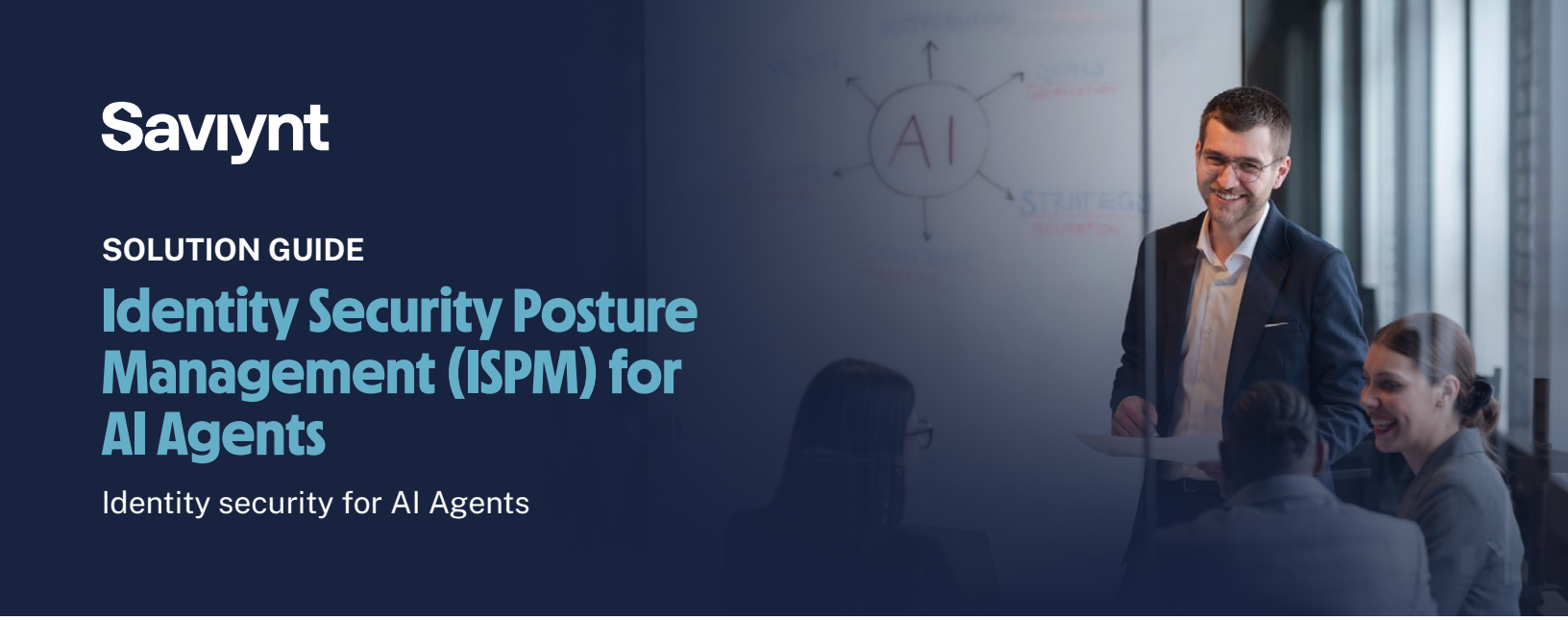
Saviynt

SOLUTION GUIDE

Identity Security Posture Management (ISPM) for AI Agents

Identity security for AI Agents

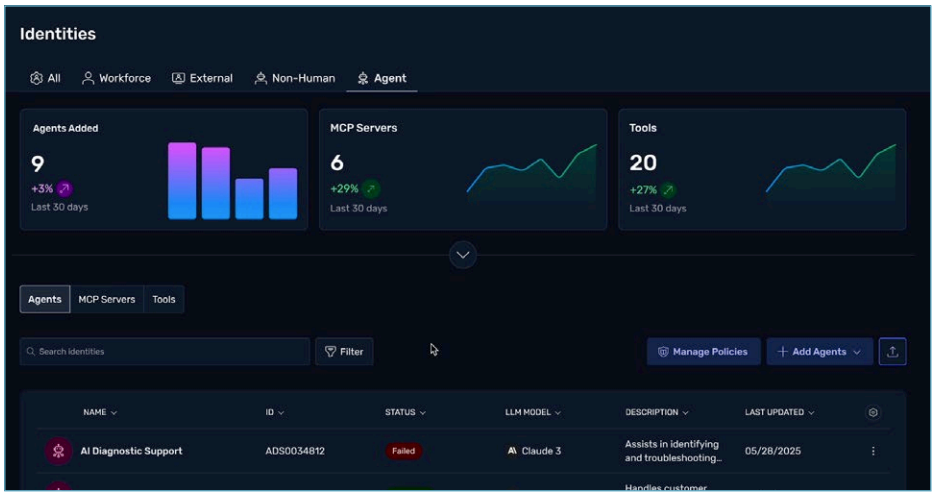


AI agents are reshaping the identity landscape, introducing new layers of security complexity. Unlike static non-human identities such as service accounts, they behave more like humans — making autonomous, dynamic decisions and requesting access without oversight. These behaviors create risks that today's NHI management solutions cannot address. To secure AI agents, identity must evolve beyond the perimeter to become the governance fabric — governing every agent, entitlement, and action, while ensuring all decisions are auditable and traceable.

Securing AI agents requires tailored controls across their core elements, including APIs that integrate LLMs with resources, Model Context Protocol (MCP) servers, tools, and the AI frameworks used to build these systems. Identity security for AI agents begins with discovering all components across both the infrastructure and intelligence layers — and understanding what each can