OWASP® | GenAI SECURITY PROJECT
TOP 10 FOR LLM AND GENERATIVE AI

AI SECURITY SOLUTIONS INITIATIVE

Q3 ‘2025

AI Security Solutions Landscape

For Agentic AI Applications

The Solutions Landscape monitors and maps the full Agentic AI lifecycle, focusing on the DevOps–SecOps intersection to meet evolving security needs. Guided by the Agentic AI Threats and Mitigations guide and SecOps tasks, it highlights open-source and commercial solutions by stage, identifying their coverage of Agentic SecOps duties and threat mitigation, and leverages industry and community input as a peer-reviewed resource for navigating agentic AI’s shifting security challenges. Updated Quarterly.

https://genai.owasp.org/ai-security-solutions-landscape/

OWASP | GenAI SECURITY PROJECT
TOP 10 FOR LLM AND GENERATIVE AI

CHEAT SHEET genai.owasp.org

Agentic AI Security Landscape - Q2/3 2025

https://genai.owasp.org/ai-security-solutions-landscape/

Operate

- NRI SECURE
- GuardionAI
- Citadel AI
- HiveTrace
- Pillar
- MINDGARD
- Microsoft
- CORTEX CLOUD BY PALO ALTO NETWORKS
- NOMA
- TROJ.AI
- Straiker
- splx AI

Deploy

- splx AI
- Straiker
- HiveTrace
- TREND MICRO
- NOMA
- POMERIUM
- Pillar
- pangea
- Citadel AI
- TROJ.AI

Release

- splx AI
- CORTEX CLOUD BY PALO ALTO NETWORKS
- NOMA
- Pillar
- zenity

Test & Evaluate

- NOMA
- VULCAN
- splx AI
- ai+me
- EUNOMATIX
- TREND MICRO
- CORTEX CLOUD BY PALO ALTO NETWORKS
- Citadel AI
- ADVERSA
- TROJ.AI
- zenity
- Straiker
- HiveTrace
- MINDGARD
- Pillar

Develop & Experiment

- Pensar
- NOMA
- Pillar
- Straiker
- TROJ.AI
- MINDGARD

Augment, Fine Tune Data

- CORTEX CLOUD BY PALO ALTO NETWORKS
- NOMA
- zenity
- Pillar

Scope & Plan

- CORTEX CLOUD BY PALO ALTO NETWORKS
- NOMA
- Straiker
- zenity
- Pillar

Monitor

- tenable AI
- Tumeryk AI Trust Score™
- INFOTECT
- fiddler
- zenity
- NOMA
- HiveTrace
- pangea
- TROJ.AI
- Straiker
- CORTEX CLOUD BY PALO ALTO NETWORKS
- NRI SECURE
- GuardionAI
- Microsoft

Govern

- HUMAN
- unbound
- fiddler
- zenity
- pangea
- TROJ.AI
- Straiker
- splx AI
- Pillar
- Prompt:
- TROJ.AI
- NOMA
- GuardionAI

