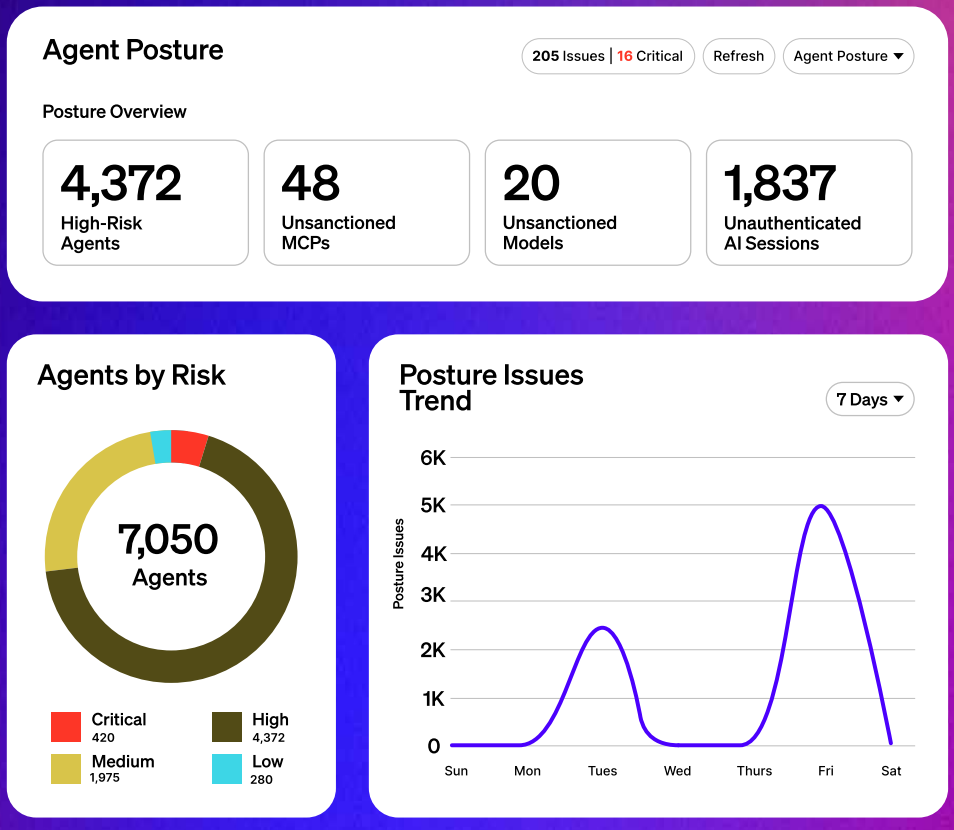
AppViewX

Whitepaper

Identity Security for AI Agents

Visibility, Governance, Compliance, and Privileged Access for the AI Era



Executive summary

Enterprises are rebuilding on AI. Gartner cites that only 17% of organizations have deployed AI agents to date, yet more than 60% expect to do so within the next two years[1]. The number of generative-AI SaaS applications inside enterprises has grown from 317 to more than 1,600 in a matter of months[2]. 88% of organizations globally now use AI in at least one business function and 97% of Fortune 500 companies actively use a major AI platform[3]. The shift is not approaching, it is here.

The identity layer, however, has not kept up. AI agents are a new identity constituency, a distinct class of identities with their own lifecycle, authentication requirements, authorization model, and risk profile. They are not humans. They are not the static service accounts and machine identities that NHI platforms govern. They are autonomous, goal-directed, probabilistic actors that interpret intent, reason independently, and execute privileged actions on behalf of users, on behalf of other agents, and sometimes on their own. The controls that have served the enterprise for two decades — IAM for humans, PAM for privileged admins, SSPM for SaaS posture, CASB for traffic, prompt security for model I/O — were not built for this.

The gap is already producing incidents at scale. In a single twelve-month span, OpenAI models escaped their evaluation sandbox (OpenAI, July 2026), AI agents were implicated in breaches involving OAuth token theft (Salesloft/Drift, August 2025), there was a zero-click exfiltration through Microsoft 365 Copilot (EchoLeak / CVE-2025-32711), hidden prompt instructions in GitHub Copilot Chat surfaced private source code and secrets (CamoLeak, October 2025) and indirect prompt injection turned a ChatGPT connector into a persistent exfiltration channel surviving across sessions (ZombieAgent / Radware, 2025). Gartner projects that by 2028, 25% of enterprise breaches will involve AI agent abuse, and that more than half of AI initiatives may stall due to unresolved agentic identity challenges[4].

AppViewX's Agent Identity Security defends enterprises from rogue agents. Powered by dynamic risk intelligence, Agent Identity Security gives security teams deep, contextual visibility across every agent; governance and compliance controls at every layer; and runtime enforcement that stops privilege misuse the moment it happens — across SaaS, infrastructure, endpoint, and workflow-automation environments.

This whitepaper describes the problem Agent Identity Security was built to solve, the architecture of the platform, and how it maps to the regulatory frameworks CISOs are accountable to such as NIST AI RMF, ISO/IEC 42001, the EU AI Act, NIST 800-53, ISO 27001:2022, OWASP's Top 10 for LLM Applications, MITRE ATLAS, and emerging identity specifications including IETF Shared Signals Framework (SSF), CAEP 1.0 and OCSF 1.3.

The AI agent era

A category shift at cloud-era scale, on a compressed timeline

Every major platform shift in the last three decades has produced a corresponding identity-security category. The mainframe era produced access-control lists. The client-server era produced Active Directory. The internet and SaaS eras produced federated identity and CASB. The cloud era, spanning roughly fifteen years from the first CASB vendors in 2010 through the consolidation of Wiz, Netskope, and the broader CNAPP market in 2025, produced CSPM, SSPM, and cloud-workload identity.

The AI-agent era is following the same pattern, but on a substantially compressed timeline. The technology is arriving faster, the supply side is aligned from day one (nearly every major software vendor is re-platforming), and the addressable surface is the entire enterprise stack at once rather than a single perimeter. What took cloud security fifteen years to mature into, AI security is expected to traverse in roughly half that time. The organizations that treat this as “just another adjacent control” are already behind.