access. Equally important is proactively establishing the right guardrails, mapping the full access path, and continuously tracking changes with a timeline for complete auditability.

Discover All AI Agents in a Single View

Gain complete visibility into every identity across all layers of your AI agents. Automatically discover not only the agents themselves but also their associated identities across infrastructure and intelligence stacks — such as MCP servers and tools. Easily identify newly added agents and register them in just a few clicks, reducing blind spots as your AI agent population continues to grow.



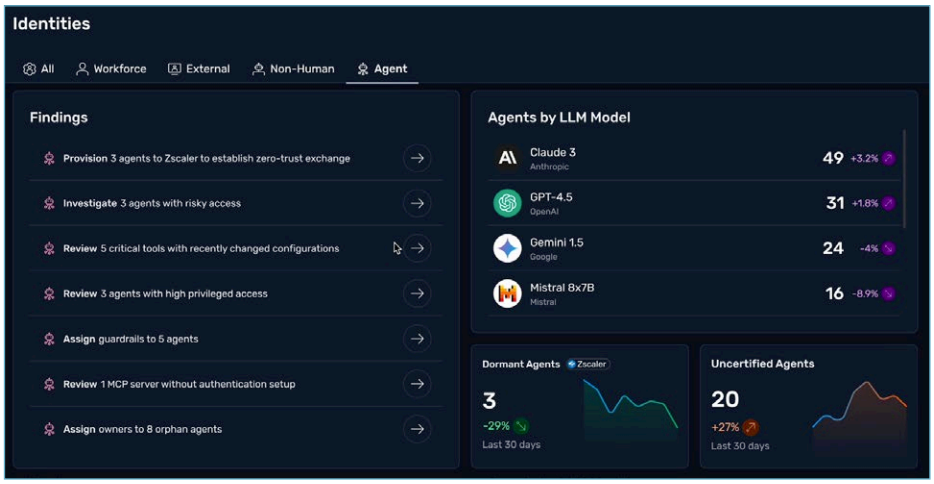
Prioritize and Eliminate AI Agent Vulnerabilities

Highlight critical AI agent risks that, if left unaddressed, can create serious vulnerabilities — such as missing guardrails against prompt injection attacks, excessive privileged roles, misconfigured settings, and gaps in governance across AI agents, MCP servers, and tools. With just a few clicks, you can move from reviewing findings to taking guided actions that effectively mitigate these risks.

What is AI identity?

