Zero Trust AI Security - Build to Runtime

AI-SPM | AI-DR | AI-RED TEAMING | AI-FIREWALL | AI-BOM* | AI-IDENTITIES* | AI-GRC*

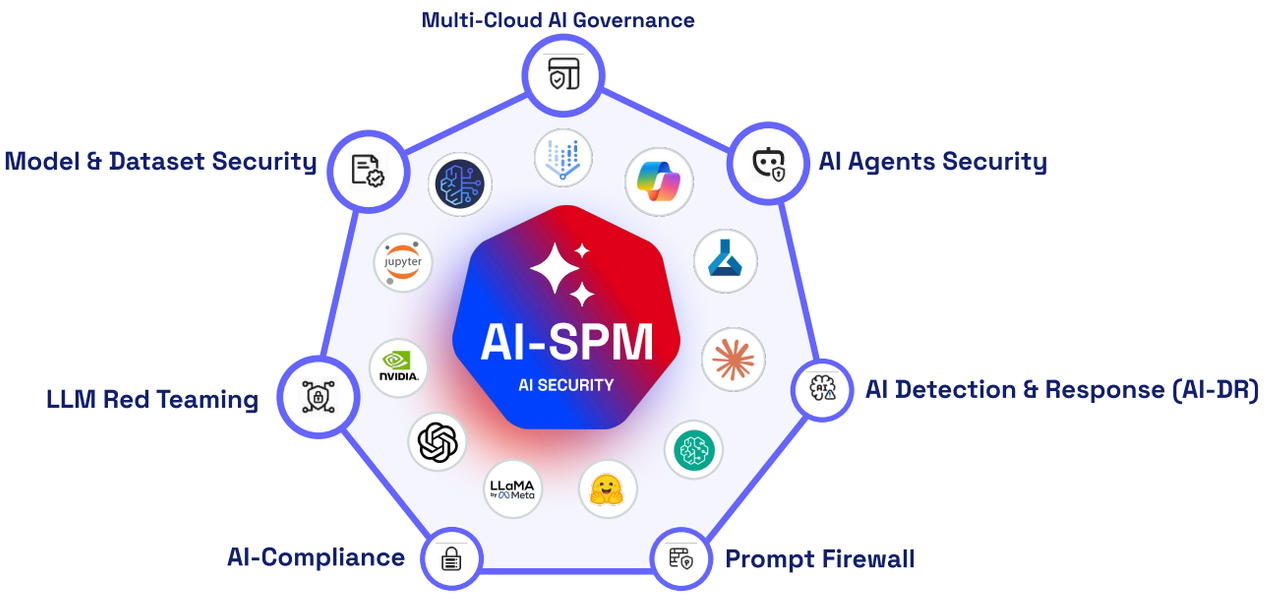
AccuKnox
Secure Code to Cognition™

Security Across Every Layer of AI

Traditional cloud security tools were not designed for AI workloads. AccuKnox closes the gap — from **shadow AI discovery** to kernel-level **agent sandboxing** — with runtime enforcement, not periodic scanning.

78% enterprises run shadow AI unnoticed

Aug '26 EU AI Act enforcement deadline



Platform Wide Support

AI Compliance
ISO27001, AVID, OWASP
+33 Frameworks

AI-DR
AI Assets Misconfiguration
Detection & Real-time Response

Data Security for AI

AI Copilot

Module	What It Does
AI Security Posture Management (AI-SPM)	Discover and govern all AI assets across your environment
AI Detection & Response (AI-DR)	Detect and respond to threats targeting AI workloads in runtime
AI Red Teaming	Automated adversarial testing for jailbreaks and abuse paths
Prompt Firewall	Block prompt injection and data leakage at the input layer
AI Pen Testing / CTEM	Continuous exposure testing across AI attack surfaces
AI-DSPM	Map sensitive data exposure paths in AI pipelines
AI-GRC	Automate compliance evidence generation for AI systems
AI Identity Controls	Enforce least privilege and audit entitlements for AI agents

Secure **AI Infrastructure**

Component	AccuKnox Coverage
vLLM	Process isolation · Syscall filtering
Ollama	ClawArmor sandbox · Network policy
NVIDIA TRITON INFERENCE SERVER	GPU namespace isolation · CVE tracking
LangChain MCP	Shadow discovery · Prompt interception

Achieve **AI Compliance**

Framework	Scope
MITRE ATLAS	Adversarial AI tactics and techniques mapping
NIST AI RMF	AI risk governance and management
ISO 42001	AI management system standard
EU Artificial Intelligence Act	High-risk AI system compliance