🟦 Open Source 🟨 Gold Sponsors ⬜ Silver Sponsors

Source; OWASP Gen AI Security Solutions Landscape Guide 2025. Q1

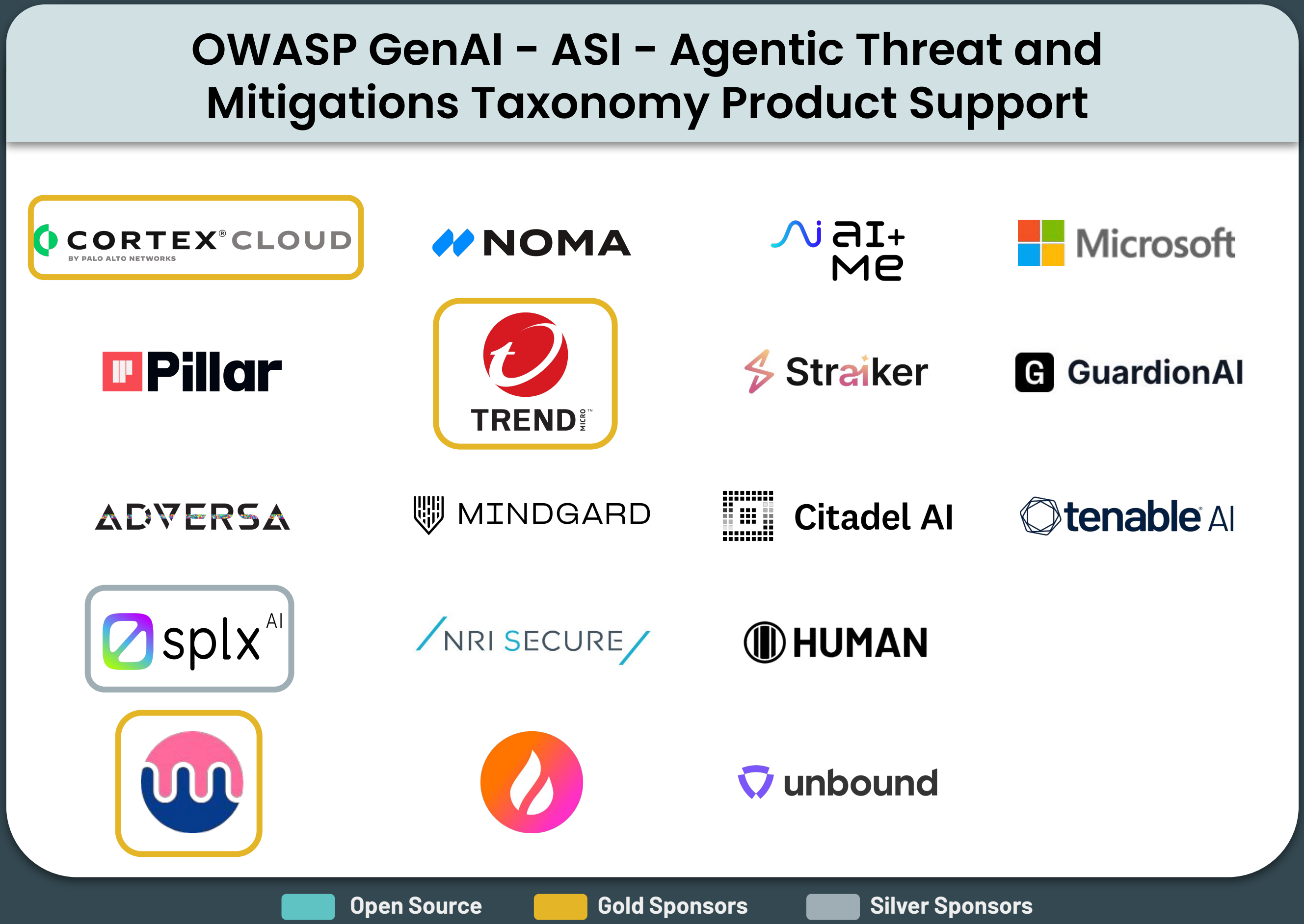
OWASP | GenAI SECURITY PROJECT TOP 10 FOR LLM AND GENERATIVE AI

CHEAT SHEET genai.owasp.org

https://genai.owasp.org/ai-security-solutions-landscape/

OWASP ASI Agentic Taxonomy Reporting & Support Built Into These Products.

18 Solution providers and open source projects have implemented the OWASP Agentic Risk and mitigations taxonomy directly into their products to help organizations identify and measure their security posture and readiness related to Agentic Ai applications and systems.

