Prisma AIRS®
AI Runtime Security

Defend Your AI Ecosystem Against a New Wave of Threats

Enterprise AI applications face an evolving threat landscape. Organizations are rapidly adopting AI into their applications to help with tasks such as customer service and increasing employee productivity. However, these AI applications face AI-specific threats that traditional security solutions can't help protect.

Integrating AI into your applications is more than a simple plug-and-play process of adding an AI model. For AI applications to deliver the most precise and valuable responses, they need to be built as "compound systems." This requires implementing an entire AI stack, which inherently involves managing and safeguarding access to sensitive internal data. That means seamlessly integrating multiple AI models, plugins, vector databases, and even internet search functionalities. Each new element in your AI application introduces potential vulnerabilities for attackers to exploit during runtime operations.

To be prepared to defend their AI environments, organizations like yours need an enterprise AI security solution that enables you to:

- **Discover** your AI ecosystem (AI applications, models, users, and datasets) automatically.
- **Protect** your models, data, and apps against AI-specific and foundational attacks.
- **Monitor** the runtime risk exposure of your AI ecosystem continuously.

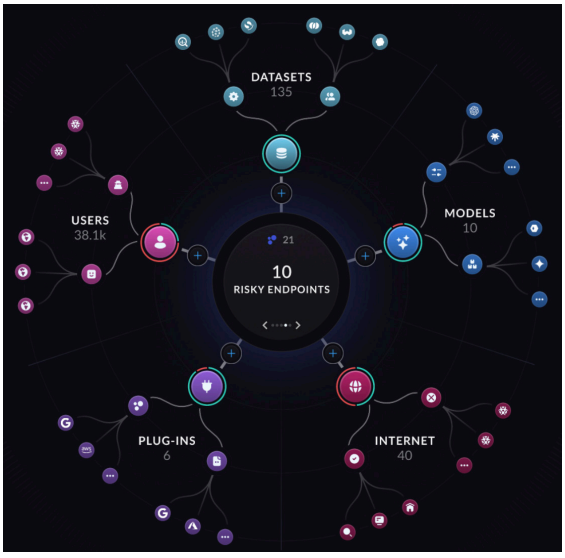
Key Benefits	
Network Intercept	**API Intercept**
• Application-layer decoding and segmentation for thousands of apps and protocols. • 40+ models protected across Google Cloud, AWS, and Azure, as well as traffic making calls directly to the OpenAI API. • Secure network, foundational, and AI-specific threats with one solution. • Protect containerized and virtualized workloads with built-in east-west traffic inspection. • Consistent best-in-class protection everywhere with Cloud-Delivered Security Services (CDSS): › 40% better protection from web-based attacks, including 25+ DNS attack types. › AI-powered prevention of >90% of zero-day application command and SQL injection attacks. • Best-in-class data protection: › 2X greater coverage than other cloud-based data leakage prevention solutions. › 1,000+ predefined data patterns. › 99% malware detection accuracy, with 26% more detections than traditional sandboxes.	• Fully agnostic deployment—secure any public or private model. • Protect AI apps, models, and data without decryption overhead. • Secure AI agents in low-code/no-code environments. • Agentic Threats • Granular detection and protection for every AI application: › Return custom error responses to users based on detected threats.

Product Capabilities

Discover and Understand the AI Ecosystem

With this key component of the Prisma AIRS platform you'll be able to understand how your cloud environments across AWS, Microsoft Azure, and GCP are using AI applications. With an intuitive onboarding process, users can quickly create Terraform templates to analyze data flows to visualize the asset inventory and understand how they communicate with each other. This allows for detailed mapping of AI applications to the relevant models, users, external sites, plugins, and data sources, revealing complex and sometimes unknown interconnections.

By visualizing and comprehending these relationships, users gain deep insights into their AI infrastructure, enabling informed decisions for placing AI Runtime Security instances and proactive management to boost efficiency, security, and compliance in their cloud-based AI implementations.



Protect Against AI-Specific Threats

Application Protection

Protect your AI applications by leveraging our state-of-the-art CDSS to ensure robust defense against malicious URLs, with Advanced URL Filtering at the core of our capabilities. Scan and detect URLs going between your AI applications and models to ensure that you can block (or flag) your applications from displaying or fetching malicious URLs. This prevents the applications or end users from receiving malicious URLs, which may appear in model output due to poisoned retrieval augmented generation (RAG) or training datasets.