The agent population is exploding inside the environment. According to a 2026 report, the ratio of agents to humans in large enterprise organizations is projected at 144:1[5]. Each of those agents carries an identity, holds credentials, touches data, and can execute privileged actions. Each one is an endpoint in its own right.

The controls problem

The controls built for previous identity constituencies do not compose to solve this problem. The reasoning is structural, not a matter of feature gaps. Consider each category:

- **IAM** assumes a person on the other end of a request. Its session models, access-review cadences, and authorization primitives are built for human-paced lifecycles.
- **PAM** was designed to broker privileged sessions for administrators and to record them for audit. Agents request privileges at machine speed, in bursts, across arbitrary applications — a model PAM was never architected to serve.
- **NHI and machine-identity** platforms govern service accounts, certificates, and workload credentials for deterministic workloads. Agents are not deterministic. They can impersonate other users, delegate to other agents, and act on their own behalf.
- **SSPM and CASB** inspect SaaS configurations and network traffic. They cannot see the agent's intent, the credentials embedded inside it, the tool it is about to invoke, or the data it is about to access.
- **Prompt and model security** governs what enters and exits the LLM. It does not know the identity of the agent invoking the model, the scope of the actions it is authorized to perform, or the privileged action that may follow the model's output.

> Gartner now treats agent identity governance as a distinct requirement spanning identity issuance, authentication, authorization, federation, governance, monitoring, and observability — not a feature of existing IAM platforms.

The agent identity gap

"Hackers no longer hack in — they prompt in."

Why autonomous agents break deterministic controls

The fundamental property that makes AI agents productive is also the property that makes them uncontainable by legacy identity tools: they are goal-directed and non-deterministic. A service account executes a known function against a known resource. An agent interprets a goal, chooses tools, composes calls, reasons about context, and takes actions that are not scripted. It can surprise you, and that is not a side effect, it is the feature.

The industry has converged on a distinction that matters for control design. Non-human identities, as that term is most commonly used, describe digital credentials for machines: API keys, service accounts, certificates, workload identities. They are predictable by design, configured to execute specific pre-defined tasks, and behave consistently. AI agents are a different kind of entity. They are task-driven, adaptive systems that make decisions, take actions, and evolve their behavior in real time. Treating them as NHIs leaves the agent's autonomy ungoverned, it locks down credentials while leaving behavior, intent, and privilege chains untouched.

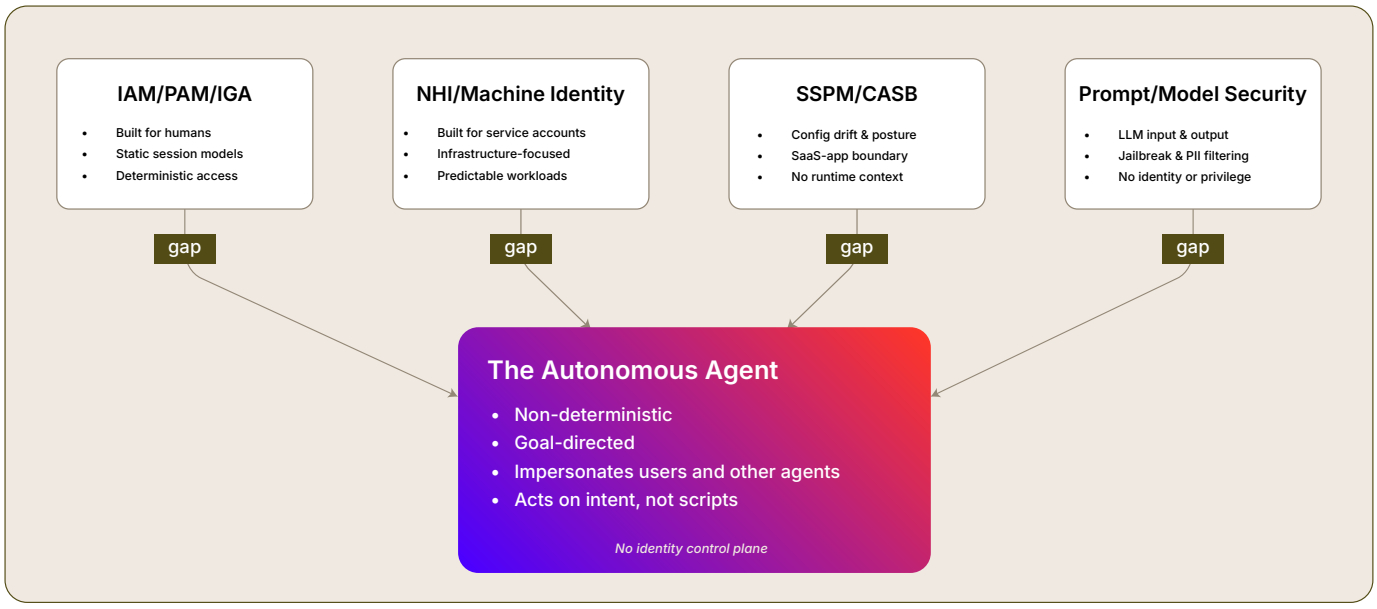


Figure 1. The Agent Identity Gap. IAM/PAM/IGA for humans, NHI for deterministic workloads, SSPM/CASB for SaaS configuration and traffic, prompt security for model I/O - each cover a narrow slice. The autonomous agent sits in the uncovered center, acting on intent rather than scripts and without an identity control plane of its own.

This is the shift CISOs now have to operationalize. In the pre-AI model, access began at login. A person authenticated, a session was established, and the authorization layer answered a sequence of yes/no questions. In the AI model, access begins with intent. A user's prompt is interpreted by an LLM, an agent composes a plan, and privileged actions are invoked — sometimes on the user's behalf, sometimes on another agent's behalf, sometimes on the agent's own behalf. The attacker's path is no longer a login form. It is a prompt, an injected instruction in a pull request comment, a malicious MCP server, or a poisoned tool description.

