Cisco Al Defense is an end-to-end solution to secure enterprise use of Al models and applications. The product protects against novel safety and security threats that are introduced through the development and deployment of Al applications, as well as employee use of third-party shadow Al apps.

Benefits:

Al Access: Find third-party and shadow Al apps being used by employees across your organization and protect against sensitive data loss, threats, and safety risks. Extend secure service edge (SSE) capabilities for targeted defense and consistent policies specific to GenAl apps, like malicious user intent or harmful prompts and responses. Al Access is included as part of the Cisco Secure Access Advantage package.

Al Cloud Visibility: Discover enterprise Al assets running in your organization's public and virtual cloud environments, including unsanctioned models. Customers with this feature will be provisioned a <u>Cisco Multicloud Defense</u> tenant, as included with the relevant offer.

Al Model and Application Validation: Perform automated vulnerability assessment of Al models and applications to identify hundreds of potential safety and security risks.

Al Runtime Protection: Safeguard production Al applications against adversarial attacks, sensitive data leaks, and harmful responses in real-time. Customers with this feature will receive Cisco Multicloud Defense gateway hours to support enforcement capabilities, as included with the relevant offer.

FAQs:

- Q: What is Cisco AI Defense?
- **A:** An AI security solution that addresses the risks introduced by the development, deployment, and usage of AI. AI Defense solves for three key areas in AI security:
- **Discovery** of Al workloads, applications, models, data, and user access across distributed cloud environments
- Detection of model and application security vulnerabilities and adversarial attacks that put AI systems at risk
- Protection for runtime AI applications against rapidly-evolving threats, including prompt injections, denial of service, and sensitive data leakage
- Q: How is AI Defense deployed?
- **A:** Deployment options include SaaS or in a private VPC (hybrid), with a cloud-based control plane and self-hosted data plane to keep models and data within the customer's network (hybrid deployment available after GA).
- **Q:** What are some examples of the applications that are secured by the AI Access feature of AI Defense?
- A: ChatGPT, Writer, CoPilot, Notion AI, etc.

Packages: Product:	Feature:	Validation Essentials	Runtime Essentials	Advantage
AI Defense	Al Cloud Visibility	✓	✓	✓
	AI Validation	✓		✓
	AI Runtime		✓	✓
	Splunk Technical Add-on ¹		✓	✓
Secure Access	Al Access	Al Access is included in the Cisco Secure Access Advantage Tier.		

¹ Splunk Enterprise Security subscription required (not included with Al Defense)

Features:

Continuous discovery of shadow AI: Discover, report, and manage 750+ third-party generative AI apps, across both sanctioned and unsanctioned, shadow AI categories. AI Access tracks the prompt and request formats of each AI for deep inspection of traffic for security policy enforcement.

Risk assessment for GenAl apps: Gain visibility into the details around app usage, including activity, device, location etc. A dashboard with risk scores and information about each GenAl app informs your security program and policies.

Access controls: Control access to apps based on user identity and context like device, network status, and more.

Consistent protection for sensitive data loss, threats, and safety: Enforce consistent guardrails for protection across all categories of security, privacy, safety, and relevancy. Use Cisco's proprietary ML models for detections.

Customizable user security workflows: Guide employees with policies that use a variety of workflows, including completely blocking access, coaching a user with custom messages and redirect to sanctioned app, or fine-grained protection that monitors prompts and responses restricts only when a detected violation is triggered.



Resources:

Example SKU: AIDEF-SEC-SUB