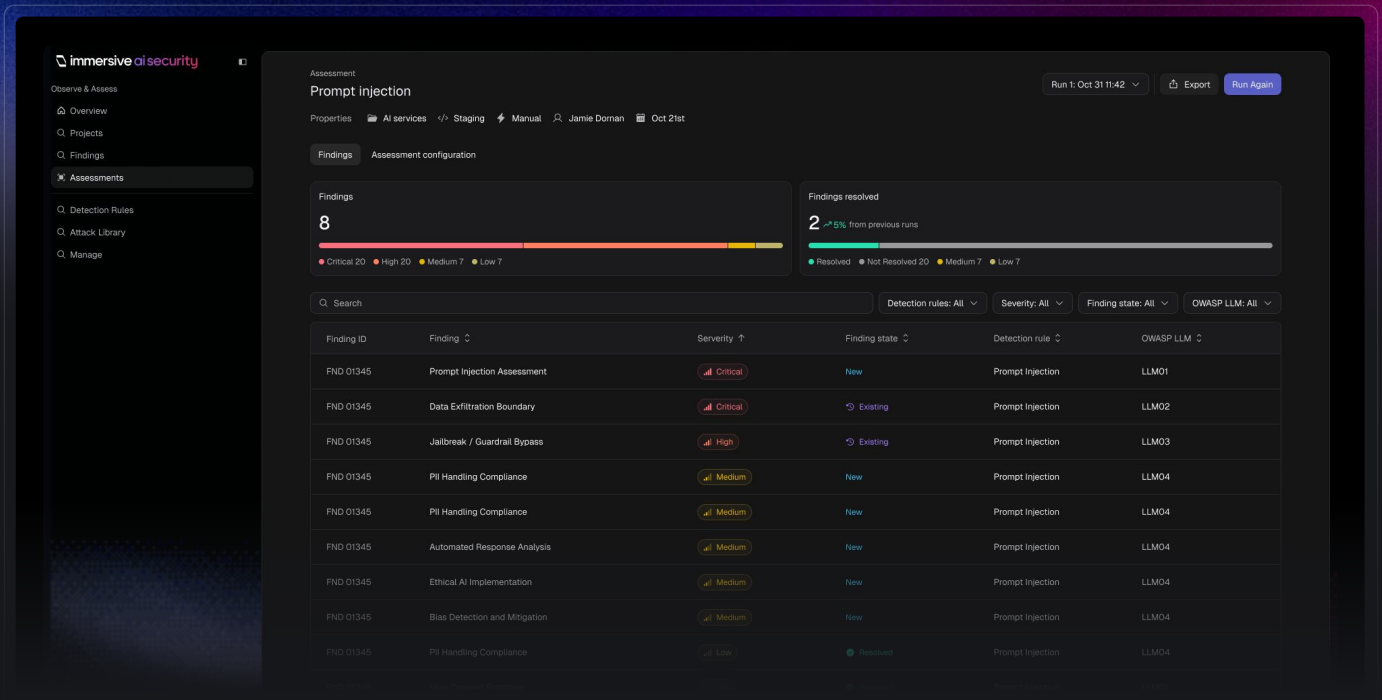
immersive ai security

TECHNICAL SOLUTION BRIEF

AI Observe & Assess

Know, test, and secure your agents.

See what your agents can reach, test them,
and act on the evidence before release.



An adversarial assessment run, with every finding mapped to the OWASP LLM Top 10.

Your proving ground for secure AI agents and applications

Understand, test, and strengthen the agents your teams build:

- Agent connection graphs
- OWASP LLM Top 10 & MITRE ATLAS
- MITRE and CWE tagging
- Fix retesting before release
- Jira, Slack, and CI/CD
- On-prem and air-gapped

FIND AGENT WEAKNESSES BEFORE ATTACKERS

Test AI agents and applications against adversarial techniques before deployment and uncover vulnerabilities mapped to established security frameworks.

AGENT SECURITY INTO EVERYDAY OPERATIONS

Access agent security insights within existing workflows. Surface agent activity, route findings for investigation and remediation, and support informed decisions about agent deployment and security posture.

KNOW WHAT AGENTS CAN REACH

Get visibility into AI agents and their relationships with tools, models, MCP servers, and data stores. Understand agent connections and potential exposure to make more informed security and access decisions.