Secure **AI Providers / Apps**

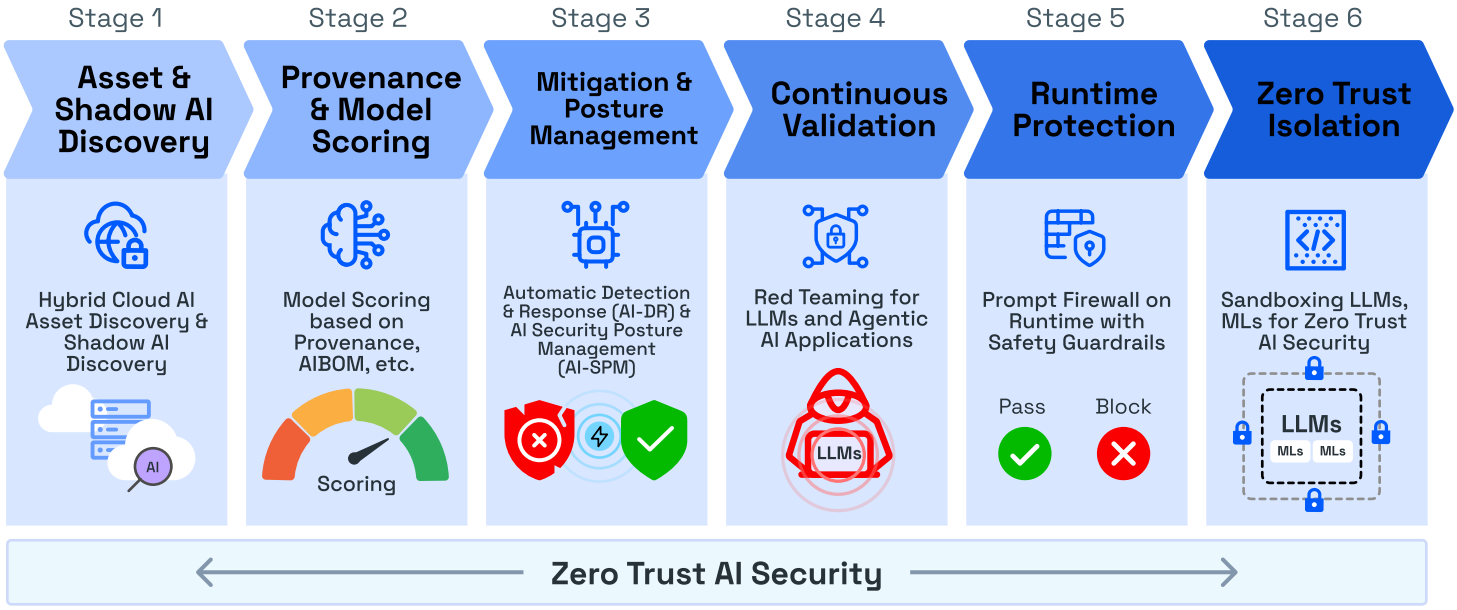
- Azure Copilot Studio
- AWS API Gateway
- Bedrock-Agentcore
- LiteLLM
- Power Apps
- Azure APIM
- Bifrost AI
- OpenAI Prompt Security

Zero Trust AI Security - Build to Runtime

AI-SPM | AI-DR | AI-RED TEAMING | AI-FIREWALL | AI-BOM* | AI-IDENTITIES* | AI-GRC*

ACCUKNOX
Secure Code to Cognition™

Multi-Stage AI Security: From Shadow Discovery to Zero Trust Isolation



Governance, Risk & Compliance

ISSUES	SOLUTIONS		
LAYER 4 **AGENTS** Prompt Injection & Malicious Prompt Abuse	**Agentic AI Security** - Multi cloud ai agent visibility, real time auditing & monitoring - Sandboxing unsafe use of tools - Sandboxing untrusted auto gen code	**Predictive ML Security** - Runtime visibility of model execution - Protecting untrusted model execution ML/LLM hijacking	
LAYER 3 **LLM/ML MODELS** Model Exploitation via Adversarial Inputs	**ML/LLM Prompt Firewall** - Real time prompt Visibility - Safety Guardrails against malicious prompts	**LLM Security** - Automated Red Teaming - Code Injection, hallucination Sentiment Analysis	**ML/LLM Dataset Scanning** - PII / PHI/Critical Data scanning
LAYER 2 **COMPUTE** Runtime Vulnerabilities & Malware Exposure	**Compute Security Posture** - Compute Issues on Containers, VMs - AKS running CICD Agents, IaC Agents or any managed/unmanaged Agents.		
LAYER 1 **INFRASTRUCTURE** Cloud Misconfiguration & Weak Posture	**AI Security Posture Management** - AI Infra Misconfiguration - AI/ML onboarding for cloud/onprem - Pipeline, Security graph view - Automapping deployed applications	**AIDR** - Cloud & AI event ingestion - AI-aware policy evaluation - AI misuse detection - Real-time alerts - Automated remediation	