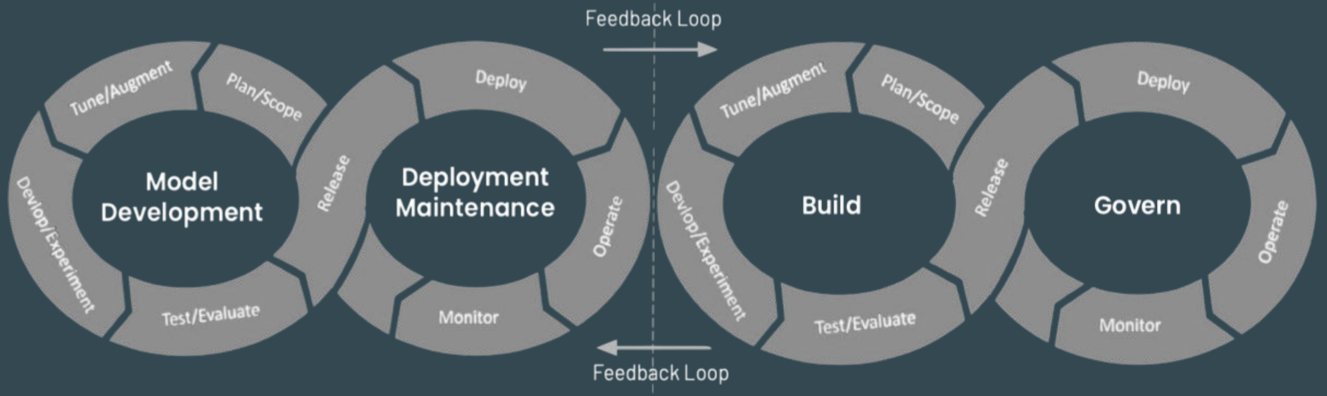

OWASP GenAI - ASI - Agentic Threat and Mitigations Taxonomy Product Support

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

CHEAT SHEET genai.owasp.org

https://genai.owasp.org/ai-security-solutions-landscape/

OWASP Agentic AI SecOps Framework



The **Agentic AI SecOps Framework** addresses the evolving security demands of next-generation AI systems as they transition from simple large language model (LLM) calls to fully autonomous, multi-agent architectures. The framework extends existing DevOps and SecOps methodologies to promote secure Agentic AI development, ensuring that organizations can maintain security, reliability, compliance, and auditability while safely embracing the capabilities of agentic AI.

Plan & Scope

- Conduct agentic threat modeling
- Identify system-wide non-human identities and authentication protocols.
- Draft policies for agent privilege boundaries, tool scopes and delegation logic.
- Define controls for memory scoping, isolation, and long-term persistence rules.

Augment & Fine Tune Data

- Apply differential privacy or obfuscation on sensitive knowledge injected into agent memory.
- Agent Action Audit

Dev & Experiment

- Perform SAST/DAST on agent planning code, tool wrappers, and plugin interfaces.
- Harden agent loop logic against infinite loops, unsafe function routing, unauth self-modification.
- Validate connector contracts
- Implement policy enforcement hooks in App Frameworks
    - e.g. LangGraph, CrewAI, or Semantic Kernel flows.

Test & Evaluation

- Available Agent Scanning
- Conduct adversarial red-teaming: goal drift
- Run multi-agent scenario simulations for collusion, misalignment, or deception detection.
- Validate agent decisions against expected goal plans.
- Sandboxed testing of all tool calls— code execution or cloud API triggers

Release

- Generate and verify model + agent + tool SBOMs - shared responsibility
- Sign model weights, plugin manifests, and memory snapshots.
- Ensure policy bundles are cryptographically validated at deploy time.
- Register all agents in an internal trust registry with capability descriptors

Deploy

- Enforce zero-trust policies between agents, tools, and external APIs
- Rotate all shared secrets, keys, and tokens with ephemeral, scoped credentials.
- Apply runtime guardrails (e.g., LLM firewalls, tool allowlists)
- Configure inter-agent authorization policies based on capabilities and roles

Operate

- Monitor agent memory mutation patterns for drift, poisoning, unauth overwrites.
- Detect task replay, infinite delegation, or hallucination.
- Enable human-in-the-loop override thresholds on high-risk actions.
- Continuously scan loaded plugins for CVEs and privilege escalation vectors.
- Runtime guardrails and moderation, and tool use.

Monitor

- Correlate telemetry from agent step tracing, tool execution, and message logs.
- Alert on anomalies like goal reversal, unexpected plan depth, adversarial-input, excessive tool usage, or rapid inter-agent chatter.
- SAudit reflection accuracy by comparing stated and observed planning outcomes.
- Ensure use of immutable logs (e.g., Sigstore, Immudb) for forensic readiness.

