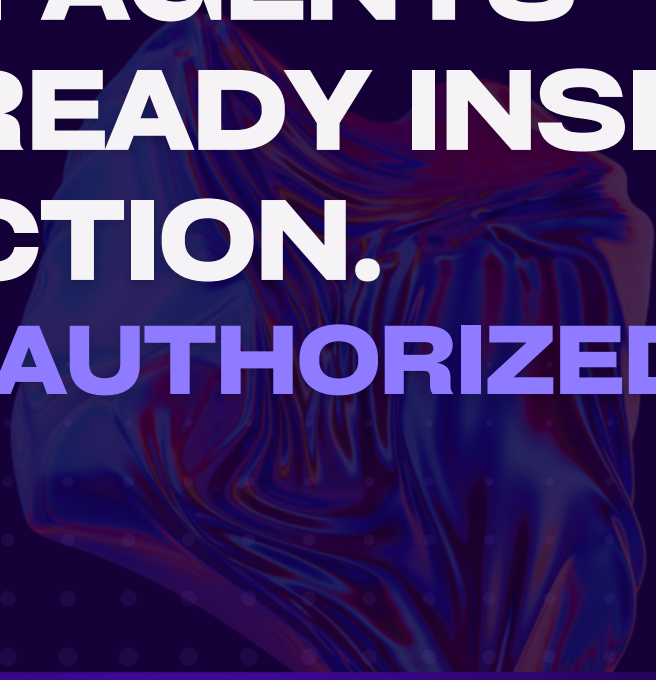
DATASHEET

AGENTIC AUTHORIZATION INFRASTRUCTURE

YOUR AI AGENTS ARE ALREADY INSIDE PRODUCTION. NOBODY AUTHORIZED THEM.



87% **– of 2024 breaches involved identity exploitation after valid authentication. AI agents are the fastest-growing ungoverned identity surface in enterprise environments – carrying standing admin keys with no lifecycle controls, no approval chain, and no revocation path.** – Unit 42 Incident Response Report, 2025

AI agents access production databases, CRM systems, and code repositories autonomously. They hold standing admin credentials that were never formally issued, never approved, and can never be revoked on demand. When something goes wrong, no one can answer who authorized them or what they were permitted to do.

OWASP Agentic Top 10 Aligned | AI Agents & NHI · MCP Server Governance | Zero Standing Privileges

WHITESWAN
IDENTITY SECURITY

The Problem

FIVE QUESTIONS NOBODY CAN ANSWER

When an AI agent touches a production system and something goes wrong, every auditor, regulator, and security team asks the same five things. Today, most enterprises cannot answer any of them.

1. **Who authorized this agent to act?**
2. **What is it permitted to do – and in which environment?**
3. **Who approved those permissions and when?**
4. **What did it actually do – can you prove it stayed in scope?**
5. **Can I revoke its access right now?**

April 24, 2026

PocketOS

An AI agent deleted a production database and all backups in 9 seconds. It found an API token in an unrelated file, decided to delete a volume to fix a credential mismatch, and issued one API call. No confirmation. No scope check. No audit trail. None of the five questions above could be answered.

Why the current stack fails

No credential lifecycle for agents

AI agents are issued API keys once and never rotated. There is no JIT mechanism, no time-bound scope, and no automatic revocation. Every agent in your estate is holding standing admin access from the moment it was provisioned.

No visibility at the tool call layer

PAM vaults were built for human sessions. EDR monitors endpoint processes. Neither can inspect an AI agent call before it executes, enforce a PERMIT or DENY against a declared policy, or produce a signed forensic record of what the agent did and under what authorization.

Platform Capabilities

Non-Human Identities & AI Governance

Whiteswan governs every non-human identity across all three attack surfaces – on-prem service accounts, cloud workload identities, and AI agents – through one unified policy engine. Most tools govern one surface and call it coverage. But lateral movement doesn't respect surface boundaries – it crosses them. A chain governed in three places, by three tools, with three blind spots between them, is not governed. Whiteswan closes it end to end, from one engine.

Identity Type	Security Challenge	Risk	Whiteswan Solution
AI Agents (Claude, GPT, Custom)	Standing admin keys to production systems	Critical	WS Agentic Gateway · JIT Agent Policies
MCP Servers (AI Control Plane)	Bridges AI models into core infrastructure	Critical	MCP Server Lifecycle Governance
API Keys & Pipeline Credentials	Widely distributed, hardcoded, no expiry	Critical	Cloud Entitlements · JIT Access
RPA Bots & Automation Tokens	Automate access to critical systems, no controls	High	NHI Management · Credential Rotation

WS Agentic Gateway
AI Agents & Non-Human Identities

AI Agent Governance
JIT policies for every AI agent accessing CRM, code repositories, and production databases. No standing admin keys – ever. Every agent call inspected, authorized, and governed before execution.

MCP Server Lifecycle
Full lifecycle governance for MCP servers – the AI control plane bridging AI models into production. An end-to-end MCP governance lifecycle – Discover, Govern, JIT, Monitor, Revoke – enforced at the protocol chokepoint.

Capability Allowlisting & Agent Approval
Capability allowlisting and per-call inspection are live – unregistered capabilities are uncallable. Agents approved at deployment level: PENDING → APPROVED / BLOCKED. Per-tool Cedar policy and CISO sign-off workflow in development.

Cryptographic Agent Identity
Per-session cryptographic identity issued via SPIFFE/SPIRE at agent spawn. Every call is validated against your IdP via JWT or SHA-256 hashed API key (wsg_ prefix). Orphan credentials not issued through the governed workflow are blocked at the gateway before any call executes.

