AI SECURITY WITH ORCA

Detect Shadow AI

in code, cloud, AI-native platforms, and runtime

Secure AI

across every builder

15 min

to full deployment across a customer's environment

THE CHALLENGE

AI Risk Has Grown Exponentially Across Two Groups of Builders

AI is no longer a pilot. 90% of developers regularly use at least one AI coding tool at work (JetBrains, 2026), and by 2027, at least 30% of application security exposures will result from usage of vibe coding practices (Gartner, "Hype Cycle for Application Security, 2025," July 22, 2025). The attack surface is growing in two directions, both inside and outside your DevOps pipelines.

Traditional builders: AI inside your pipelines

Engineering and data teams call managed models like Amazon Bedrock, Azure OpenAI, and Google Vertex AI. They self-host open models on compute nobody registered. They install AI packages with known vulnerabilities, and connect coding agents to repos and APIs through MCP servers. Increasingly, those teams deploy AI agents that act on their own: calling tools, reading data, and running under the IAM role of whatever workload hosts them, often with the same identity as the humans who use them. An agent's risk is in everything the agent can reach.

The structural problem is that most AI security point tools cover one layer. One sees the model. Another sees the agent. A third sees the code. None of them knows that the exposed model endpoint sits two hops from a production database.

AI AppGen builders: AI outside your pipelines

Citizen builders generate apps by prompting AI, publish them to infrastructure they don't necessarily own, and provision a database along the way. A single AI-generated app has five parts: the agent or builder that made it, the models it uses, the database it stores PII in, the hosting service, and the cloud resources it runs on. Five vendors, five owners. None of them is security.

These apps don't ship with high security settings by default. Public URLs, plaintext secrets, missing authentication, and no session limits are platform defaults, not deliberate choices. And the Verizon 2026 DBIR ranked shadow AI as the third most common non-malicious insider action, a 4x increase in one year.

©2026 ORCA SECURITY. ALL RIGHTS RESERVED.

THE SOLUTION

Unified Security for Every Layer of Your AI Footprint

Orca secures every layer of AI your organization builds with: the model, the agent, the code it writes, and the app it ships. Every AI finding flows into the same Unified Data Model, attack path analysis, and alert workflows you already use to secure your cloud environment. AI risk gets prioritized against real exploitability, not a separate severity scale.

Three principles shape how it works:

01

Complete coverage of the AI footprint. One inventory of every AI surface, including the ones no one told you about.

02

AI risk in cloud context. Attack paths connect an exposed model to the database it touches, or an AppGen tool to the exposed database with PII.

03

Security that fits how builders work. Agentless, read-only connections. No new workflow for engineers or citizen developers, and findings land in the queue your team already works.

PILLAR 1

For Traditional Builders

For the models, agents, and AI workloads your engineering and data teams run in DevOps pipelines, AI Security Posture Management (AI-SPM) scans your environment agentlessly and surfaces managed AI services, self-hosted models, and the AI packages and agents running on your workloads.

The AI Bill of Materials (AIBOM) turns that into a searchable inventory you can filter by account, business unit, or model provider. For teams that need runtime depth, Orca Sensor adds Runtime AI Security, which observes LLM and MCP activity while flagging prompt injection, jailbreak attempts, and sensitive data in prompts as they happen.

ORCA ANSWERS

Which of our AI models and agents are exposed, over-privileged, or within reach of sensitive data, and which of those matter first?

PILLAR 2

For AI AppGen Builders

For the apps that skip your pipeline entirely, AI AppGen Security covers every stage an AI-generated app travels and connects them, so security teams see the whole app instead of five disconnected pieces: the agent or builder that made it, the models it uses, the database it stores PII in, the hosting service, and the cloud resources it runs on.

