industries and sectors.

Cloudflare AI Security Suite can enable safe, controlled, and efficient transformation to meet accelerating customer expectations and market requirements in intensely competitive environments.

- **Faster AI innovation:** When security enables safe usage rather than blocking adoption, employees and teams can explore new AI apps to become more productive in their daily work. Proper security frameworks give developers confidence to build ambitious AI features without jeopardizing sensitive data or systems.
- **Reduced AI-related risk:** Organizations can manage the full spectrum of risks that come with AI adoption, including the AI-related risks inherent in AI-enabled web applications. AI apps under development can be secured to ensure that employees don't leak sensitive data or expose it in AI training sets. SecOps can proactively identify and mitigate AI-specific threats and vulnerabilities, minimizing the attack surface and protecting critical data and models.
- **Streamlined security operations:** Centralized visibility and control over your AI security posture simplifies management and streamlines incident response. Your SecOps team can focus on strategic initiatives instead of constantly firefighting AI-related incidents.
- **Robust data governance and compliance:** AI-specific data protection controls help you secure sensitive information and meet changing regulatory requirements across the AI lifecycle.
- **Lower total cost of ownership (TCO):** Leveraging a consolidated platform that integrates existing security investments is more economical than implementing separate point solutions for every AI security challenge.

Model use cases for Cloudflare AI Security Suite

Cloudflare AI Security Suite is ideally suited to address essential requirements around AI adoption.

- **Securing workforce AI tool usage:** Enforce zero trust policies for workforce access to both public generative AI tools like ChatGPT and internally developed AI-powered apps.
- **Protecting public-facing, AI-enabled apps:** Safeguard web applications and APIs that incorporate AI models — such as chatbots and recommendation engines—from attacks that could expose sensitive data or abuse the model.
- **Shadow AI management:** Automatically discover unsanctioned AI tools across your organization and applying appropriate controls to allow continued innovation within managed risk boundaries.
- **AI-powered DLP for AI interactions:** Prevent sensitive data from being exposed in AI prompts or responses, ensuring PII and confidential information remain protected.
- **AI development security:** Provide engineering teams with frameworks that make secure development the default approach, enabling them to build AI features quickly without compromising security.

Considerations for implementation

As a foundational security strategy, AI security should be approached thoughtfully.

Start with your current infrastructure	The most successful AI security implementations build on existing SASE and application security tools rather than replacing them. This approach leverages current investments while extending protection to cover AI-specific risks.
Deploy unified inline protections	The ability to block malicious activity at the network edge, as it happens, is crucial for AI security. Deploy real-time, inline controls, and complement them with API-based monitoring.
Ensure complete coverage	Your AI security strategy should address the full spectrum of requirements: workforce use of GenAI tools, protection for AI-powered apps and workflows, securing agentic AI workflows, and security for AI development workflows.
Plan for enterprise scale	Choose solutions that can grow with your AI adoption. What works for one pilot program must also scale across the entire organization as AI usage expands.
Validate your success criteria	Before making a full commitment, validate the solution in real-world scenarios. Choose a platform that allows for free, self-serve activation of its enterprise capabilities. This lets you test the full security suite on a small scale — like for a single team or app — to rapidly prove its value and ensure it meets your success criteria.

Taking the next step with Cloudflare AI Security Suite

As AI adoption accelerates worldwide, the organizations who can scale AI securely will have a significant competitive advantage. The key is recognizing that AI security isn't about preventing AI use — it's about enabling it intelligently.

Cloudflare AI Security Suite allows organizations to innovate with confidence. Built on and extending our widely adopted SASE and Application Security platforms, our AI Security Suite provides integrated AI discovery, zero trust access controls, intelligence-led proactive threat defense, and robust data governance. By securing AI usage and models from all angles, organizations can empower developers to build faster, and enable employees to be more productive, all without sacrificing the end-user experience.

Schedule a consultation

Explore how Cloudflare AI Security Suite can transform your organization's approach to secure AI adoption.

→ **1 888 99 FLARE**

✉ **enterprise@cloudflare.com**

🌐 **www.cloudflare.com**

1. ManageEngine. The Shadow AI Surge in Enterprises: Insights from the US & Canada
2. Check Point Research. Q1 2025 Global Cyber Attack Report from Check Point Software: An Almost 50% Surge in Cyber Threats Worldwide, with a Rise of 126% in Ransomware Attacks