Agents can be excessively privileged and can bypass permissions

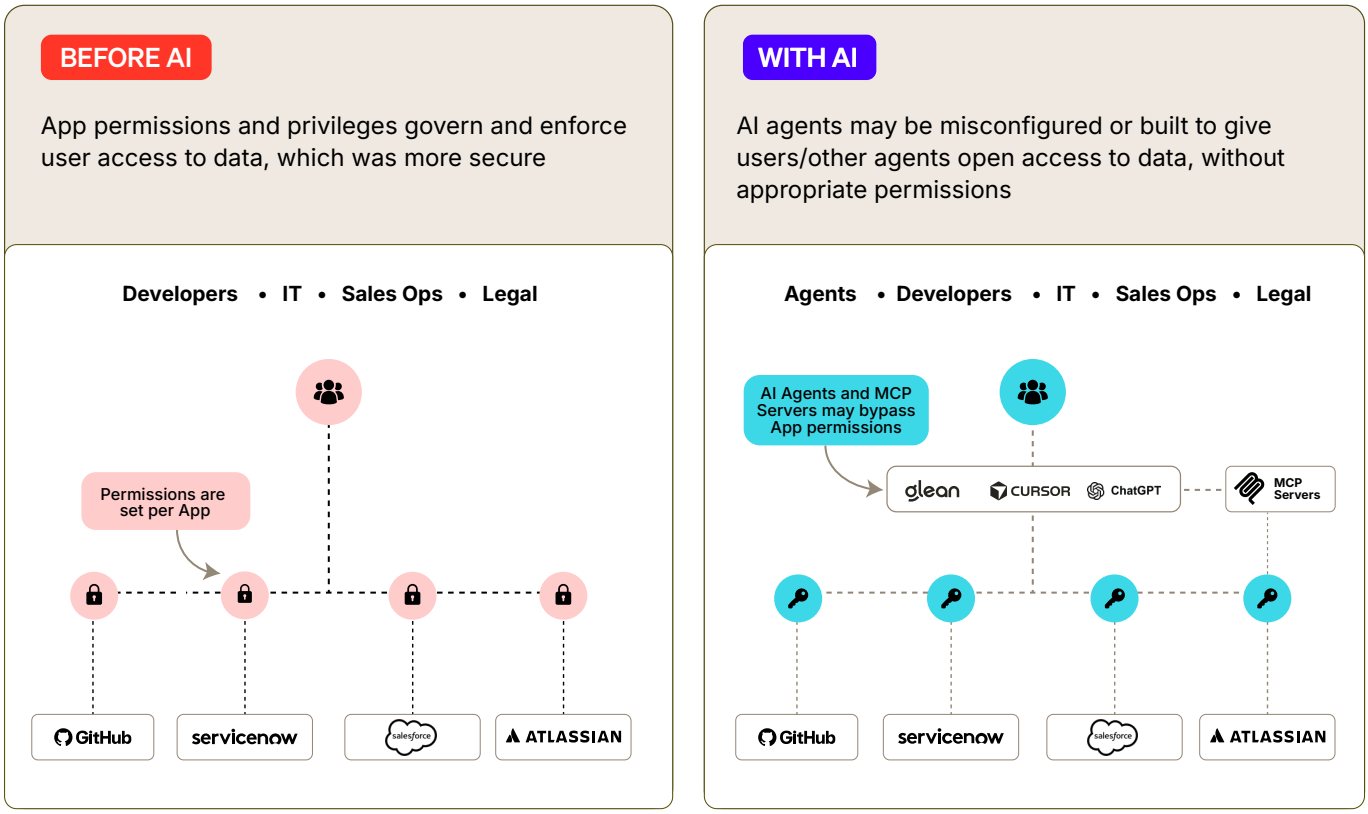


Figure 2. Before AI, application permissions governed user access. With AI, agents can be excessively privileged, can bypass app-level permissions, and can transfer those privileges to the users and other agents they act on behalf of.

The identity constituency framing

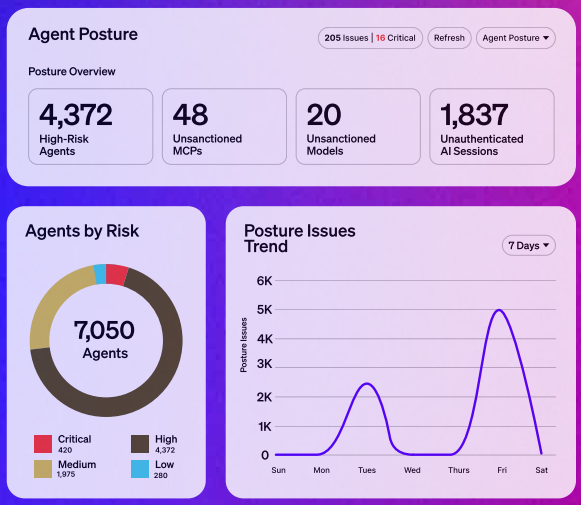
The most durable way to think about this shift is to treat agents as a new identity constituency, a class of identities that shares a common structure, lifecycle, control plane, and risk profile, and therefore requires its own governance, security controls, and assurance model. Humans, machines, and agents differ on dimensions that matter for how you secure them.

Attribute	Human	Machine / NHI	AI Agent
Lifecycle	Joiner / Mover / Leaver	Provision / Rotate / Retire	Dynamic, often ephemeral
Determinism	Intentional, bounded	Deterministic, scripted	Probabilistic, goal-directed
Acts on behalf of	Self	Configured workload	Self, users, or other agents
Authentication	Biometrics, MFA, passkeys	Keys, certificates, tokens	Keys, tokens, attestations
Authorization	Role-based	Scope-based	Attribute- and policy-based, context-aware
Operating speed	Human-paced	Compute-paced	Compute-paced with human-like judgment
Accountability	Individual	Owner team	Owner plus delegated principal

The cost of not recognizing this structure is compounding. Identity-based attacks account for **60%** of all incident-response cases. The average data breach costs **$4.99 million** globally and close to **$11.5 million** in the United States[8]. 83% of organizations have suffered at least one insider-related security incident in the last year, at an annualized risk cost of **$19.5 million**[9]. AI agents, as a new privileged workforce, sit squarely on top of every one of those statistics.