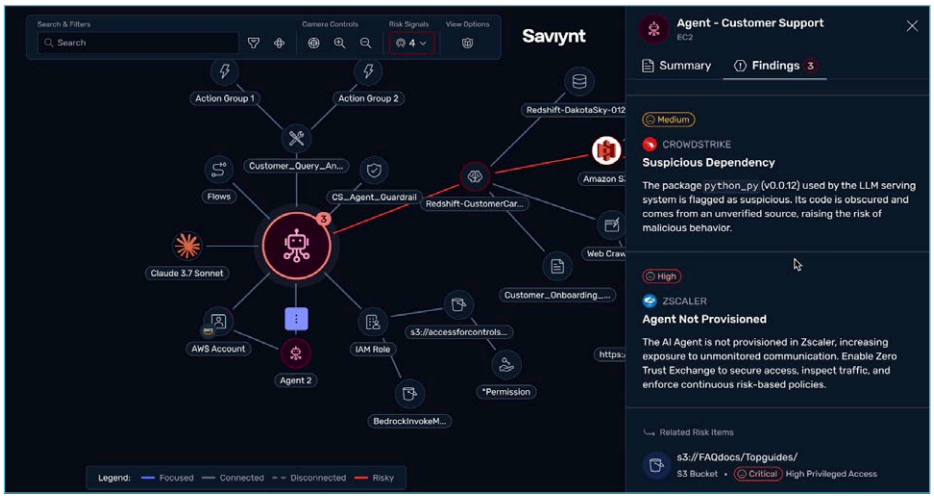
An AI agent isn't defined by a single identity — it's composed of many. Each brings its own access patterns, lifecycle requirements, and security risks. These identities can consume sensitive data, initiate actions, and operate autonomously. With AI, enterprises must now account for new classes of identities:

- **AI Agents** – copilots, bots, and reasoning engines that act on behalf of users.
- **MCP Servers** – orchestration layers connecting agents and models.
- **Tools** - external capabilities or functions that an AI agent can call to extend its reasoning and actions
- **Model Endpoints** – APIs that serve LLMs, often versioned with different contexts.
- **Agent Frameworks** – platforms that define agent behavior.



Map and Understand Risk Connections

Visualize every relationship your AI agents maintain — across roles, accounts, knowledge bases, guardrails, and more — in a single unified view. Drill down into child connections as needed, and visually trace risk signals back to their source for easy tracking. With a single click, access detailed insights and risk evaluation tags for each issue. You can also shift focus seamlessly — from an agent to an MCP server, for example — to investigate and remediate detected risks.



Stay Audit-Ready with Timeline Views

Capture every change to your AI agents — including newly added tools, created MCP servers, updated configurations, and more — while maintaining a clear, chronological record to stay audit-ready at all times. Each event is tracked for quick context, and with a single click, you can access a comprehensive summary of any change for deeper insights.

Discovery and Risk Prioritization

- Identify AI agents and their core elements, while detecting newly added, orphaned, or uncertified agents with associated risks. This provides a complete view of their AI agent landscape, eliminating blind spots.

Guardrail Enforcement

- Set guardrails to block harmful inputs and unsafe outputs, stopping prompt injection, data leaks, and misuse of sensitive resources while ensuring safe, accountable AI agent operations.

Access Map

- Visualize how each component across AI agent layers is connected, along with the entitlements, roles, and guardrails assigned. This visibility makes it easier to investigate risks and take effective remediation actions.

Timeline

- Display key lifecycle events, such as creation, knowledge base changes, guardrail updates, and attribute modifications, on a clear timeline. This enables organizations to stay audit-ready.

Next Steps

- Visit www.saviynt.com to learn more about The Identity Cloud from Saviynt
- Schedule a demo
- Contact us to learn more

ABOUT SAVIYNT

Saviynt's AI-powered identity platform manages and governs human and non-human access to all of an organization's applications, data, and business processes. Customers trust Saviynt to safeguard their digital assets, drive operational efficiency, and reduce compliance costs. Built for the AI age, Saviynt is today helping organizations safely accelerate their deployment and usage of AI. Saviynt is recognized as the leader in identity security, with solutions that protect and empower the world's leading brands, Fortune 500 companies and government institutions. For more information, please visit www.saviynt.com.

Saviynt

Headquarters, 1301 E
El Segundo Bl, Suite D, El Segundo, CA
90245, United States

310. 641. 1664 | info@saviynt.com
www.saviynt.com