Prisma AIRS AI Model Security

Secure Every Model Across the AI Lifecycle from Development to Deployment

AI innovation is transforming every industry, but beneath that progress lies a largely invisible attack surface—the models themselves. AI models can contain embedded malicious code, use unsafe serialization formats, or contain unverified dependencies that execute silently at load or inference time. Traditional security tools can't see inside these AI models—"black boxes" that leave organizations blind to threats hidden within the systems driving their AI strategy.

Prisma® AIRS™ AI Model Security brings visibility, validation, and control to every model across its lifecycle. It detects, prevents, and contains model-borne threats before they reach production so you can innovate confidently, protect your intellectual property (IP), and maintain regulatory trust.

Key Benefits

Prisma AIRS AI Model Security closes these gaps with a unified, automated approach that strengthens security, accelerates development, and preserves the confidentiality of your AI assets.

- **Prevent threats before deployment:** Identify malicious code, poisoned data, and unsafe formats before models reach production.
- **Enforce consistent standards:** Apply organization-wide policies to every model, regardless of origin.
- **Accelerate secure AI adoption:** Automate validation workflows to reduce manual review time and speed release cycles.
- **Preserve IP and confidentiality:** Scan models locally so sensitive data never leaves your control.
- **Demonstrate governance and compliance:** Produce audit-ready evidence for regulators, partners, and internal oversight.

What Is an AI Model?

A model is the collection of files required to perform a single inference pass, including weights, configurations, and dependencies. Each model includes:

- **Source:** Where it resides, for example: Hugging Face, Amazon S3, or local storage.
- **Version:** The specific version of a logical entity.
- **Files:** Artifacts such as binaries, weights, and metadata.

Because these components can execute code, Prisma AIRS AI Model Security treats models as critical software assets that scan and validate every file, dependency, and serialization format before use.

Key Challenges: Critical Security Gaps in the AI Model Lifecycle

AI models introduce unique, often invisible risks that traditional security tools are not designed to detect or control, creating critical gaps across development and deployment.

Black Box: You Can't See Inside the AI

AI models are complex and opaque. Without deep visibility, their hidden vulnerabilities, malicious code, or embedded threats go undetected, leaving teams unsure about what's safe to deploy.

Supply Chain Exposure: External Models Introduce Hidden Risks

Although open-source and third-party models accelerate innovation, they can import poisoned data, embedded malware, or compromised components. Each unverified dependency expands systemic risk to your business, customers, and brand.

IP Exposure: Your Competitive Advantage Is at Risk

Your proprietary models and training data are your edge. Moving them outside secure environments for scanning increases IP leakage and compliance risk, forcing a trade-off between safety and confidentiality.

Operational Friction: Security Slows Down AI Development

Security reviews are manual, slow, and disconnected from fast AI release cycles. This creates bottlenecks, inconsistent coverage, and friction between security, data science, and product teams.

Prisma AIRS AI Model Security Capabilities

Prisma AIRS AI Model Security delivers a comprehensive set of capabilities that secure every model, from intake to deployment, combining deep analysis, adaptive enforcement, continuous monitoring, and seamless integration into existing AI workflows.

Comprehensive Model Scanning

Prisma AIRS AI Model Security performs deep static and dynamic analysis of model artifacts to detect:

- **Deserialization threats:** Unsafe formats, such as Pickle, that enable arbitrary code execution.
- **Malicious payloads and backdoors:** Hidden logic that manipulates inference or exfiltrates data.
- **Insecure dependencies and libraries:** Untrusted or outdated components that create exploit paths.
- **License violations and metadata gaps:** Ensures proper attribution and legal compliance.

For example, a poisoned fraud-detection model could alter decision thresholds or silently bypass checks. Prisma AIRS AI Model Security detects these anomalies through architecture and serialization inspection.

Policy Enforcement and Adaptive Controls

Prisma AIRS enforces context-aware policies through AI Model Security groups that align to organizational risk, ensuring that every model is evaluated according to its source and environment.

Prisma AIRS AI Model Security tailors enforcement rules based on where a model comes from and how it will be used:

- **External models (Hugging Face or partner repositories):** Strict blocking rules for malware, license validation, and author verification.
- **Internal models (Amazon S3 or local storage):** Flexible controls focused on provenance, dependency validation, and version integrity.
- **Hybrid environments:** Dynamic rules adjust to source metadata and deployment context.

Rules can be blocking (prevent deployment) or nonblocking (generate alerts), supporting risk-based decisions without slowing innovation.

Continuous Monitoring and Compliance

Once models are approved and deployed, Prisma AIRS AI Model Security continuously monitors their security posture through:

- **Real-time dashboards** that display organizational risk trends.
- **Threat-intelligence feeds** that detect new attack patterns and alert when models become vulnerable.
- **Audit and evidence logging** that captures every scan verdict, rule change, and policy decision for governance and regulatory reporting.

Compliance teams can export evidence to support ISO, SOC 2, and industry frameworks.