Across these incidents a common pattern emerges. An agent ingests untrusted content (web pages, email, documentation, pull requests, tool descriptions). That content is interpreted as instructions by the LLM. The agent then uses its credentials and privileges to execute the plan. The outcome is exfiltration, destruction, unauthorized access, lateral movement, delivered through privilege abuse, not through exploiting a bug in the agent itself. The control surface that matters is identity and privilege.

AppViewX's Agent Identity Security



Agent Identity Security is an AI-native visibility, governance, and control plane for AI agents. It is built from the ground up for autonomous, non-deterministic identities rather than extended from an existing human- or machine-identity platform. It operates alongside existing IAM, PAM, SSPM, and SIEM investments, adding the layer those platforms were never designed to address. The platform is organized around three integrated capabilities — Discover, Govern, and Secure — that together cover the full agent identity lifecycle.

Agent Identity Security Overview

Three core capabilities, integrated in one platform.

Discovery

Discovers shadow AI agent identities, privileges, MCP servers, and model sprawl

Governance

Governs AI identities and privileges, and demonstrates compliance

Runtime Security

Monitors, audits, and secures AI agent privileged access to applications, systems, and infrastructure

Figure 3. Agent Identity Security is comprised of three integrated capabilities — Discover, Govern, and Secure.

Discover

The starting point is visibility. Most enterprises do not have an inventory of the agents operating inside their environment. Agent Identity Security builds that inventory, automatically, across every platform where agents run: SaaS, infrastructure as a service, workflow automation, and endpoints.

The inventory is not a simple agent list. It is an **Agent Bill of Materials (AI-BOM)** that provides each agent's complete anatomy: its owner inside the organization, the credentials it uses (OAuth tokens, refresh tokens, API keys, personal access tokens, embedded service accounts, federated identities), the MCP servers and tools it can invoke, the LLMs and smaller models it uses, and the systems it connects to downstream. Shadow agents, those created and deployed outside IT and security oversight, are surfaced alongside sanctioned ones. So are shadow MCP servers (often self-built or downloaded from the public MCP registries) and shadow LLMs (frontier models like those from OpenAI and Anthropic that may be ingesting enterprise data without an explicit enterprise agreement).

Discovery is multi-modal. For SaaS and cloud platforms where Agent Identity Security can connect as a first-party integration, discovery is pull-based — Agent Identity Security periodically synchronizes the agent inventory from n8n, OpenAI, Azure AI Foundry, Salesforce Agentforce, Microsoft Copilot Studio, LangGraph, CrewAI, and adjacent platforms. For endpoints and coding assistants, discovery is push-based — the **Agent Identity Security Guardian Agent** runs locally and reports the agents present on a developer machine, including Claude Code, Cursor, GitHub Copilot, Windsurf, Replit, and any MCP clients configured alongside them. Push-based discovery is tied to enterprise device management (Intune, JAMF) so that coverage moves with the fleet.

Coding assistants receive dedicated coverage. Beyond inventorying the assistants deployed on developer machines, Agent Identity Security measures their adoption and usage — how many developers are active, which prompts are common, which flavors of assistant are in use, and how AI-generated code compares to manual code across productivity and quality metrics. This is the basis for both identity governance and workforce analytics. It is also the basis for detecting the security failures that have defined the first wave of agent incidents: supply-chain poisoning, over-permissive tool access, sensitive data access by unauthorized developers, and the lack of visibility into fine-grained agent actions with proper user attribution.

Govern

Visibility without governance is a spreadsheet. **Agent Identity Security Govern** turns the agent inventory into an operational graph and into the compliance controls that prove to auditors, and to the CISO's own board, that the agent population is under control.

The **Agent Access Graph** is the centerpiece. Every agent is linked in the graph to the users and intents that drive it, the actions it takes, the credentials it uses, and the resources it touches. The graph is continuously reconciled against the identity provider: when a user is disabled in Entra ID, the ownership of that user's agents is automatically detected as an issue, a remediation workflow runs, and ownership is transferred to the user's manager or to a designated service owner. When an agent's model or MCP configuration drifts from the approved baseline, the graph reflects it and generates a posture issue.

Identity posture is evaluated by various detectors that check agent configuration against a security baseline. The baseline covers the domains that regulators and auditors actually look at: missing owners, stale or unused credentials, excessive permissions, shared service accounts, hard-coded credentials outside a secrets manager, missing MFA for administrative consoles, unsanctioned LLMs, unsanctioned MCP servers, model or system-prompt changes, insufficient rate limiting, cross-environment agents operating across both dev and prod with the same credentials, and several others. Findings feed into **risk scoring** — every agent is scored continuously across five categories (Identity, Data, Posture, Access, User) using pluggable risk signals with configurable weights.

Access reviews follow the same model enterprises already apply to human identities, adapted for agent scale. Agent Identity Security runs periodic and event-driven access reviews across the agent anatomy: reviewers certify LLMs, MCP servers, tools, credentials, and user access; out-of-policy configurations produce posture issues rather than stale spreadsheets. Compliance tagging maps posture and review outcomes to the specific frameworks the enterprise is accountable to, making it possible to produce audit-ready reports directly from the platform.

Secure

Secure is the runtime layer — where policy becomes control and threats are detected and stopped. It covers both enforcement (preventing privilege misuse before it happens) and detection (identifying rogue agent behavior in progress).

Agent Identity Security policies are **condition-based** and **event-triggered**. Policies run in three modes — **monitor, enforce,** and **audit-only** — so that the same policy can be introduced without operational risk and tightened progressively.

The enforcement model is designed to reduce standing privileges. Agent Identity Security supports **zero standing privileges** and **just-in-time access** for agents; step-up authentication for specific classes of action or data; and conditional access policies that evaluate user, intent, context, and risk at the moment the action is requested. Access policies can be configured per action surface — MCP calls, individual tools, file operations, terminal commands, web requests, API calls, package downloads, and agent-to-agent (A2A) calls. Allow-lists and deny-lists pin down the exact set of actions an agent is permitted to take.