Govern

- Enforce role- and task-based access policies across agent populations and tool access.
- Automate agent versioning, expiration, and rotation policies.
- Align control evidence with frameworks like EU AI Act, NIST AI RMF, and ISO/IEC 42001.
- Automate goal alignment audits, including adversarial review of long-term agent memory.

Source; OWASP Gen AI Security Solutions Landscape Guide 2025. Q2

Agentic AI Security Landscape - Q2/3 2025

https://genai.owasp.org/ai-security-solutions-landscape/

Scope & Plan

Pillar

NOMA

Straiker

zenity

CORTEX® CLOUD
BY PALO ALTO NETWORKS

🟦 Open Source 🟨 Gold Sponsors ⬜ Silver Sponsors

During planning of agentic AI apps, SecOps and DevOps must embed security in the design, focusing on non-human identities, agent threat modeling, privilege boundaries, and authentication. Memory scoping and isolation are critical to prevent data leaks. Early collaboration aligns agent workflows and tools with enforceable security, unlike traditional post-design security..

Agentic DevOps	Agentic SecOps
• Define the business goal and translate into agent goals & roles • Choose model families (chat-LLM vs. multimodal) & hosting mode. • Define agent architecture patterns (single, hierarchical, swarm) • Identify external services and tooling • Design inter-agent communication and tool workflows • Select memory pattern (short-term context vs long-term e.g. vector DB). • Create initial threat model and Service Level Objectives.	• Conduct agentic threat modeling (referencing the threat modeling approach from the GenAI Security Project - Agentic Security Initiative) • Identify system-wide non-human identities (NHIs) and determine authentication protocols (e.g., SPIFFE, mTLS). • Draft policies for agent privilege boundaries, tool scopes (e.g., MCP), and delegation logic. • Define controls for memory scoping, isolation, and long-term persistence rules.

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

Agentic AI Security Landscape - Q2/3 2025

https://genai.owasp.org/ai-security-solutions-landscape/

Augment, Fine Tune Data

CORTEX® CLOUD BY PALO ALTO NETWORKS

Pillar

zenity

NOMA

🟦 Open Source 🟨 Gold Sponsors ⬜ Silver Sponsors

In data augmentation & fine-tuning, SecOps works with DevOps to prevent risks from poisoned data, adversarial tuning, and reasoning traces. They sanitize datasets, validate alignment, log provenance, and protect sensitive memory with privacy controls. This ensures compliant, trustworthy agentic AI before deployment.

Agentic DevOps	Agentic SecOps
• Collect domain-specific corpora that agents will reference during planning & reflection. • Generate tool-schema embeddings so planners can choose the right action. • Fine-tune/refine LLM on task-specific dialogues that include multi-step reasoning traces (ReAct, Tree-of-Thought). • Populate seed "agent memory" (company knowledge, rules).	• Scan datasets for prompt-poisoning, biased instructions, or encoded policy bypasses. • Validate RLHF traces for ethical alignment, adversarial manipulation, or leakage of secrets. • Register data lineage and provenance in immutable logs. • Apply differential privacy or obfuscation on sensitive knowledge injected into agent memory. • Agent Action Audit

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

OWASP | GenAI SECURITY PROJECT
TOP 10 FOR LLM AND GENERATIVE AI

https://genai.owasp.org/ai-security-solutions-landscape/

CHEAT SHEET genai.owasp.org

Agentic AI Security Landscape - Q2/3 2025

https://genai.owasp.org/ai-security-solutions-landscape/

Develop & Experiment

Pensar

NOMA

Pillar

Straiker

TROJ.AI

MINDGARD

In the Development & Experimentation phase of agentic AI, SecOps partners with DevOps to secure dynamic agent loops, inter-agent comms, and API/plugin use. They validate I/O contracts, embed policy hooks, and test resilience to prevent unsafe behaviors. Security shifts from static code focus to real-time orchestration and co-engineering secure experimentation.