Compliance Mapping Coverage

MITRE ATLAS
Adversarial tactics & techniques mapped to detection rules

OWASP LLM Top 10
All 10 categories covered with runtime guardrails

AVID (AI Vulnerability DB)
Vulnerability tracking integrated into AI-SPM

EU AI Act
High-risk system classification & audit readiness

NIST AI RMF
Govern, Map, Measure, Manage functions addressed

ISO 42001
AI management system controls with evidence generation

Zero Trust AI Security - Build to Runtime

AI-SPM | AI-DR | AI-RED TEAMING | AI-FIREWALL | AI-BOM* | AI-IDENTITIES* | AI-GRC*

AccuKnox
Secure Code to Cognition™

AI Security Features Coming Soon (Beta)

AI Pen Testing / CTEM — H2 2026

- Prompt injection, jailbreak, sandbox escape simulations
- Attack surface tracked with full audit trails

Unvalidated models ship with exploitable blind spots

AI-DSPM — H2 2026

- Finds PII in cloud data stores feeding AI pipelines
- GDPR, CCPA, HIPAA, and residency coverage

Regulated data in training sets - a violation before any breach

AI-GRC — H2 2026

- Posture mapped to NIST AI RMF, ISO 42001, EU AI Act
- Auto-generated audit evidence and compliance reports

No AI compliance posture means fines and lost deals

AI-Identities — H1 2026

- API keys, service accounts, inter-agent trust in one view
- Least-privilege enforcement for agentic workflows

Over-permissioned agents move laterally without detection

GPU Prompt Firewall — H2 2026

- Screens prompts at GPU scale before model inference
- Blocks injections from docs, emails, and APIs

Poisoned inputs reach your model unfiltered

Local AI Discovery — H1 2026

- Finds Ollama, LangChain agents, vector DBs outside IAM
- Inventory across dev machines and on-prem

78% of shadow AI runs with zero visibility or controls

AI-BOM — H2 2026

- SBOMs for models, frameworks, inference servers
- CVE detection and drift monitoring over time

Known CVEs in AI components already in production

Intelligent Red Teaming — H2 2026

- Continuous probing mapped to MITRE ATLAS at machine speed
- Dynamic probe generation across model providers

Memory poisoning corrupts agents between sessions

AI Agent Inventory — H2 2026

- All AI assets unified — cloud, on-prem, dev machines
- Flags endpoints in restricted or non-compliant regions

Unauthorized AI endpoints in restricted regions

Why AccuKnox AI Security Outperforms

ClawArmor Agent Sandbox

- Blocks unauthorized outbound calls from local LLMs at the kernel level before any packet leaves the host
- No user-space bypass possible — Ollama exfiltration via HTTP is denied at syscall

AI-DR: Threat Detection Built for AI Workloads

- Detects prompt anomalies, unusual token patterns, model inference spikes, and shadow API calls in real time
- Correlates AI-specific telemetry signals that standard cloud logs never see

Runtime Enforcement in Action

- PII-containing prompts are intercepted and rejected at the gateway — zero data reaches the LLM
- SSNs, credit card numbers, and regulated identifiers are logged and blocked before model inference

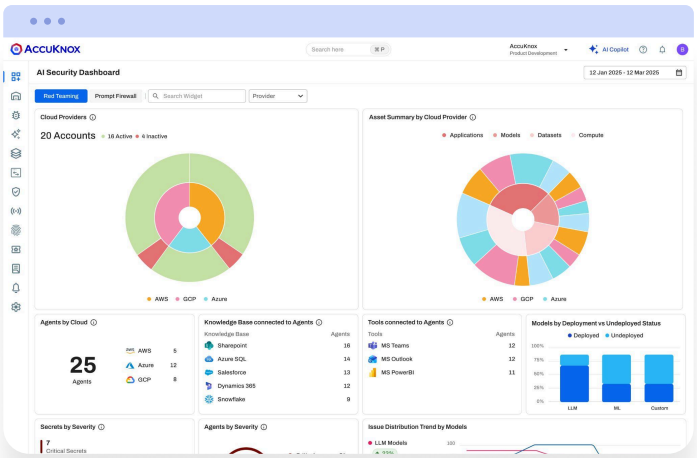
Prompt Firewall: Built for Production Scale

- Sub-5ms median latency at 10K+ requests per second across GPU inference endpoints
- Zero false-positive rate on known injection patterns — scales horizontally with no tuning required