Seamless Integration Across Workflows

Integrate Prisma AIRS AI Model Security directly into existing toolchains, including:

- **CI/CD pipelines:** GitHub Actions, GitLab CI, Jenkins, and Azure DevOps.
- **Model registries:** MLflow, Kubeflow, Vertex AI, and custom registries.
- **Development environments:** Jupyter, VS Code, and PyCharm.
- **Deployment automation tools:** Terraform, Ansible, and Kubernetes.
- **SIEM platforms:** Cortex XSIAM® and third-party aggregators.

Models never leave your environment; they preserve IP and data sovereignty while maintaining speed.

How Prisma AIRS AI Model Security Protects Models

Prisma AIRS AI Model Security applies the right security controls based on where models originate and how they are stored, ensuring each model is validated with the level of scrutiny its risk profile demands.

External Models

Downloaded from repositories, like Hugging Face, or partner vendors, these models are scanned for malware, unsafe formats, and license violations. Strict blocking rules and provenance checks ensure that only verified models enter your pipeline.

Internal Models

Stored in enterprise Amazon S3 buckets or local object storage, these models undergo integrity and dependency validation without leaving secure infrastructure. Local scanning eliminates the need to export confidential artifacts.

Deployment Flexibility

Prisma AIRS AI Model Security adapts to your existing infrastructure by deploying wherever your models reside, ensuring consistent protection without disrupting workflows.

- **Local scanning:** Runs scans entirely on-premises to protect IP-sensitive models.
- **CI/CD pipeline integration:** Performs automated testing and model validation at build time.
- **High-performance API:** Delivers fast, reliable model assessments in milliseconds. It's backed by a trusted database of over 1.9 million public AI models and continuously scans new releases from major repositories like Hugging Face. And, it helps security teams identify risks and make deployment decisions in real time.

Benefits by Role: Value for Every Team Driving AI Innovation

Prisma AIRS AI Model Security delivers tailored benefits across the organization—aligning security, engineering, and data science teams on a shared foundation of visibility, trust, and responsible AI development.

CISOs and Security Leaders

Prisma AIRS AI Model Security provides a governed, auditable, and enterprise-scale approach to model security. CISOs can define organization-wide standards, set risk tolerance across development and production, and ensure consistent enforcement at every checkpoint. All validation results, approvals, and policy decisions feed directly into compliance and audit programs, supporting regulatory readiness and executive oversight.

Security Engineers and Developers

For engineering and security teams, Prisma AIRS AI Model Security integrates directly into existing automation. Scanning runs at CI/CD, registry, and deployment stages, with automated gates that block vulnerable models before they move downstream. Continuous monitoring, real-time violation alerts, and SIEM integration centralize visibility and streamline triage across distributed pipelines.

Data Science and AI Teams

Prisma AIRS AI Model Security enables data scientists to innovate quickly without compromising safety. AI Teams can validate open-source and third-party models without moving sensitive data or interrupting experimentation. Shared dashboards and automated alerts create alignment between data science and security, supporting faster prototyping, safer model reuse, and smoother transitions into production.

Flexible Consumption Model: Security That Fits Your Workflow

Prisma AIRS AI Model Security is delivered through Strata™ Cloud Manager and offers flexible activation options so teams can embed model security directly into existing development and MLOps processes. Teams can:

- Activate scanning using **APIs or SDKs** for programmatic workflows.
- Provision **developer API keys** directly from Strata Cloud Manager.
- Integrate security checks without procurement delays or tooling changes.

This consumption model ensures security can be adopted immediately, scales with demand, and aligns with how modern AI teams build and deploy models.

Product Specifications: Built for Enterprise Scale and Performance

Prisma AIRS AI Model Security provides the speed, compatibility, and reliability required for high-scale AI environments. It includes:

- **Support for over 35 model formats**, including Hugging Face Transformers, Keras, ONNX, PyTorch, Scikit-Learn, and TensorFlow.
- **API response times 200 ms or less** with a **99.9% uptime SLA**.
- **Local and cloud deployment options** for complete data control.
- **Integration templates** for CI/CD and model registry platforms.
- **Exportable reports**, **audit logs**, **and evidence packages** to support governance and compliance.

Why Prisma AIRS AI Model Security: The Most Complete Approach to Securing AI Models

Prisma AIRS AI Model Security delivers unmatched depth, trust, and operational scale, providing organizations with the end-to-end assurances required to safely adopt AI.

- **Depth:** Model-aware inspection goes beyond signatures to understand serialization, dependencies, and embedded code.
- **Trust:** Local scanning ensures proprietary models and data stay within secure environments.
- **Speed:** Automated controls embed security directly into development workflows without slowing delivery.
- **Scale:** Unified policy management and visibility across models, agents, and runtime—delivered through one platform.

Additional Resources

- Prisma AIRS product page
- Prisma AIRS AI Model Security in TechDocs

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-ai-model-security_112425