

REVERSINGLABS FOR NVIDIA OPEN AGENT SAFETY PLATFORM

Trusted artifacts for autonomous agents

NVIDIA OpenShell governs what agents are allowed to do. ReversingLabs helps determine whether the software, packages, models, and files they depend on are safe to use.

Agents depend on more than prompts

Autonomous agents download packages, install skills, load models, execute software, and move files as part of normal operation. Each artifact becomes part of the agent's effective attack surface. NVIDIA OpenShell provides sandboxed execution and policy controls that govern what an agent may do. ReversingLabs adds intelligence about the artifacts those agents interact with, so security teams can apply policy with deeper context before software runs, while agents operate, and when they investigate.

Artifacts agents depend on

Sandbox images

Before agents run

Agent skills

Installed and updated by the agent

Packages

Open source dependencies

Models

AI and ML artifacts

Downloads

Files entering workflows

Uploads

Files leaving workflows

Three points of protection across the agent lifecycle

01 · BEFORE EXECUTION

Admit only what has been analyzed

ReversingLabs analyzes sandbox images, agent skills, packages, and models before they become part of an agent environment, so only approved artifacts move forward into OpenShell.

02 · DURING OPERATION

Inspect what agents send and receive

Artifact inspection and reputation add context as agents retrieve, send, or request files, packages, and dependencies, helping OpenShell policy decisions at runtime.

03 · INVESTIGATION

Investigate with evidence

Verdicts, reputation, and artifact traits enrich NVIDIA telemetry and your SIEM, giving security teams artifact-level evidence to explain what an agent encountered and respond.

Why ReversingLabs

Coverage

Intelligence on 422B+ searchable files, the largest private corpus of goodwill and malware

Speed

7.91 ms average hash lookup keeps checks lightweight

Locality

Cloud, on-premises, or air-gapped deployment options

Placement

Applied from admission through runtime and investigation

Get started with 100k lookups for free

Step 1 · Scan what your agents already use

Point Spectra Assure at your images, skills, packages, and models

Get a verdict on every artifact your agents pull in today

Starts day one, with no changes to your NVIDIA deployment

Step 2 · Pilot inspection on one agent workflow

Turn on artifact inspection for one OpenShell workflow

See what agents send, receive, and request before it is allowed

Measure latency and results on your own workloads

Step 3 · Bring the evidence to your SOC

Feed verdicts and reputation into your SIEM, SOAR, or XDR

Correlate agent activity with artifact-level context

Investigate anything stopped with Spectra Analyze

Build more trust into autonomous agent workflows.

Create a free account and start with 100k lookups. Questions? sales@reversinglabs.com

Get Started Today