Agentic DevOps	Agentic SecOps
● Implement agent loops (Observe-Plan-Act-Reflect) with frameworks such as LangGraph / AutoGen. ● Build manager-worker graphs; encode delegation policies. ● Wire plugins for each external API (e.g., MCP) and enforce input/output schemas. ● Prototype interagent protocol (e.g. A2A) handshake and capability negotiation. ● Iterate on prompts, system instructions, and guard-functions; run sandbox tests.	● Perform SAST/DAST on agent planning code, tool wrappers, and plugin interfaces. ● Harden agent loop logic against infinite loops, unsafe function routing, and unauthorized self-modification. ● Validate connector (e.g., MCP) contracts (input/output schemas and permissions). ● Implement policy enforcement hooks in Frameworks ○ e.g. LangGraph, CrewAI, or Semantic Kernel flows.

🟦 Open Source 🟨 Gold Sponsors ⬜ Silver Sponsors

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

OWASP | GenAI SECURITY PROJECT TOP 10 FOR LLM AND GENERATIVE AI

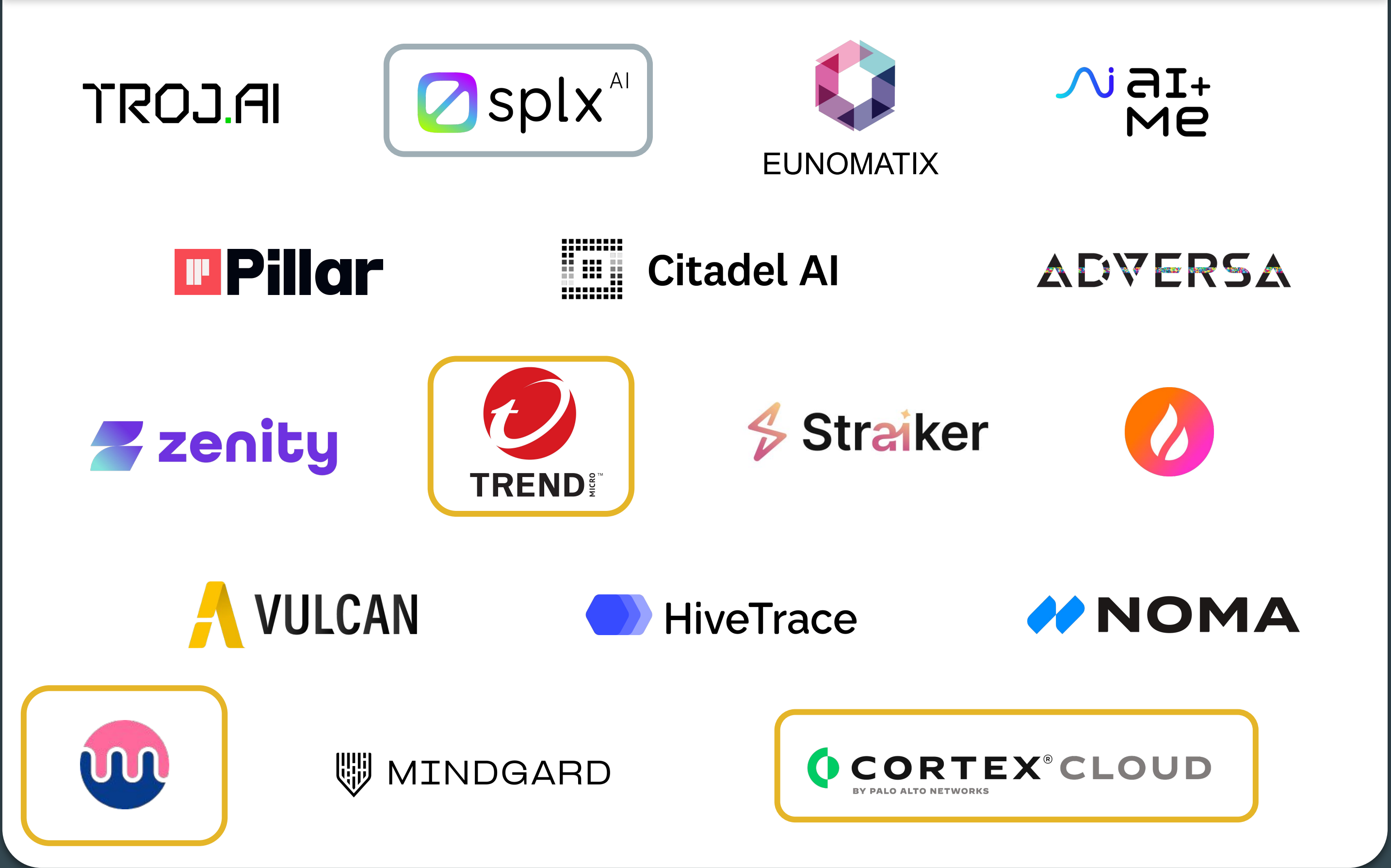
https://genai.owasp.org/ai-security-solutions-landscape/

