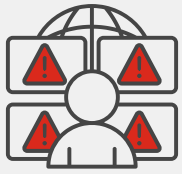
SOLUTION BRIEF

Secure and Govern AI Applications and Agent Traffic with FortiAIGate

Executive Summary

Nearly 9 in 10 organizations now use AI regularly in at least one business function.[1] And these AI applications and agents call tools, APIs, and Model Context Protocol (MCP) servers. Each new connection opens a path for data loss, misuse, and runaway cost.

FortiAIGate is a self-hosted AI gateway for large language models (LLMs), agents, and MCP tools. From one control point, organizations can stop AI attacks and data leaks, enforce usage policy, and cap spend by organization, workspace, project, and user. Guardrails block harmful prompts, outputs, and MCP tool calls before they cause damage. Prompts and data stay safely inside your network. FortiGate detects shadow AI, and FortiAIGate inspects it, blocking or redacting risky content, so your organization can safely adopt AI at its own pace.



More than 20% of organizations reported a breach that targeted AI models or applications in 2026.[2]

AI Risk Outpaces Current Controls

AI has moved from answering questions to taking action. Agents read email, query data, and call tools on a user's behalf. Attackers target the gap between the prompt and the output.

Network firewalls, web application firewalls, and endpoint tools inspect IP addresses, signatures, and payloads, but they do not read the meaning of a prompt or a model response. Many AI security tools inspect the user prompt and the final output, but attacks hidden in tool responses and tool calls pass through unseen.

Separate point tools for guardrails, data loss prevention (DLP), and cost tracking spread governance across disconnected consoles. The result is gaps in coverage, rising costs, and no single record for your auditors.

One Control Point Between AI Applications, Agents, and Models

FortiAIGate runs as a secure reverse proxy on your infrastructure. AI applications and agents send each request to FortiAIGate, and FortiAIGate forwards it to the model. FortiAIGate works with public, private, and self-hosted models, regardless of where they run. Guardrails inspect each prompt, response, tool call, and tool response in context. Intelligent routing sends each request to the right model, based on prompt content, the requested model, and HTTP headers.

Applications and agents connect through one OpenAI-compatible API, so no re-architecture is required. FortiAIGate gives security teams:

- **Clear risk coverage:** Guardrails aligned with the OWASP Top 10 lists for LLM, MCP, and agentic applications
- **One source of truth:** Logs that capture the user, the rule triggered, and the full prompt and response
- **High-performance guardrails:** Dynamic inspection and low-latency guardrails to remove bottlenecks, with acceleration across multiple graphics processing units (GPUs) for high-volume traffic
- **Simple management:** A built-in graphical user interface that sets policy with no Python coding

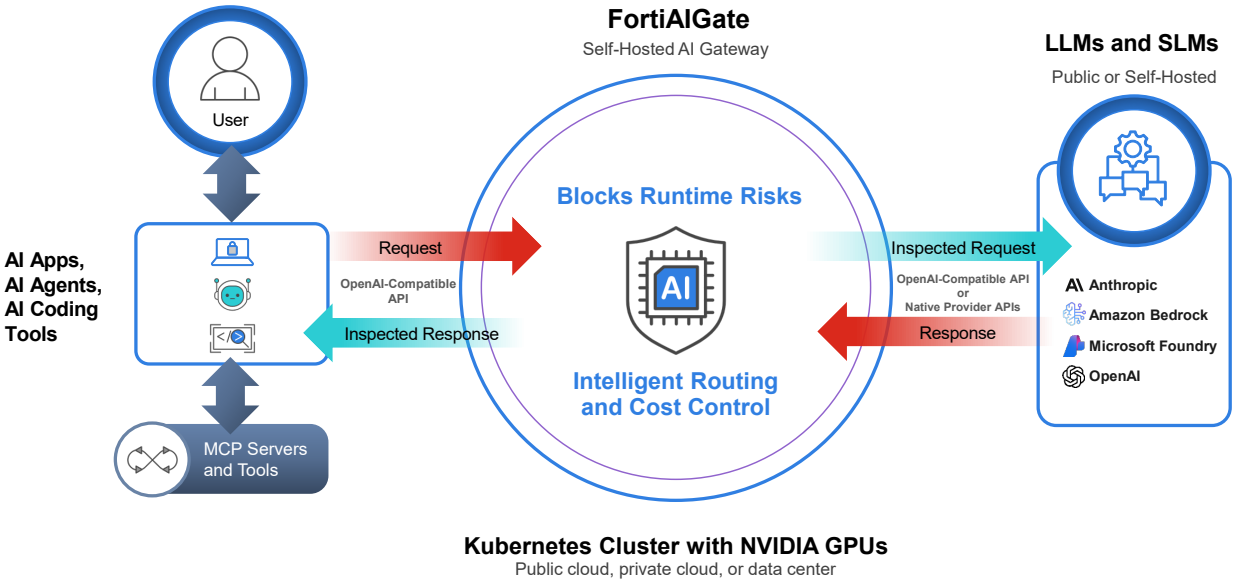


Figure 1: FortiAIGate secures and governs AI traffic between applications, agents, and models.