Furthermore, URL security for AI apps prevents data exfiltration attacks that trick an AI model to compile a URL containing an attacker-owned domain with sensitive data embedded in URL parameters, which the app or end user may attempt to fetch, sending the data to the attacker's server. You can also configure policies to allow, alert, or block specific URL domains that appear in model input or output, which enables enforcement of fine-grained RAG web access control.

AI Runtime Security also enables detailed segmentation of all your application components within an environment to secure every communication pathway, from port-to-port to namespace-to-namespace traffic, effectively preventing both known and zero-day application-layer attacks.

Figure 1. Dashboard showing current state of protection and alerted threats

Model Protection

Defend your AI applications against threats such as direct and indirect prompt injections. To maintain the integrity of these systems, it's critical to defend against various types of prompt injection attacks, which extend beyond simple impersonation. AI Runtime Security can block prompt injections such as goal hijacking and do-anything-now attacks.

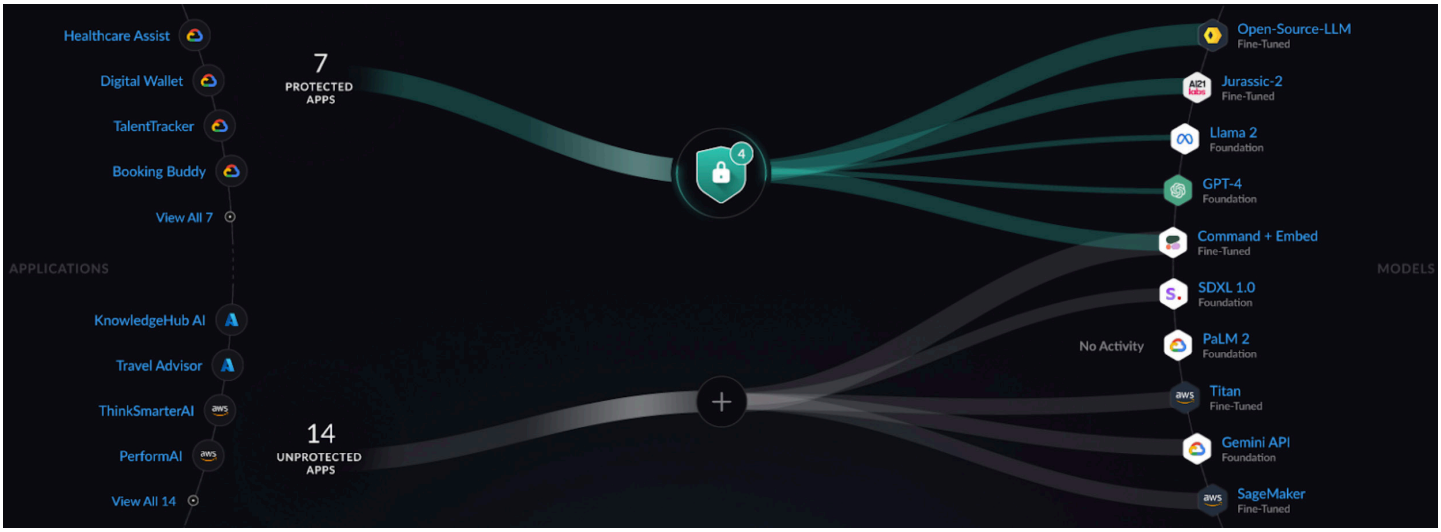


Figure 2. Model view showing the placement of AI Runtime Security instances and connections between applications and models

Data Protection

Stop data leakage to and from your AI applications with the built-in Enterprise Data Loss Prevention (DLP) CDSS. As you work to deploy and train finely tuned models in your environment, you want to ensure that the training data isn't susceptible to being leaked in application outputs. The data protection features built into AI Runtime Security can detect over 1,000 predefined data patterns (regex and ML-based), support custom data patterns in prompts and responses, and has double the coverage compared to other cloud-based data leakage prevention solutions.

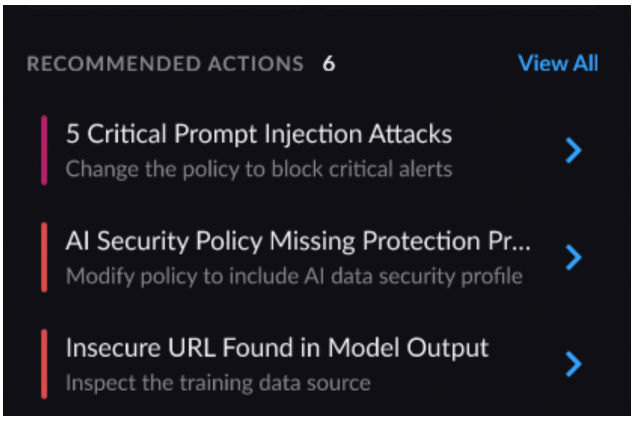
For AI applications that use AI models to generate database queries, prevent unauthorized changes to your databases by regulating what types of query scripts (Create, Read, Update, Delete) can be returned from a model (currently SQL supported).