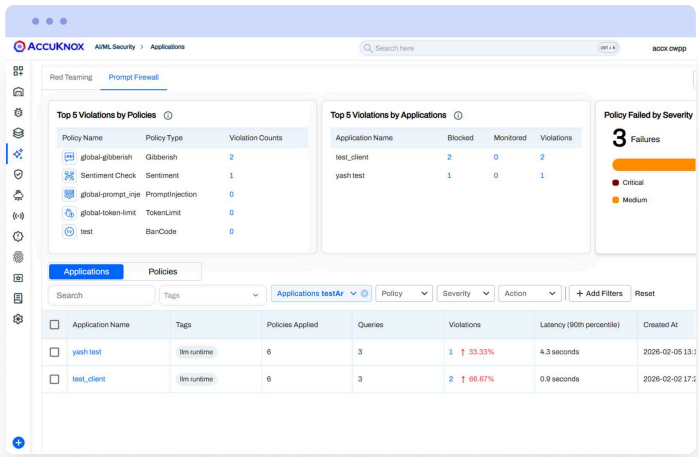
Zero Trust AI Security - Build to Runtime

AI-SPM | AI-DR | AI-RED TEAMING | AI-FIREWALL | AI-BOM* | AI-IDENTITIES* | AI-GRC*

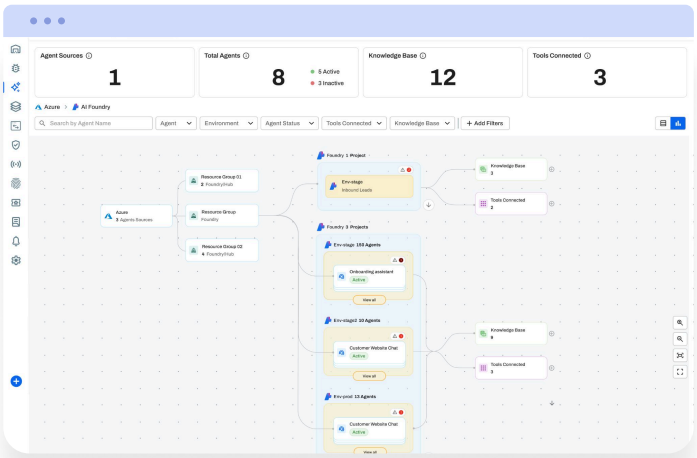
ACCUKNOX
Secure Code to Cognition™



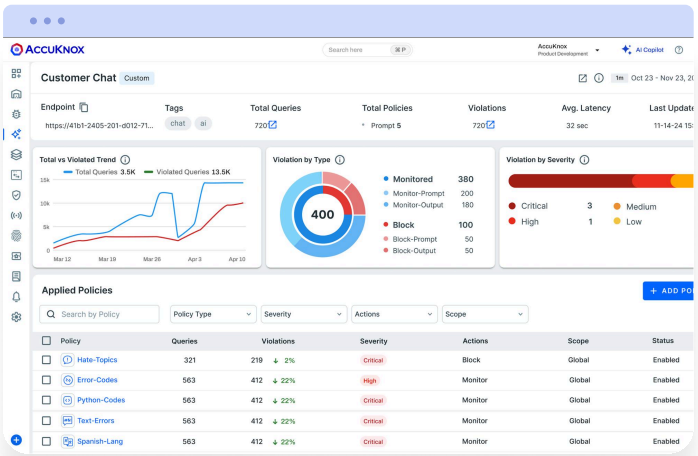
Dashboard: Your entire AI security posture, at a glance.



Application View: Violations by app, policy, and severity. All in one view.



Agents Graph View: Map every agent risk. Catch every attack surface.



LLM Defence: Real-time prompt firewall enforcement for your LLM interactions.

Customer Reviews

TAG

"AI is being deployed faster than it can be secured. Models, agents, and data pipelines introduce risks that legacy cloud tools weren't built for. AccuKnox AI-Security 2.0 brings Zero Trust to the AI layer, where the attack surface lives."

Dr. Edward Amoroso,
CEO TAG Infosphere | Former CSO, AT&T

SONESTA INTERNATIONAL HOTELS

"AccuKnox's strong roadmap offerings in API Security, AI/LLM Security made AccuKnox the best choice for AppSec/CloudSec platform."

David Billeter,
Cybersecurity Leader

About AccuKnox

AccuKnox is a Zero Trust CNAPP Cloud Security protects API Security, CDR, SIEM, Public clouds, Private clouds, Kubernetes, VMs, Bare metals, IoT Edge, and 5G security.

kubernetes CERTIFIED SERVICE PROVIDER | AICPA SOC 2 Formerly SAS 70 Reports | CLOUD NATIVE COMPUTING FOUNDATION SILVER MEMBER | NEPHIO | aws PARTNER Containers Software Competency

ACCUKNOX

linkedin.com/accuknox

@AccuKnox