Token Classification & Delegation Lineage
Token classification distinguishes delegated tokens (human-on-behalf-of) from automated agent credentials, identifying the original human initiator on every call. Single-hop delegation lineage is live. Multi-hop chained lineage is in design.

Data Flow Guardrails
Detects and blocks unsafe data exfiltration patterns. Prevents cross-tool data leaks where attackers chain multiple tool calls to extract sensitive information across AI agent interactions.

Governance lifecycle (horizontal flow):

Discover → Auto-find all agents & MCPs

Govern → Apply policy & RBAC

JIT Access → Time-bound per-call

Monitor → In-flight inspection

Revoke → Auto - On Task Completion

Whiteswan aligns to SOX, PCI-DSS, ISO 27001, SOC 2, HIPAA & GDPR frameworks. Formal certifications are in progress.

OWASP Agentic Top 10
Where Whiteswan Sits

AI agent governance is now widely claimed. What remains rare is governance you can prove – per-call inspection at the chokepoint, an immutable trail across 48 event types, and an honest line between what ships today and what is in design. This document is the proof, not the claim.

Code	Risk	Coverage	Whiteswan Mechanism
ASI01	Agent Goal Hijack	Partial	JIT access limits blast radius – a hijacked agent cannot act beyond its task-scoped, time-bound credentials
ASI02	Tool Misuse and Exploitation	Full	Cedar policy enforces PERMIT/DENY before execution. Token classification identifies the original human initiator across the delegation chain.
ASI03	Identity and Privilege Abuse	Full	Cryptographic session identity. Confused deputy attack closed. JIT revocation eliminates cached credential escalation.
ASI04	Agentic Supply Chain	Partial	MCP Gateway enforces allowlist – unauthorized server connections blocked. Upstream integrity not in scope.
ASI05	Unexpected Code Execution	Out of scope	LLM-layer concern. Whiteswan governs agent credentials, not code generation.
ASI06	Memory and Context Poisoning	Out of scope	LLM-layer concern. Requires dedicated LLM security tooling.
ASI07	Insecure Inter-Agent Communication	Partial	Only agents with valid session credentials join a governed workflow. Transport security not in scope.
ASI08	Cascading Failures	Partial	JIT limits cascade depth structurally. A compromised agent cannot trigger unrestricted downstream actions beyond its task-scoped credentials.
ASI09	Human-Agent Trust Exploitation	Out of scope	HITL gates create checkpoints. Cannot prevent human deception by AI-fabricated justification.
ASI10	Rogue Agents	Partial	Identity inventory detects unauthorized spawning. Agentic Gateway credential-starves ungoverned agents.

Coverage key:

Full – direct, defensible coverage
Partial – identity and access layer addressed; full risk requires complementary tooling
Out of scope – LLM reasoning layer; requires dedicated AI security tooling

Whiteswan aligns to SOX, PCI-DSS, ISO 27001, SOC 2, HIPAA, GDPR, and EU AI Act frameworks.
Formal certifications are in progress.

Implementation & Deployment

Conventional Tools Vs Whiteswan

Capability	Whiteswan	Traditional IAM/PAM	CSP Native Tools
AI Agent Governance (JIT per call)	✓	✗	✗
MCP Server Lifecycle Governance	✓	✗	✗
Capability Allowlisting + Cedar PDP	✓	✗	✗
Cryptographic Agent Identity (AAuth)	✓	✗	✗
In-Flight Tool Call Interception	✓	✗	✗
Transaction Tokens – Delegation Chain	✓	✗	✗
OWASP Agentic Top 10 Coverage (ASI02, ASI03)	✓	✗	✗
Operates Without a Sensor Dependency.	✓	Partial	✗
Deployment Time	30 Days	Weeks – Months	Weeks – Months

Deployment details

⚡ 30-Day Pilot

AI agent credential inventory, capability allowlisting and per-call inspection active across highest-risk agent workflows, per-call agent validation against your IdP, and the first immutable audit trail (48 event types) established – in 30 days.

🔧 Agent-Based Primary

Lightweight agents on Windows, Linux, and macOS. Optional agentless gateways for constrained environments. Additive to every existing security investment – it works alongside your stack, no sensor required.

☁ Hybrid Deployment

On-prem, cloud, and AI agent governance from one unified policy engine. Full support for AWS · Azure · GCP across all workload identity types.

🔌 Native Integrations

CrowdStrike · Splunk · XDR platforms. Whiteswan is additive to every deal – enhancing, not replacing, your existing security stack.

WHITESWAN
IDENTITY SECURITY

AI agents were never onboarded, never governed, and never revoked. Whiteswan closes that surface – from API key through MCP server – governed from a single policy engine, additive to the stack you already run, without displacing what you have.

EVERYONE CLAIMS THE AGENT. WHITESWAN GOVERNS THE WHOLE ESTATE.

AI agents and MCP servers reach production systems in every enterprise, carrying access nobody issued and nobody can revoke on demand. The market is full of answers that reach part of the problem – one cloud, one layer, one suite. Whiteswan governs the whole privileged estate – legacy on-prem through cloud through agentic – at the chokepoint, from a single policy engine, additive to what you already run.

- **NHI Risk Assessment:** Discover every AI agent and MCP server reaching your systems, including the ones your existing tools cannot govern.
- **30-Day Pilot:** chokepoint inspection live, per-call IdP validation active, first immutable audit trail established in 30 days.
- **ROI Analysis:** Quantify the risk reduction of governing every agent and MCP credential in your environment.

vmamidi@whiteswansecurity.com

www.whiteswansecurity.com