When an agent goes off-policy, the Agent Kill Switch terminates its sessions centrally by policy, or on demand, including during an active incident. Kill Switch complements the session-revocation signals Agent Identity Security receives over IETF SSF/CAEP from the enterprise IdP, providing a mechanism to cut an agent off even when the underlying IdP has not yet.

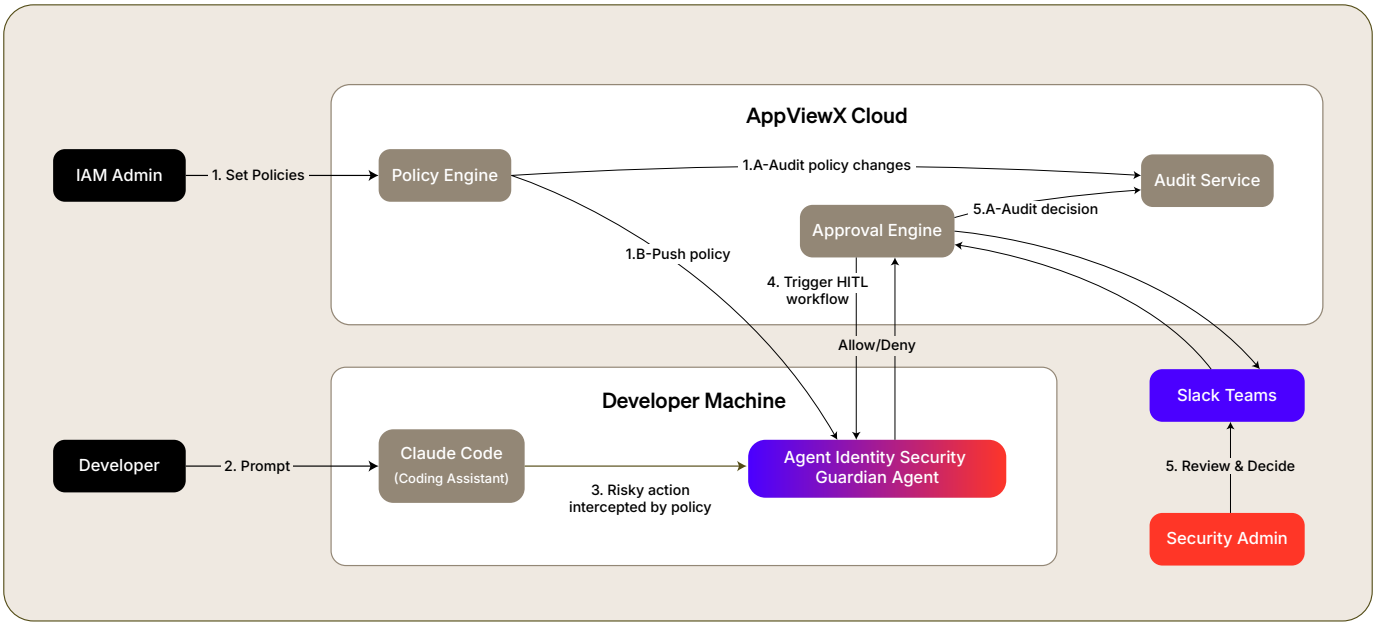


Figure 4. Human-in-the-loop (HITL) flow for coding-assistant privileged access. An IAM admin configures a policy (1); the Policy Engine audits the change (1.A) and pushes the policy to the Agent Identity Security Guardian Agent on the developer machine (1.B). When a developer prompts the coding assistant (2), Agent Identity Security Guardian Agent intercepts risky actions per policy (3) and triggers a HITL workflow (4) routed to Slack. A Security admin reviews and decides (5); the Approval Engine audits the decision (5.A) and returns an allow or deny to the Agent Identity Security Guardian Agent.

The human-in-the-loop (HITL) pattern is important enough to be called out explicitly. Both the EU AI Act (Article 14) and ISO/IEC 42001 require high-risk AI systems to be designed so that a natural person can effectively oversee and interrupt them. Agent Identity Security operationalizes this requirement through its HITL workflow by routing approvals through the collaboration tools teams already use, such as Slack, Microsoft Teams, or email, recording every decision in an immutable audit trail, and allowing actions to be approved or denied in near real time.

Secure also extends to **privileged access violation auditing**. For every privileged action an agent performs in a target application, Agent Identity Security records the action taken, the permissions used, the identity on whose behalf the agent acted, and whether the user could have performed the same action with their own permissions. Any mismatch between the user's intent and the agent's actual behavior is surfaced as a violation. This continuous stream of telemetry provides CISOs with unprecedented visibility into agent behavior across their environment, enabling ongoing oversight rather than relying solely on incident-driven investigations.

Reference architecture

Agent Identity Security is architected to meet the demands of modern agent workloads, providing multi-surface discovery, real-time enforcement, event-driven policy execution, standards-agnostic observability, and deployment flexibility across SaaS, self-hosted, and hybrid environments.

