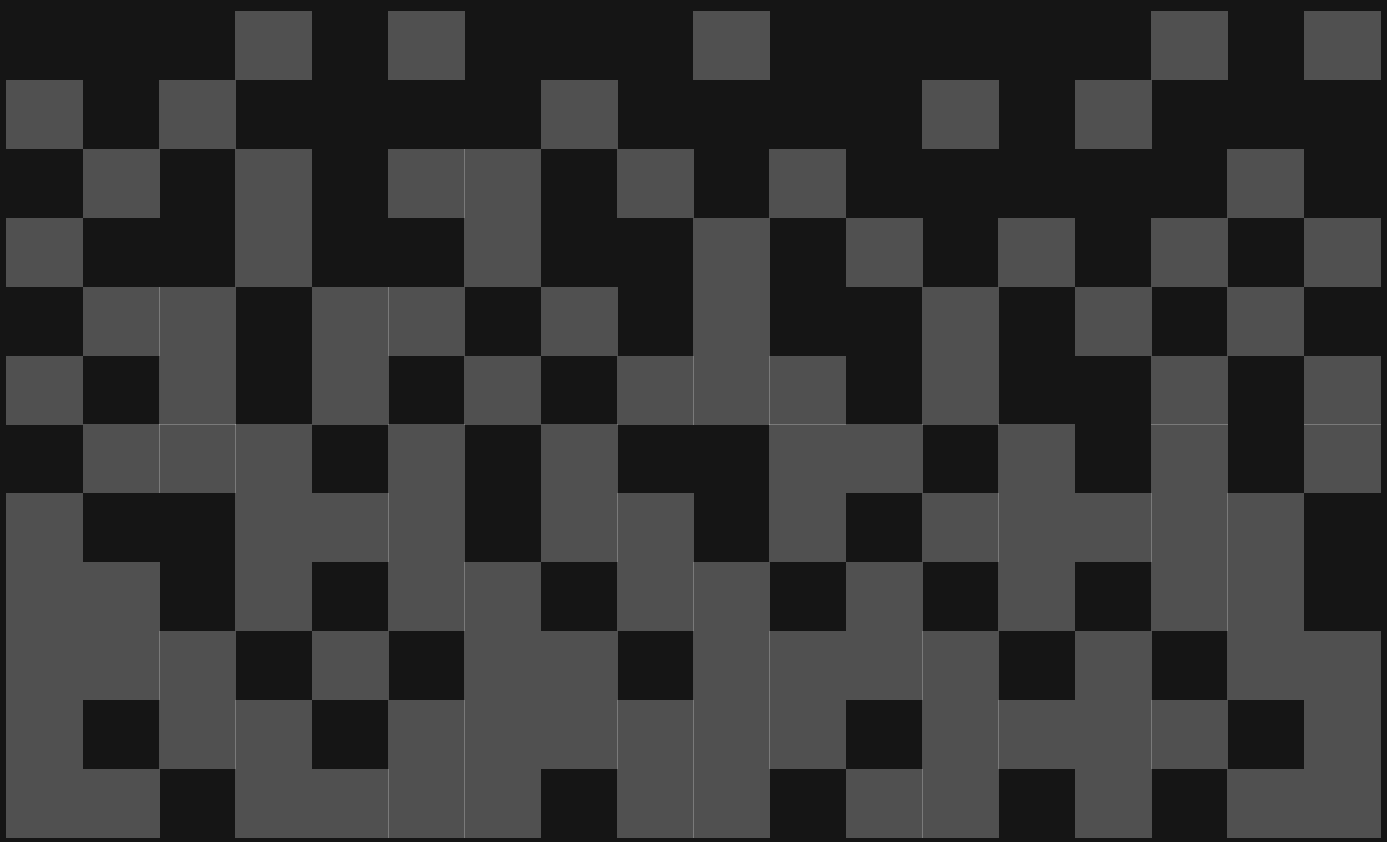
SOLUTION BRIEF

Aim Security's AI Firewall

Runtime protection for AI Applications You Build



THE CHALLENGE

Automate protection and ensure governance for AI apps in production

As organizations increasingly invest in building and deploying custom applications using AI technologies like LLMs and agents, they both contend with a new attack surface potentially vulnerable to exploits like prompt injections - as well as run into compliance and regulatory concerns about how agents and users interact or expose sensitive and proprietary data.