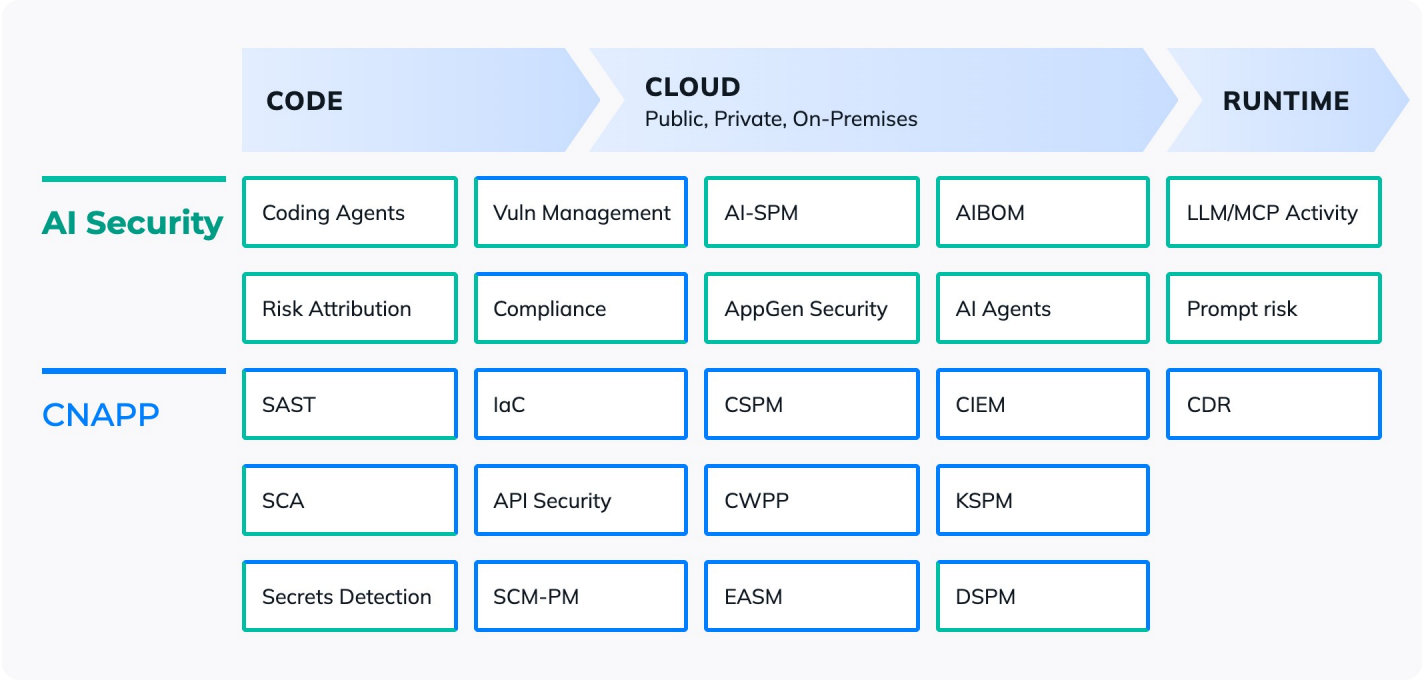
Orca agentlessly discovers the platforms builders already use, like Lovable, Supabase, OpenAI and Vercel, and detects what resources are misconfigured or exposed. Through native integrations, security teams can extend the Unified Data Model, connect the context back to the root cause, and drive remediation on those platforms.

ORCA ANSWERS

What AppGen tools are being used, where are misconfigurations creating compliance gaps, and what sensitive data is exposed?

COVERAGE

AI Security Across Code, Cloud, and Runtime



AI Security for Traditional Builders

- Inventory of cloud-managed and self-hosted AI models, packages, and agents, including shadow AI
- AI-BOM, including MCP servers and AI agents
- Sensitive data detection: leaked AI credentials, PII in training data sets
- Coding agent visibility and risk attribution
- Runtime AI Security with Orca Sensor
- AI compliance frameworks & customizations
- Attack path analysis for AI workloads

AI Security for AppGen Builders

- Discover what AI tools citizen builders are using
- Detect leaked credentials and secrets
- Fix misconfigured identities and projects
- Find exposed databases with PII

THE PLATFORM

How Orca Secures AI

- **SideScanning™:** Orca's patented, agentless technology reads workload data directly from the cloud provider's block runtime storage layer, out-of-band and read-only, so full deployment takes minutes instead of the weeks or months agent rollouts typically require, with zero performance impact on production systems.
- **Unified Data Model:** Cloud, identity, workload, data, code, and AI context all live in one graph instead of siloed tools, so an over-privileged role is automatically connected to the AWS Bedrock agent API that it can use, the exposed workload it sits on, and the data it can reach.
- **Orca AI:** A full suite of AI tools built for security teams. Pre-built and custom AI agents drive investigation and response, remediation, detection and prioritization. Native MCP support connects Orca's context into the tools your team already uses. AI Code Fix drafts remediations and enables a one-click pull request. An AI chatbot enables natural language searching and more.
- **Orca Sensor:** An optional, lightweight eBPF sensor added on top of the agentless foundation for teams that need real-time, in-workload detection, so runtime alerts inherit full cloud, identity, and code context.
- **Dynamic Risk Scoring:** Every finding is scored on reachability, exploitability, and blast radius across the full attack path, not a static CVSS label, with the full reasoning behind the score shown alongside it, not just the number. The score updates automatically the moment the environment changes, so a "critical" finding doesn't stay critical after it's patched or isolated.

About Orca

Orca Security delivers security for the companies that build. As cloud, code, AI and app generation expand the attack surface, Orca transforms security risk into the context teams need to act. The Orca Platform provides complete visibility across cloud, AI, and application environments, correlates risk across every layer, and prioritizes the exposures that matter most so organizations can remediate faster and innovate with confidence. Trusted by hundreds of organizations, including SAP, Autodesk, Gannett, Lemonade, and Digital Turbine, Orca is backed by leading investors including Temasek, CapitalG, ICONIQ Capital, and Redpoint Ventures. Learn more at orca.security and book a personalized demo.

Ready to try it out?
Sign up for a demo. Visit orca.security/demo

©2026 ORCA SECURITY. ALL RIGHTS RESERVED.