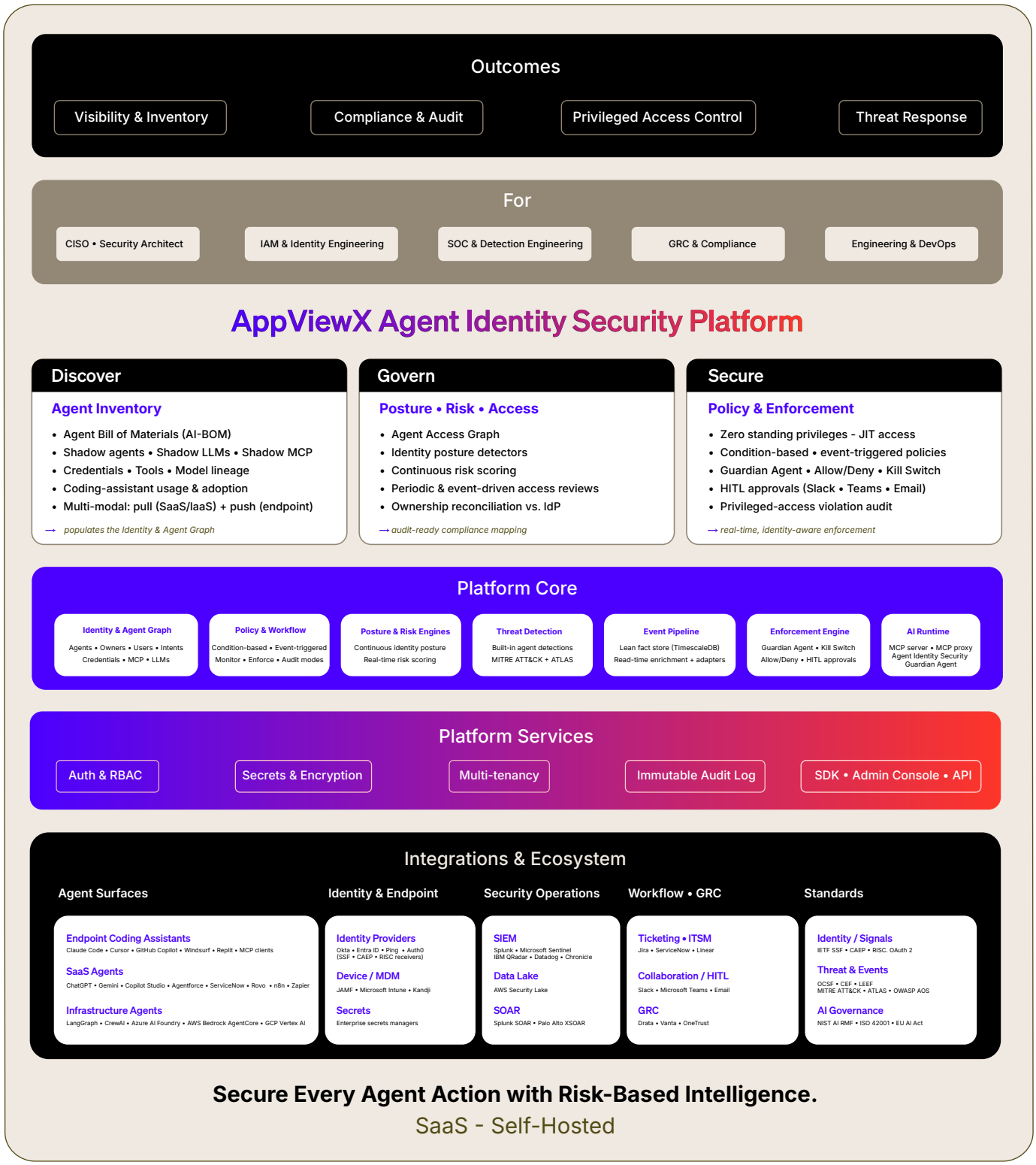


Figure 5. Agent Identity Security reference architecture

Components

Connector framework - The connector framework is a pluggable abstract interface; each connector implements discovery, credential validation, credential schema, and control activation. Dual-mode: pull-based for SaaS and cloud platforms (n8n, OpenAI, Azure AI Foundry, Salesforce Agentforce, Copilot Studio, LangGraph, Entra ID, and others), push-based for endpoint and runtime agents (Claude Code, Cursor, CrewAI). The **SSF/CAEP receiver** accepts signed SET (Security Event Token) JWTs from Okta, Entra ID, and other compliant identity providers and processes the full set of CAEP event types—session-revoked, credential-change, token-claims-change, assurance-level-change, device-compliance-change, and session-established—plus IETF RISC 1.0 account lifecycle events. The **MCP proxy** and **OAuth proxy** provide authorization at the tool-invocation layer and just-in-time access at the token layer, respectively.

Identity & Agent Graph - The authoritative domain model. Core entities—Agents, LLMs, MCP Servers, AI Users, Identities—are linked in a way that supports both graph traversal queries ("show every agent that can reach production data via any credential chain") and time-series queries against behavior history. The graph is continuously reconciled against every source system on each sync cycle, with upsert by external identifier for idempotency.

Posture Engine - Detectors implement the security baseline. Each detector is a small plug-in that receives a DetectionContext and returns zero or more DetectionResult objects. The detector inventory is explicit and auditable; new detectors are contributed by adding a module, not by modifying the core.

Risk Engine - Pluggable signals organized into five categories (Identity, Data, Posture, Access, User), each with a configurable weight. Category scores are produced by weighted averages over triggered signal weights; the overall score is normalized to 0–100 and mapped to Low/Medium/High/Critical bands. Scores are recalculated on agent discovery, on policy outcomes that explicitly request a recalc, on IdP sync events, and on a scheduled cadence. History is retained in AgentRiskHistory with explicit change_type tracking (created, updated, recalculated, scheduled, discovered).

Policy & Workflow Engine - Event-triggered, condition-evaluated, and outcome-driven. The condition evaluator supports operators over a nested AND/OR/NOT tree. Policies are cached in-memory with a short TTL for hot-path access-control decisions targeting sub-100-millisecond evaluation. Workflow execution runs on a pluggable engine for flows that include approvals, time-bound waits, or external calls. The engine provides checkpointed steps, durable sleep that survives restarts, and typed channel signaling for human-in-the-loop approvals.

Event Pipeline - The platform's most architecturally distinctive component. Events are stored in a lean, standard-agnostic schema (twenty columns plus a JSONB payload) in a TimescaleDB hypertable partitioned by time. Indexes support the high-volume query patterns (time-range, event-type, severity, domain, actor, resource, agent, trace, domain+entity). Read-time enrichment applies the full library of output standards: **OCSF 1.3** (mapped to IAM event classes including Account Change 3001, Authentication 3002, Authorize Session 3003, User Access Management 3005), **MITRE ATT&CK v15** and **MITRE ATLAS, OWASP AOS, IETF SSF/CAEP 1.0**, and **IETF RISC 1.0.** Because standards live in the enrichment layer rather than the schema, new standards can be added without a database migration.