CHEAT SHEET genai.owasp.org

Agentic AI Security Landscape - Q2/3 2025

https://genai.owasp.org/ai-security-solutions-landscape/

Test & Evaluate



🟦 Open Source 🟨 Gold Sponsors ⬜ Silver Sponsors

In Test & Evaluation, SecOps partners with DevOps to stress-test agentic AI in adversarial conditions, targeting emergent risks like goal drift, prompt injection, and tool misuse. They run red-team simulations, sandbox tool/API calls, and validate decisions in multi-agent setups, focusing on behavioral security beyond traditional QA or pen testing..

Agentic DevOps	Agentic SecOps
• Spin up synthetic multi-agent arenas to stress-test negotiation, bidding and consensus flows. • Run goal-drift, prompt-injection, and resource-exhaustion scenarios against the planner. • Benchmark reflection latency and memory-poisoning resilience. • Validate generated tool calls in a sandbox for RCE / over-scope.	• Available Agent Scanning • Conduct adversarial red-teaming: goal drift, prompt injection, hallucination chaining, and over-permissioned tool usage. • Run multi-agent scenario simulations for collusion, misalignment, or deception detection. • Validate agent decisions against expected goal plans. • Sandboxed testing of all tool calls—particularly code execution or cloud API triggers.

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

OWASP® | GenAI SECURITY PROJECT
TOP 10 FOR LLM AND GENERATIVE AI

https://genai.owasp.org/ai-security-solutions-landscape/

CHEAT SHEET genai.owasp.org

Agentic AI Security Landscape – Q2/3 2025

https://genai.owasp.org/ai-security-solutions-landscape/

Release

splx AI

CORTEX® CLOUD BY PALO ALTO NETWORKS

zenity

NOMA

Pillar

🟦 Open Source 🟨 Gold Sponsors ⬜ Silver Sponsors

In the Release phase, SecOps teams work with DevOps to securely package, validate, and register agentic AI apps. They sign model weights, plugins, and memory to prevent tampering, verify SBOMs for all components, enforce cryptographically validated policies, and register agents in secure capability registries, ensuring trusted, auditable deployments.

Agentic DevOps	Agentic SecOps
• Package agent graphs, plugins, policies, and memory snapshots • Generate Model & Tool SBOMs; sign artefacts (Sigstore). - shared responsibility • Publish agent capability-cards to an internal A2A registry.	• Generate and verify model + agent + tool SBOMs - shared responsibility • Sign model weights, plugin manifests, and memory snapshots. • Ensure policy bundles (e.g., OPA/Rego) are cryptographically validated at deploy time. • Register all agents in an internal trust registry with capability descriptors.

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

Agentic AI Security Landscape – Q2/3 2025

https://genai.owasp.org/ai-security-solutions-landscape/

Deploy

splx AI

NOMA

TREND MICRO

POMERIUM

HiveTrace

Pillar

Straiker

Citadel AI

🟦 Open Source 🟨 Gold Sponsors ⬜ Silver Sponsors

In the Deploy phase, SecOps partners with DevOps to enable secure, policy-compliant activation of agentic AI. They enforce zero-trust comms, rotate ephemeral credentials, set LLM firewalls and allowlists, and apply fine-grained authorization so each agent runs with least privilege, reducing risks in multi-agent environments..

