Cloudflare

CLOUDFLARE AI SECURITY

Securing AI Across Your Entire Organisation

Jason Clarke · Cloudflare · Western Australia

THE CHALLENGE

AI Adoption Is Outpacing Security

Three distinct threat surfaces have emerged — each requiring a different response. Most organisations are exposed on all three simultaneously.

Your People

Employees are using ChatGPT, Copilot, and Claude every day — often without IT visibility. Corporate data, customer PII, and source code are leaving the organisation in prompts.

Your Applications

Development teams are building AI-powered products that call LLM APIs directly. Without a control layer, there is no visibility into cost, data exposure, or what happens when a provider goes down.

Your AI Agents

AI agents now take actions — reading files, sending emails, querying databases. Without governance, a single compromised agent can exfiltrate data or execute unauthorised actions at machine speed.

The common thread: Traditional perimeter security was not designed for AI. Cloudflare addresses all three surfaces from a single platform — without adding new vendors or complexity.

Cloudflare

CLOUDFLARE AI SECURITY

Three Pillars. One Platform. One Dashboard.

PILLAR 1

End-User Protection

Employees using AI tools in their browser

Products:

Cloudflare One · Gateway · DLP · Browser Isolation · CASB

Outcome: Shadow AI visibility + data loss prevention

PILLAR 2

App & API Security

Applications calling LLM APIs programmatically

Products:

AI Gateway · DLP on prompts & responses · Rate limiting · Observability

Outcome: Cost control + resilience + full audit trail

PILLAR 3

Agentic & MCP Security

AI agents accessing tools, APIs, and data

Products:

MCP Portal (open beta) · AI Gateway · AI Security for Apps · Gateway

Outcome: Zero Trust governance for every agent action

Cloudflare

PILLAR 1 — END-USER PROTECTION

Let Your People Use AI. Safely.

The Business Risk Today

When employees use AI tools without IT oversight:

- Customer PII enters third-party AI systems
- Source code and IP shared in prompts
- No audit trail for compliance or breach response
- No way to enforce policy across 40+ AI applications

75% of employees use AI tools at work. **60%** do so without IT approval.

Industry surveys, 2024

What Cloudflare Delivers

Complete Visibility

See every AI tool in use across your organisation — including tools IT didn't approve

Data Loss Prevention

Scan prompts and file uploads for PII, credentials, and proprietary data before they reach any AI provider

Granular Policy

Allow, block, or restrict AI tools by user, group, device, or application — without blocking productivity

Compliance Audit Trail

Full conversation logging with user attribution — supports Privacy Act NDB obligations and internal investigations

Cloudflare

PILLAR 2 — APP & API SECURITY

Control Every LLM Call Your Applications Make

The Business Risk Today

When applications call LLM APIs without a control layer:

- A single runaway agent loop can cost $10,000+ in minutes
- No visibility into what data is being sent to which model
- Single provider dependency — one outage breaks your product
- Prompt injection attacks can hijack application behaviour

AI Gateway core features are free — observability, caching, and rate limiting at no additional cost on any Cloudflare plan.

What Cloudflare Delivers

Cost Control

Rate limits and budget caps per model, per user, or per application. Semantic caching cuts repeat API costs by up to 90%

Resilience

Automatic failover across 20+ LLM providers. If OpenAI is down, route to Anthropic or Workers AI — your application keeps running

Full Observability

Every prompt and response logged — token counts, latency, cost per request, and guardrail triggers. No more black-box AI spend

DLP on Both Directions

Scan prompts before they reach the LLM and responses before they reach your users — no TLS interception required

Cloudflare

PILLAR 3 — AGENTIC & MCP SECURITY

AI Agents Take Actions. Govern Them.

The Business Risk Today

AI agents connected to your systems via MCP can:

- Access tools and data beyond their intended scope
- Be hijacked by malicious data returned from a tool call
- Connect to unvetted third-party MCP servers
- Exfiltrate sensitive data at machine speed, without human review

MCP (Model Context Protocol) is the emerging standard for connecting AI agents to external tools. Adoption is accelerating — governance frameworks are not keeping pace.

What Cloudflare Delivers

Governed Tool Access

MCP Portal (open beta) gives agents a single, identity-enforced gateway to approved tools — only vetted servers are accessible

Shadow MCP Detection

Cloudflare Gateway detects and blocks employees connecting AI clients to unauthorised MCP servers outside the approved portal

Inbound App Protection

WAF-based AI Security scores every inbound prompt for injection risk and PII — protecting your public-facing AI applications

Complete Audit Trail

Every tool call and every LLM call logged centrally — who did what, when, with which agent. Supports board-level accountability requirements

Cloudflare

WHY CLOUDFLARE

One Platform. Proven at Scale. Ready Now.

227%
ROI over 3 years
Forrester TEI, Jan 2026

<6 mo
Payback period
Forrester TEI, Jan 2026

35%
Reduction in IT ops time
Forrester TEI, Jan 2026

38%
of Fortune 500 are
Cloudflare customers

Single dashboard. Zero Trust, AI Gateway, and WAF are managed from one place — no separate consoles, no separate vendors, no integration projects.

→ Suggested Next Steps

1. **Discovery Workshop (1 hr)**
   Map your current AI tool usage, application LLM calls, and any agentic workflows in flight — identify which pillar is most urgent

2. **Proof of Concept (2–4 weeks)**
   Deploy the highest-priority pillar in your environment — measure shadow AI discovery, cost reduction, or agent governance against your current baseline

3. **TCO Review**
   We'll build a model specific to your environment — headcount, AI spend, and existing security tooling — using the Forrester TEI methodology