Agentic Threat Protection

As enterprises increasingly use AI agents (including those built on no-code/low-code platforms), securing these agents themselves becomes paramount. Prisma AIRS provides AI Agent Security to defend against new agentic threats, such as identity impersonation, and memory manipulation. Plus, it protects against tool misuse by ensuring that tools and APIs connected to AI agents are not abused.

Monitor Your Runtime Risk

AI Runtime Security is designed to safeguard your AI environments with constant analysis of your AI runtime risk posture, offering clear insights into vulnerabilities within operational AI systems. It continuously evaluates unprotected AI applications across your organization, identifying any that lack essential security measures. To ensure comprehensive protection, the software pinpoints risky communication pathways emerging from AI applications, identifying potential entry points for malicious activity. This proactive approach allows organizations to secure their AI applications, reduce risk exposure, and maintain robust defenses in an evolving cybersecurity landscape.



Flexible Deployment Options

Deploy AI Runtime Security to protect the data and traffic flows between AI applications and their associated models, users, external sites, plugins, agents, and data sources. Palo Alto Networks enables AI security that aligns with your infrastructure and application needs. Secure your applications through a centralized network intercept (for AWS, Azure, and GCP environments) and/or through granular protection via our API intercept (for all apps, environments, and models). Organizations can choose either one or both options but only need to define policies once and apply as and where needed.

Automated or Manual Network Deployments

Strata™ Cloud Manager orchestrates automated network-based deployments into Google Cloud, AWS, and Azure environments by providing Terraform templates with recommended best practices architecture scripts for the target cloud environment. For manual deployments, select your image and bootstrapping variables, and manually deploy AI Runtime Security instances into public clouds. It provides seamless security for virtualized and container environments, ensuring comprehensive traffic protection—including east-west, outbound, and inbound protection within the Kubernetes clusters and app-to-app, app-to-model, and app-to-database interactions. Your AI Runtime Security instances not only shield against north-south attacks but also offer port-to-port isolation—crucial for maintaining containerized and noncontainerized application integrity.

For additional information, please see the deployment guides for AWS, Microsoft Azure, and Google Cloud.

API-Based Deployments for Developers and AI Agents

Developers can use software development practices to automate the management and deployment of AI security in AI applications and AI agents. They gain a quick and simple way to embed AI security into applications by using code to protect against AI-specific attacks. This API-based functionality enables developers to protect AI applications and agents in real time from known and unknown threats specific to AI, while having granular protection including custom error behaviors and user-specific policies. Alternatively, developers (and detection engineers) gather data asynchronously for analysis. Additionally, the API allows developers to scan RAG data sources for poisoning or PII.

Palo Alto Networks enables you to secure any AI app, agent, workload, and model in any environment. Deploy using a network-based (already launched) and/or code-based (new) intercept to align with the needs of your application and infrastructure.

Flexible Consumption Model

Customers can acquire or use their existing Flexible Software NGFW (FW-Flex) credits to deploy AI Runtime Security instances from Strata Cloud Manager and reconfigure security quickly with minimal procurement roadblocks. For API-based instances, developers can create a deployment profile in their support portal via FW-Flex credits, which allows them to create an API key in Strata Cloud Manager. They can then use this key to make API calls from application code.

Product Specifications

For cloud-specific AI model support on network-based intercepts, please check the Support Table.

For capacity and throughput information, please see the VM-Series datasheet.

Resources

- Prisma AIRS product page
- Software NGFW Credit Estimator
- Software Firewall Selector
- VM-Series datasheet

About This Datasheet

The information provided with this paper that concerns technical or professional subject matter is for general awareness only, may be subject to change, and does not constitute legal or professional advice, nor warranty of fitness for a particular purpose or compliance with applicable laws.

paloalto NETWORKS

3000 Tannery Way
Santa Clara, CA 95054

Main: +1.408.753.4000
Sales: +1.866.320.4788
Support: +1.866.898.9087

www.paloaltonetworks.com

prisma_ds_prisma-airs_110525