Agentic DevOps	Agentic SecOps
● Provision vector DB, memory store, tool side-cars, and service-mesh with mTLS for A2A traffic. ● Apply least-privilege IAM roles to every agent (non-human identities). ● Load initial long-term memory and register agents with discovery service. ● Enable runtime guardrails / LLM firewall	● Enforce zero-trust policies between agents, tools, and external APIs via mTLS and fine-grained RBAC. ● Rotate all shared secrets, keys, and tokens with ephemeral, scoped credentials. ● Apply runtime guardrails (e.g., LLM firewalls, tool allowlists) before production traffic is enabled. ● Configure inter-agent authorization policies based on capabilities and roles

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

OWASP | GenAI SECURITY PROJECT
TOP 10 FOR LLM AND GENERATIVE AI

https://genai.owasp.org/ai-security-solutions-landscape/

CHEAT SHEET genai.owasp.org

Agentic AI Security Landscape - Q2/3 2025

https://genai.owasp.org/ai-security-solutions-landscape/

Operate

NRI SECURE

GuardionAI

Microsoft

HiveTrace

Citadel AI

MINDGARD

NOMA

Straiker

Pillar

CORTEX® CLOUD BY PALO ALTO NETWORKS

splx AI

🟦 Open Source 🟨 Gold Sponsors ⬜ Silver Sponsors

In the Operate phase, SecOps teams partner with DevOps to secure the dynamic footprint of agentic AI, where agents evolve, adapt, and act in changing environments. They monitor memory mutations to prevent drift or poisoning, detect abnormal loops or misuse, enforce HITL overrides, and scan plugins for risks. This persistent, real-time vigilance ensures secure, resilient operations as systems scale and self-orchestrate..

Agentic DevOps	Agentic SecOps
● Run SRE playbooks: auto-scale inference pods, rotate keys/tokens, prune memory. ● Collect feedback / RLHF traces; schedule periodic self-evaluation tasks. ● Trigger automated reflection or human-in-the-loop when agent confidence drops. ● - Orchestrate inter-agent workflows.	● Monitor agent memory mutation patterns for drift, poisoning, or unauthorized overwrites. ● Detect task replay, infinite delegation, or hallucination loops. ● Enable human-in-the-loop (HITL) override thresholds on high-risk or ambiguous actions. ● Continuously scan loaded plugins for CVEs and privilege escalation vectors. ● Runtime guardrails & moderation; anomalous tool use.

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

OWASP | GenAI SECURITY PROJECT
TOP 10 FOR LLM AND GENERATIVE AI

https://genai.owasp.org/ai-security-solutions-landscape/

CHEAT SHEET genai.owasp.org

Agentic AI Security Landscape - Q2/3 2025

https://genai.owasp.org/ai-security-solutions-landscape/

Monitor

tenable AI

Tumeryk
AI Trust Score™

fiddler

zenity

NOMA

iNFOTECT

TROJ.AI

Straiker

HiveTrace

NRI SECURE

GuardionAI

Microsoft

CORTEX® CLOUD
BY PALO ALTO NETWORKS

pangea

🟦 Open Source 🟨 Gold Sponsors ⬜ Silver Sponsors

In the Monitor phase, SecOps teams collaborate with DevOps to secure agentic AI's dynamic, evolving footprint. Unlike static systems, these agents produce context-rich traces that shift with reasoning and interactions. SecOps correlates agent steps, tool calls, and inter-agent comms to catch anomalies like goal reversal or adversarial inputs, using immutable logs and audits to drive proactive, behavior-aware security.

Agentic DevOps	Agentic SecOps
● Stream agent-step telemetry via OpenTelemetry; correlate tool errors with planning nodes. ● Track KPIs: goal-completion rate, average reasoning depth, vector-store growth, inter-agent latency. ● Alert on anomaly patterns (looping, hallucination cascades, excessive privilege use)..	● Correlate telemetry from agent step tracing, tool execution, and message logs. ● Alert on anomalies like goal reversal, unexpected plan depth, adversarial-input, excessive tool usage, or rapid inter-agent chatter. ● Audit reflection accuracy by comparing stated and observed planning outcomes. ● Use immutable logs (e.g., Sigstore, Immudb) for forensic readiness.

