🗑 PROTECT AI

The Platform for Al and ML Security

Protect AI delivers the broadest and most comprehensive platform to secure your AI. It enables you to see, know, and manage security risks to defend against unique AI security threats, and embrace MLSecOps for a safer AI-powered world.

Guardian	Guardian > Managed policies		
OVERVIEW			
	Managed policies		
SECURITY	3 1		
- 41			
SCANNER - FIRST PARTY MODELS	• •	0	0
57			
200	246		
GATEWAY - THIRD PARTY MODELS	240		
57			
ER (1997) 188	~ · · · · · · · · · · · · · · · · · · ·		~ ~
8 🛑			
	Global threshold		
	First party models	III Third party mo	dels
	This value represents global threshold for private scanning, click on the button	This value represents global threshold for public	
	to edit the default value. Edit global threshold 🖌	to edit the default value.	
	Edit global threshold 2	Edit global threshold 2	

Addressing Unique AI Security Risks

닢



TECHNOLOGY The AI supply chain is complex, introducing unique technical risks and vulnerabilities that put your business at risk. OPERATIONS

Without adequate oversight and controls, AI systems might behave unpredictably, inaccurately 'or' simply fail.



шц

🗊 PF

Home Home Radar Layer Signtine Recon Signtine Policies

REGULATION Organizations must evaluate how AI and ML fit into regulatory structures and

take steps to stay compliant.

REPUTATION

As enterprises turn to Al for their customer-facing workflows, failures can lead to a loss of brand trust.

The Protect AI Platform

Protect AI is the only vendor neutral, cloud-agnostic MLSecOps platform that helps you to See, Know, and Manage AI Risk end-to-end.

	oles & Group Controls →← y & Risk Policies 🛛 🕀	SIEM / SOAR Connectors Global Dashboards	絞 Al/ML Vendor Neutral
J GUARDIAN	RECON	L LAYER	radar
First & Third Party Model Security Enterprise Al Controls	Automated and Human Augmented GenAl Red Teaming Comprehensive	LLM Security LLM Monitoring & Observability	AI Bill of Materials (AI-BOM) AI Risk Assessment
CI/CD Integration	Threat Coverage Across 6 Major Risks No-code Integration	30+ Unique GenAl Scanners	Al Governance, Risk & Compliance

Continuously monitor, manage, and improve the security of your AI systems with the Protect AI platform for AI-SPM

Get Started with AI Asset Security

Implement Zero Trust for the engine that powers your AI — the Machine Learning Model.



GUARDIAN

Enforce the use of secure models through managed policies and a secure gateway
Continuously scan both first and third party models and artifacts for risks before they are deployed

Implement Comprehensive GenAl Security at Scale

Defend your LLMs from risk with end-to-end observability and monitoring

for risks and vulnerabilities



LAYER

- Evaluate LLM inputs and outputs for data leakage, adversarial threats, content moderation, and more
- · Gain full visibility into the upstream and downstream actions of the LLM
- Rigorously Red Team your LLMs with automated scanning

RECON

- Comprehensively evaluate threat exposure for all GenAl, including prompt injections, jailbreaks, data leakage and more
- Optimize LLM selection and guardrails for safety and security

· Go beyond firewalls to scale LLM usage securely

· Continuously improve the security posture of your GenAl system

Protect AI is the only platform to enable advanced and comprehensive AI-SPM capabilities

Gain Full Visibility, Risk Assessment, and Management of the AI Lifecycle

Ensure continuous assessment, management, and discovery of risk across AI entities



RADAR

- Inventory all AI entities and dependencies, including datasets, models, pipelines and applications, via an AI-BOM
- Enforce policies to ensure your AI remains compliant with AI regulation





Join the community



MLSECOPS

Learn best practices in MLSecOps, listen to podcasts with thought leaders, and connect with our thriving community.

HUNTR

Join the hunt in the world's first AI Bug Bounty Platform huntr is the place to start your journey into AI threat research.



OPEN SOURCE

Contribute code or issues, discover documentation, and get started with AI security with our Apache 2.0 licensed Open Source projects for AI Security.