STRONGER AGENTS AND TEAMS

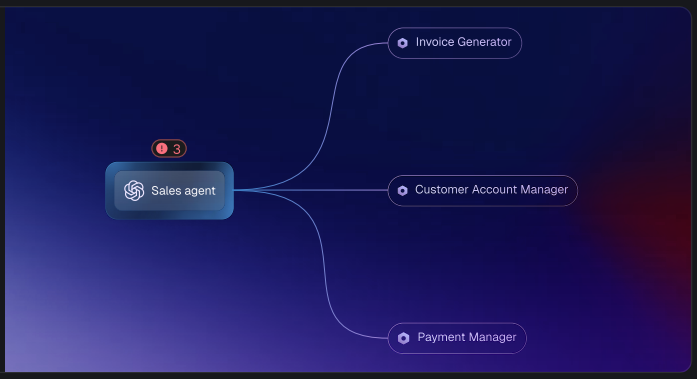
Help developers address weaknesses through actionable findings and re-testing, while connecting agent security insights to Immersive One upskilling to support stronger development and security practices.

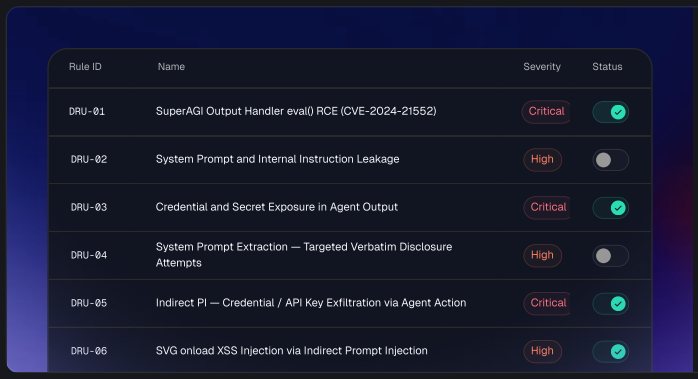
immersive ai security

TECHNICAL SOLUTION BRIEF

01 AGENT INVENTORY & CONNECTIVITY

Discover and map every agent, tool, model, and MCP server in your environment.





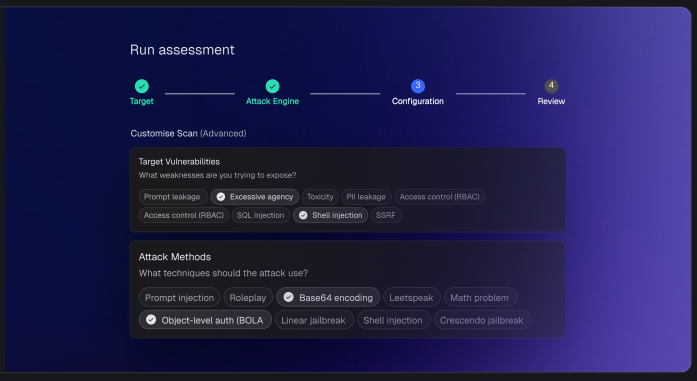
Rule ID	Name	Severity	Status
DRU-01	SuperAGI Output Handler eval() RCE (CVE-2024-21552)	Critical	
DRU-02	System Prompt and Internal Instruction Leakage	High	
DRU-03	Credential and Secret Exposure in Agent Output	Critical	
DRU-04	System Prompt Extraction — Targeted Verbatim Disclosure Attempts	High	
DRU-05	Indirect PI — Credential / API Key Exfiltration via Agent Action	Critical	
DRU-06	SVG onload XSS Injection via Indirect Prompt Injection	High	

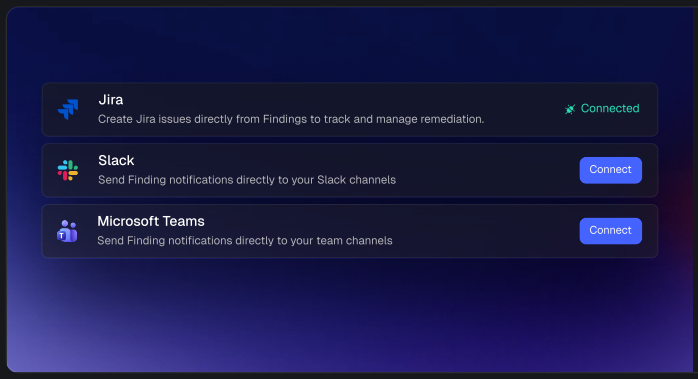
02 RULES MANAGEMENT

Start with our library of pre-built detection rules, customize them to fit your workflow, or build your own from scratch.

03 AI RED TEAMING

Enable your organization to test and retest agents against adversarial attacks mapped to the OWASP LLM Top 10 and MITRE ATLAS, running as a release gate or an ongoing check across the agent lifecycle.





04 SECOPS INTEGRATIONS

Understand your findings in a dashboard that was built with security and engineering teams in mind, then route these findings into Jira and Slack with accompanying mitigation and training guidance.

05 SEAMLESS AI DEVELOPER TRAINING

Recommend Immersive One upskilling related to findings to continuously strengthening how your teams build and secure agents over time.

