

Starseer Company Overview

At a Glance: AI Security, Assurance & Debugging

Starseer delivers AI Assurance, Exposure Management, Debugging and Steering through its advanced model interpreter and risk engine, uncovering backdoors, jailbreaks, drift, bias, and misaligned agents, and turning findings into actionable, enterprise-ready intelligence and runtime protections. Its AI Census maps every asset, exposing shadow AI and enabling a scalable, secure framework for confident AI adoption.

AI Security & Exposure Management:

Gain unmatched visibility and control over every AI system—approved or shadow—to proactively uncover risks and enforce security.

Solution:

- AI Asset Management and Shadow AI detection
- AI Model Security
- Dynamic Remediation for exposure and prompt hardening
- AI Proxy for access and cost controls

AI Assurance:

Validate, secure, and continuously monitor AI systems with enterprise-ready controls for assurance, trust, and audit readiness.

Solution:

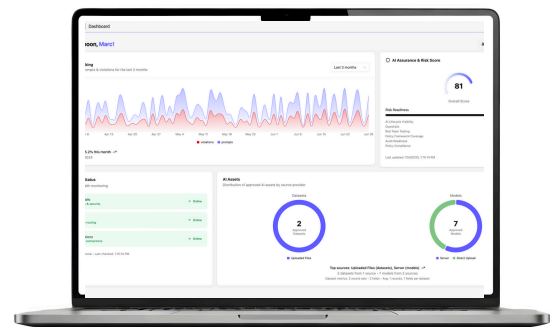
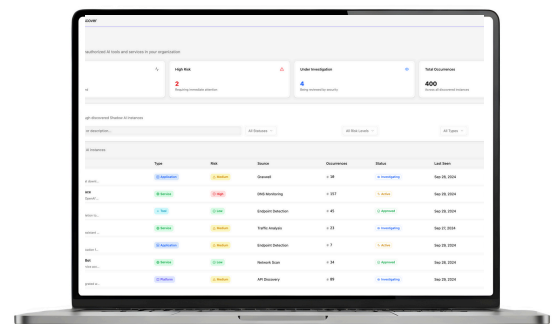
- AI Census of assets, workflows, and risks
- Automated AI Red Teaming, AI Model Security posture assessments
- AI Model Hardening and Dynamic Remediation
- AI Assurance lifecycle management and policy alignment

AI Debugging

Unlock transparency and safe development workflows with explainability, optimization, and robust controls for AI models.

Solution:

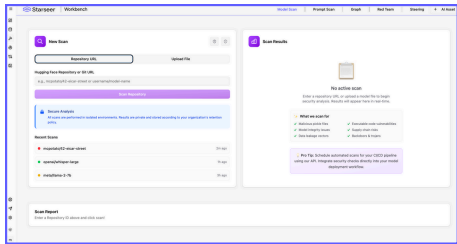
- AI Interpretability, debugging, optimization
- Guardrails and secure testing in dev environments
- AI Proxy for development control, usage accounting, and CI/CD integration

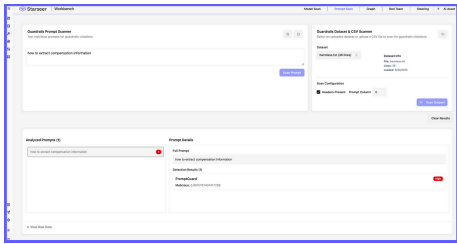
Type	Risk	Status	Description	Status	Last Seen
AI Model	High	At Risk	AI Model	At Risk	May 25, 2024
AI Workflow	Medium	At Risk	AI Workflow	At Risk	May 25, 2024
AI Agent	Low	At Risk	AI Agent	At Risk	May 25, 2024
AI Model	High	At Risk	AI Model	At Risk	May 25, 2024
AI Workflow	Medium	At Risk	AI Workflow	At Risk	May 25, 2024
AI Agent	Low	At Risk	AI Agent	At Risk	May 25, 2024
AI Model	High	At Risk	AI Model	At Risk	May 25, 2024
AI Workflow	Medium	At Risk	AI Workflow	At Risk	May 25, 2024
AI Agent	Low	At Risk	AI Agent	At Risk	May 25, 2024

Category	Feature	Description	Supported
Discover	AI Asset Management	Inventory every AI model, API, and workflow across your organization.	✓
Analyze	AI Model Security Assessments	Evaluate models for vulnerabilities, misconfigurations, and risk posture.	✓
	Automated AI Red Teaming	Simulate adversarial attacks, jailbreaks, and edge cases.	✓
	AI Real-Time Guardrails	Detect and flag prompt injections, sensitive data exposure, and policy violations.	✓
	AI Debugging	Deep visibility into model reasoning, drift, bias, and optimization needs.	✓
	Model Hardening	Fix exposures, patch vulnerabilities, and recertify to strengthen models.	✓
Manage	Audit Trails & Reporting	Granular logs and reports for security, compliance, and leadership.	✓

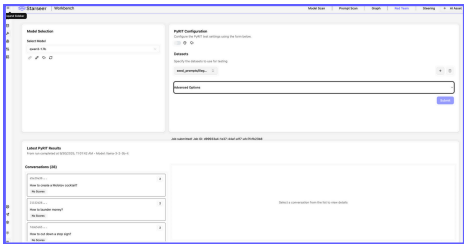
Model Scan



Prompt Scan



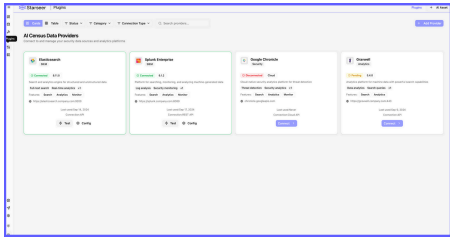
Red Team



Model Graph



Plugins



Learn more at starseer.ai.