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

OWASP | GenAI SECURITY PROJECT
TOP 10 FOR LLM AND GENERATIVE AI

https://genai.owasp.org/ai-security-solutions-landscape/

CHEAT SHEET genai.owasp.org

Agentic AI Security Landscape - Q2/3 2025

https://genai.owasp.org/ai-security-solutions-landscape/

Govern

HUMAN
splx AI
zenity
fiddler
TROJ.AI
Straiker
unbound
Pillar
Prompt:
pangea
NOMA
GuardionAI

🟦 Open Source 🟨 Gold Sponsors ⬜ Silver Sponsors

In the Govern phase, SecOps partners with DevOps to uphold compliance, access control, and lifecycle governance for evolving agentic AI. They enforce role- and task-based policies, automate agent versioning and retirement, and guard against privilege creep. With immutable logs, audits, and alignment to AI regulations, they ensure long-term security, accountability, and trust in dynamic multi-agent systems.

Agentic DevOps	LLMSecOps
● Maintain registry of agent versions, roles, and approved tools; enforce retirement policy. ● Run quarterly attestation of A2A trust graph and MCP connector scopes. ● Archive immutable logs for audit; map evidence to EU AI Act / NIST RMF controls. ● Periodically review alignment metrics and update constitutional rules.	● Enforce role- and task-based access policies across agent populations and their tool access. ● Automate agent versioning, expiration, and rotation policies. ● Align control evidence with frameworks like EU AI Act, NIST AI RMF, and ISO/IEC 42001. ● Automate goal alignment audits, including adversarial review of long-term agent memory.

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

Acknowledgement

OWASP Gen AI Solutions Landscape Initiative : https://genai.owasp.org/ai-security-solutions-landscape/

Lead: Scott Clinton

Initiative Slack Channel: #team-genai-ai-solutions-landscape-initiative

Contributors

Aurora Starita
Bryan Nakayama
Dennys Pereira
Emmanuel Guilherme
Fabrizio Cilli
Garvin LeClaire
Helen Oakley
Ishan Anand
Jason Ross
Marcel Winandy
Markus Hupfauer
Migel Fernandes
Mohit Yadav
Rachel James
Rico Komenda
Talesh Seeparsan
Teruhiro Tagomori
Todd Hathaway
Ron F. Del Rosario
Vaibhav Malik

Reviewers

Andy Smith
Arun John
Aurora Starita
Blanca Rivera Campos
Bryan Nakayama
Dan Guido
Dennys Pereira
Emmanuel Guilherme
Fabrizio Cilli
Garvin LeClaire
Heather Linn
Helen Oakley
Ishan Anand
Jason Ross
Joshua Berkoh
Marcel Winandy
Markus Hupfauer
Migel Fernandes
Mohit Yadav
Rachel James
Rammohan Thirupasur
Rico Komenda
Rammohan Thirupasur
Talesh Seeparsan
Teruhiro Tagomori
Todd Hathaway
Ron F. Del Rosario
Vaibhav Malik

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

OWASP | GenAI SECURITY PROJECT TOP 10 FOR LLM AND GENERATIVE AI

https://genai.owasp.org/ai-security-solutions-landscape/

Gen AI Security Project Sponsors

Supporting Community Operations And Outreach Through Direct Financial Sponsorship

GenAI SECURITY PROJECT GOLD PROJECT SPONSOR TOP 10 for LLM and Generative AI

ByteDance | HIDDENLAYER | Mend.io | snyk | paloalto NETWORKS | TREND MICRO

GenAI SECURITY PROJECT SILVER PROJECT SPONSOR TOP 10 for LLM and Generative AI

splx AI | Calypso AI | Acuvity | pangea | Cobalt

LAKERA | securiti | ActiveFence | Prompt: | TROJ.AI

Lasso SECURITY | PROTECTO | PROMPTARMOR | Synack | CT cyber

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31

Contributing to the Landscape Guide

For Agentic AI

Use the **QR Code** and associated form to submit an Agentic AI Security Landscape entry

Source; OWASP Gen AI Security Solutions Landscape Guide 2025.Q2/Q31