FortiAIGate Critical Capabilities

FortiAIGate addresses the new security, governance, and operational challenges that arise when organizations deploy AI models, MCP tools, and agents. Specifically, it enables organizations to:

Stop AI attacks and data leaks

FortiAIGate closes the security gap that exists between the tool response and tool call. Guardrails inspect MCP tool responses and tool calls along with prompts and outputs. Multilayered guardrails stop injection, jailbreaks, toxic content, model poisoning, and insecure or malicious code generation. The guardrails inspect text, images, audio, video, and code. DLP detects 35 types of personally identifiable information and confidential data, then alerts, denies, or redacts. Anomaly detection flags insider threats and data exfiltration attempts.



46.9% of enterprises report AI spend over budget.[3]

Cap AI spend by organization, workspace, project, and user

To keep track of AI costs, FortiAIGate shows AI spend by organization, workspace, project, and user, and it stops spending at the limit you set. Budgets reset daily, weekly, or monthly. Once spending exceeds the limit, FortiAIGate blocks requests. Rate limits also cap requests and tokens per minute. Administrators also control which models each organization, workspace, project, and user can access.

Workspaces give each team and business unit its own policies, models, budgets, and logs with central governance. Roles and permissions define what each person can see and change. Virtual keys tie each request to a user, project, and owner, providing one view of policy and cost.

Intelligent routing sends routine tasks to lower-cost models and reserves premium models for the work that needs them. Guardrails also block malicious prompts and Denial-of-Wallet (DoW) attacks before they consume tokens or GPU compute.

Enforce usage policy and keep AI data under your control

FortiAIGate is self-hosted on-premises or in your cloud, on Kubernetes with NVIDIA GPUs. The control, management, and data planes stay inside your environment, along with prompts, responses, and logs. Deployment options include bare metal, KVM, VMware ESXi, Amazon EC2, and Amazon EKS.

More than 50 predefined policies can be added with a single click, so you start from a governed baseline on day one. They span AI governance frameworks, data privacy laws, industry regulations, and security frameworks. Examples include the General Data Protection Regulation, the National Institute of Standards and Technology AI Risk Management Framework, and ISO standards. Custom rules adapt each policy to your business.

FortiAIGate logs each prompt, output, and tool call and forwards the logs to the platforms you already use, making it easy to provide evidence to auditors.

Inspect shadow AI across the organization

FortiAIGate works with FortiGate network firewalls to control shadow AI activity. FortiGate detects shadow AI activity at the network edge and sends that traffic to FortiAIGate. FortiAIGate inspects the full interaction in context, including the user, the AI service, and each prompt, response, and file. FortiAIGate can block the request, alert your team, or mask and anonymize sensitive data before it returns the traffic to FortiGate.

FortiAIGate extends the Fortinet AI-Native Security Fabric to the AI runtime environment. This enables the same visibility and control over AI applications that you have for users, devices, and traditional applications.

Scale AI without Losing Control

From one self-hosted control point, FortiAIGate stops AI attacks and data leaks, enforces usage policy, and caps AI spend. It inspects every step, including prompts, outputs, tool calls, and tool responses. To preserve data sovereignty, all prompts, sensitive data, and interaction logs remain strictly inside the host environment. In addition, FortiAIGate provides verifiable audit evidence for every interaction while keeping AI spend firmly within budget by user, team, project, and model.

Learn more about FortiAIGate

[1] Dan Tinkoff, Lieven Van der Veken, and Michael Chui, The state of AI in 2026: On the road to ROI, McKinsey & Company, August 25, 2026. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

[2] Cost of a Data Breach Report 2026, IBM, July 29, 2026. https://www.ibm.com/reports/data-breach

[3] 46.9% of Enterprises Report AI Spend Over Budget in 2H 2026, Futurum, September 10, 2026. https://futurumgroup.com/press-release/46-9-of-enterprises-report-ai-spend-over-budget-in-2h-2026/

Copyright © 2026 Fortinet, Inc. All rights reserved. Fortinet®, FortiGate®, FortiCare® and FortiGuard®, and certain other marks are registered trademarks of Fortinet, Inc., and other Fortinet names herein may also be registered and/or common law trademarks of Fortinet. All other product or company names may be trademarks of their respective owners. Performance and other metrics contained herein were attained in internal lab tests under ideal conditions, and actual performance and other results may vary. Network variables, different network environments and other conditions may affect performance results. Nothing herein represents any binding commitment by Fortinet, and Fortinet disclaims all warranties, whether express or implied, except to the extent Fortinet enters a binding written contract, signed by Fortinet's SVP Legal and above, with a purchaser that expressly warrants that the identified product will perform according to certain expressly-identified performance metrics and, in such event, only the specific performance metrics expressly identified in such binding written contract shall be binding on Fortinet. For absolute clarity, any such warranty will be limited to performance in the same ideal conditions as in Fortinet's internal lab tests. Fortinet disclaims in full any covenants, representations, and guarantees pursuant hereto, whether express or implied. Fortinet reserves the right to change, modify, transfer, or otherwise revise this publication without notice, and the most current version of the publication shall be applicable.

October 23, 2026 1:42 PM / MKTG-4453-0-0-EN