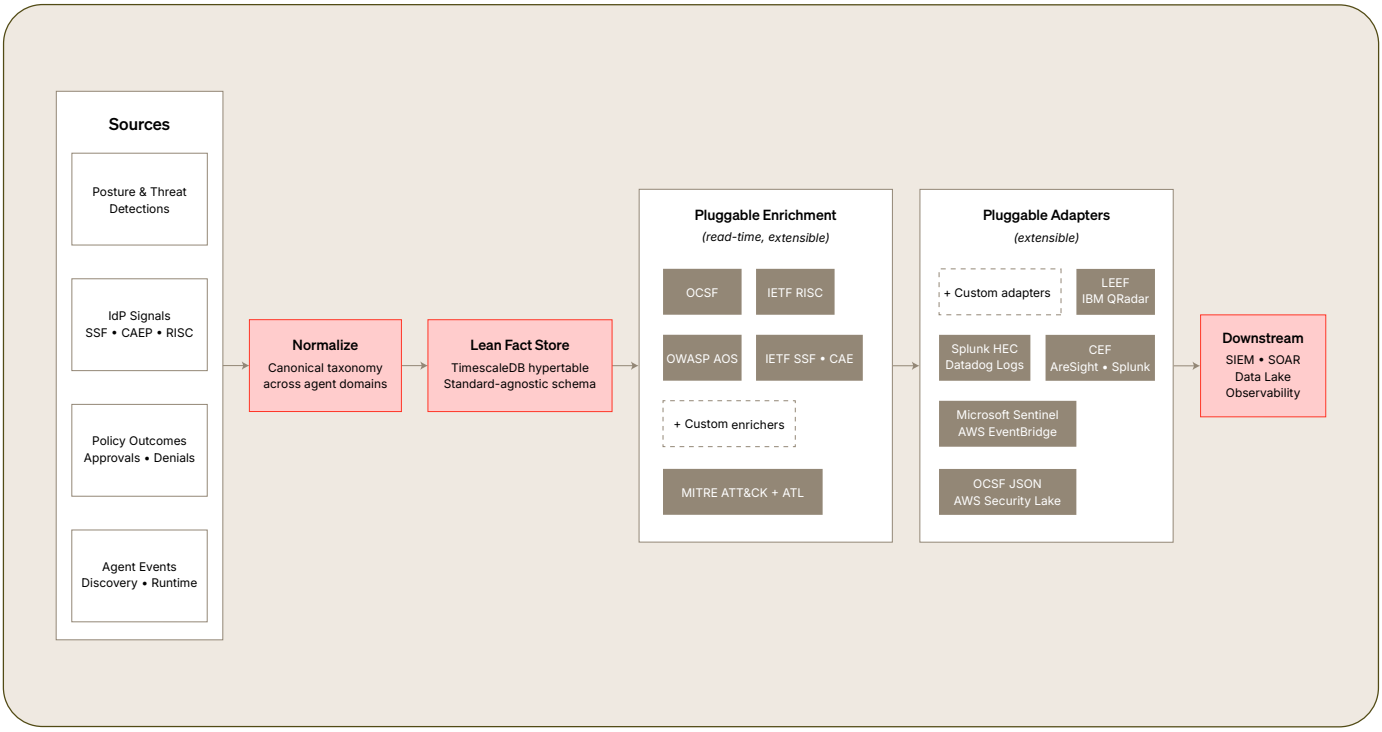


Figure 6. The Agent Identity Security event pipeline. Lean facts are written once; read-time enrichment pipelines apply OCSF, MITRE, AOS, SSF/CAEP, and RISC mappings; and pluggable export adapters emit in the format each downstream tool expects. Both the enrichment library and the adapter library are extensible.

Event taxonomy

Agent Identity Security defines a canonical event taxonomy — **multiple event types across domains** — that covers the operational surface an agent creates:

- **IAM**: User lifecycle, authentication outcomes, credential events, session revocation, account enable/disable.
- **Agent**: Agent lifecycle, runtime MCP calls, runtime execution, tool calls, model changes.
- **Policy**: Policy creation, evaluation, approval decisions, expiration.
- **Workflow**: Workflow lifecycle, step completion, step skip, execution failure.
- **Posture**: Identity posture issue detection, resolution, update.
- **Integration**: Connector lifecycle, sync lifecycle, capability activation.
- **System**: Platform health, configuration changes, migration events.
- **SSF**: Stream creation, event receipt, event processing, connect/disconnect.
- **Resource**: Generic resource events, MCP server discovery, model sanctioning, MCP tool registration, agent-parent creation.

Every event carries W3C Trace Context (trace_id, span_id, parent_span_id) so distributed traces across agent runtime, Agent Identity Security, and downstream systems can be reconstructed end-to-end.

Integration framework

Agent Identity Security operates where the agents operate. The integration framework covers three agent planes, endpoint, SaaS, and infrastructure — each reached by a platform-appropriate connector and, where runtime enforcement is required, an instance of the Agent Identity Security Guardian Agent.