To consistently realize the business value of these AI investments, security teams need a scalable and seamless platform that functions in real time to:

- Detect and prevent prompt-level attacks
- Maintain prompt and agent activity governance based on corporate policy and requirements
- Monitor and report on activity for safe and appropriate use - as well as usage trends and patterns
- Implement guardrails in collaboration with application owners and compliance teams
- Ensure that sensitive data is protected or anonymized for input and output

In addition security teams often can’t keep pace with new risk and governance demands. To support these evolving demands, security teams need purpose built technology that provides centralized policy definition, with distributed enforcement that is easy to deploy as new applications are brought on line.

INTRODUCING

The Aim Firewall: Flexible, High-performance and Seamless Enforcement

The Aim Firewall provides detection and response to AI attacks, prompt level inspection, policy enforcement for data protection and agent interaction, as well as monitoring as an integral component of custom AI app deployment.

The Firewall can be deployed where and how it makes sense for the application owner and the security team. The Firewall provides out of the box guardrails for key security frameworks, as well as supports the ability to define custom prompt-level policies based on data classifications and semantic analysis of domain and intent for both input and output.

Powered by the Aim Engine™, the Firewall automatically detects and protects against threats and performs inspection of prompts, LLM and chatbot responses, and agentic AI interactions, leveraging cutting edge research from Aim Labs on AI attacks.

The Aim Engine is maintained by a dedicated data science team that is focused on AI threats and risks, supporting a market leading AI intelligence repository. As our team identifies new threats and attack vectors, we create and deploy detections and protection policies. Aim also supports centralized data collection and reporting across all deployed firewalls, providing security, compliance and governance stakeholders with a consolidated view across all AI applications.

Key Benefits

High-performance, low false positive attack detection

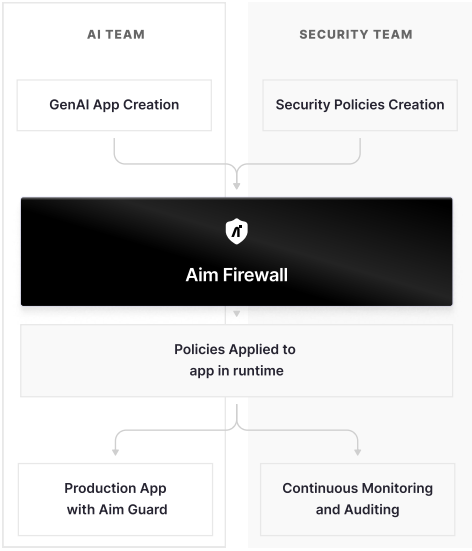
Designed to meet high-performance, low latency requirements, the Aim Engine provides seamless inspection and enforcement actions across deployed Aim firewalls - ensuring there is no disruption to app functionality and utilization

Invisible to Your Developers

Easily integrate the Aim firewall into your software stack with multiple deployment methodologies, giving security teams the autonomy to protect applications with minimal developer disruption, and securely deploy AI agents into your production environment.

Organizational Workflows and Collaborative Policy Definition

A cohesive and effective strategy requires collaboration with AI enablement teams, legal, data governance, and compliance teams. Aim allows security teams to delegate policy definition to application owners, while still maintaining a set of common predefined guardrails.



Data Protection and Anonymization

Aim provides a set of flexible policy options that allow users to interact with LLMs, while still ensuring data privacy and protection in compliance with EU GDPR and other regulations - including block, monitor and anonymize based on data classification and inferred intent.

Enforce Content Guardrails

Go beyond regex with semantic analysis to detect unsafe content in user prompts or LLM responses, including profanity, hate speech, sexual content, criminal topics, and other prohibited topics.

Flexible deployment options

- AI-GW Integration – Aim natively integrates with the AI gateway to provide inline guardrails
- Out-of-band API – Aim provides an API that can be called out-of-band by any application
- Aim AI Gateway – Aim provides an OpenAI-API endpoint that wraps a model

Aim is on a mission to empower secure AI use across the enterprise. The Aim AI Security Platform ensures the security of every AI interaction, agent, and application, safeguarding against adversarial attacks, data leaks, and compliance breaches. Get comfortable and unleash the business benefits of AI without compromising on security.

Learn more at **aim.security →**