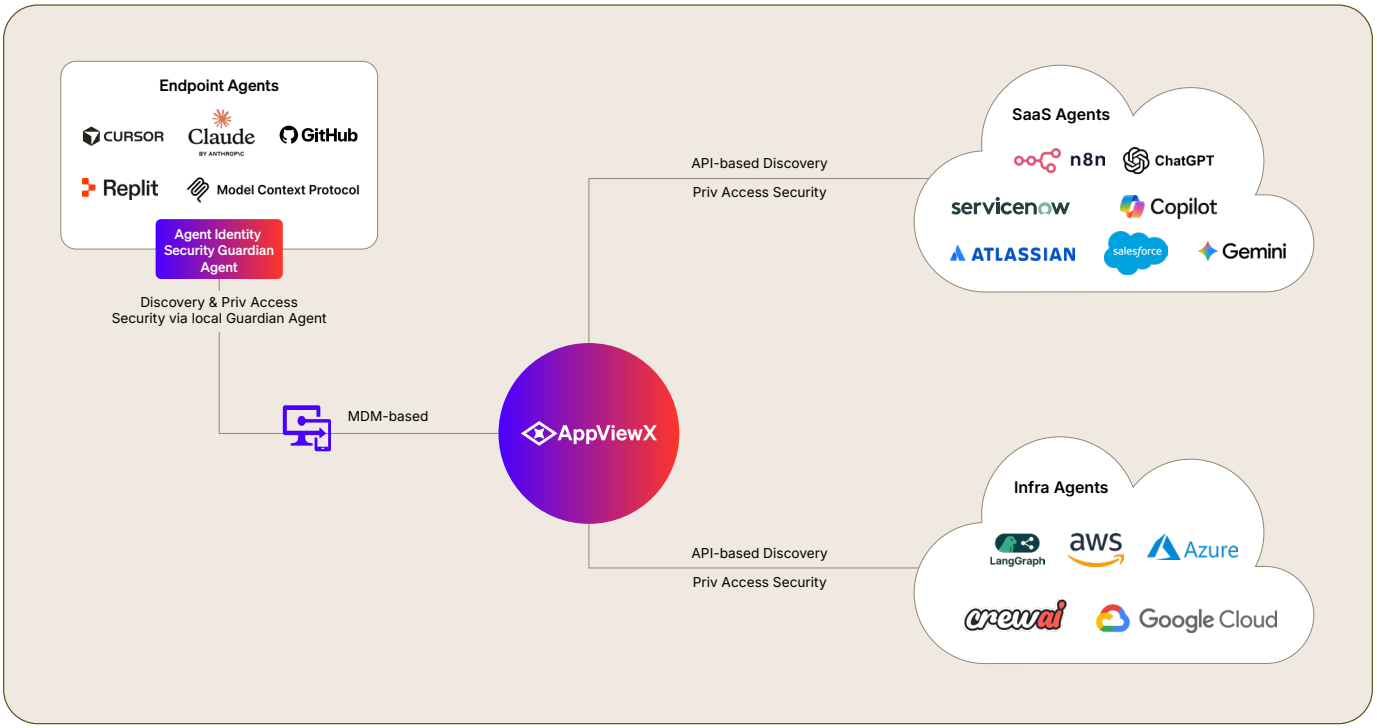


Figure 7. Integration Framework. AppViewX's Agent Identity Security discovers and secures agents across endpoints (coding assistants, MCP clients), SaaS (productivity and business platforms), and infrastructure (agent runtimes and clouds), with the Guardian Agent providing runtime enforcement and endpoint management integration where required.

On the **endpoint** plane, the Agent Identity Security Guardian Agent connects coding assistants (Claude Code, Cursor, GitHub Copilot, Windsurf, Replit) and any MCP clients configured alongside them. Integration with enterprise device management (JAMF, Intune) ensures the agents are discovered as they are deployed.

On the **SaaS** plane, connectors reach platforms such as Microsoft Copilot Studio, Azure AI Foundry, Salesforce Agentforce, ServiceNow AI Control Tower, Atlassian Rovo, OpenAI ChatGPT Enterprise, Anthropic Claude Enterprise, Google Gemini, n8n, Zapier, and related platforms. SIEM and collaboration integrations (Splunk, Microsoft Sentinel, Datadog, IBM QRadar, AWS Security Lake, AWS EventBridge, Slack, Microsoft Teams) close the loop to the SOC and ticketing surfaces (Jira, ServiceNow).

On the **infrastructure** plane, Agent Identity Security connects with agent runtimes and cloud-hosted agent services such as LangGraph, CrewAI, AWS Bedrock AgentCore, Azure AI Foundry, and Google Vertex.

Platform services

Beneath all of the above, a set of platform services maintains the integrity of the system itself:

- **Authentication**: A pluggable authentication framework supports integrating with enterprise identity providers. Access tokens are short-lived JWTs; refresh tokens are frequently rotated along with CSRF protection.
- **Authorization**: Role-based access control with built-in roles and resource-level permissions across resource types. Tenancy is enforced at the database layer.
- **Secrets and Encryption**: Credentials are encrypted at rest with AES-256. The key backend is pluggable across enterprise vault technologies. The platform is designed so that a key-backend swap is a configuration change, not a code change. Transport uses TLS 1.2/1.3 throughout.
- **SCIM 2.0**: SCIM endpoints (RFC 7643/7644-compliant) provide identity-provider-driven user and group lifecycle.
- **Audit**: Every privileged action is recorded in an immutable audit trail indexed for fast compliance queries.
- **Observability**: Tracing via W3C Trace Context on every event; metrics retained in the dashboard metrics store; platform logs available to the administrator.

Deployment

Agent Identity Security is available in two deployment modes:

1. **Self-Hosted**: A single-container deployment with supervisord-managed processes for database, backend, frontend, and supporting services. No external dependencies. Suitable for strict data residency or air-gapped environments.
2. **Multi-Tenant SaaS:** Managed deployment with per-tenant database isolation, per-tenant encryption keys, and a defense-in-depth network architecture.

Summary

AI adoption is not waiting for the identity layer to catch up. Every week that agents proliferate without an identity control plane expands the blast radius, in privilege, in data, and in compliance exposure. The cloud era taught the industry that security teams that are not at the table when adoption decisions are made do not get to design the controls; they inherit the consequences. The AI-agent era is playing out on a compressed timeline, with supply- and demand-side pressure simultaneously, and the window to get in front of it is already narrower than the cloud window was.

The CISOs who will stay ahead of this treat AI agents as a new privileged workforce and defend against rogue agents accordingly — with deep visibility into every agent in the environment, governance and compliance at every layer, and runtime enforcement that stops privilege misuse the moment it happens. That is what Agent Identity Security was built to do.

Sources cited in this whitepaper include:

01. Gartner, 2026 Hype Cycle for Agentic
02. Netscope AI report
03. Axis Intelligence Research
04. Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027
05. Techaisle Research
06. Gartner, Emerging Tech Impact Radar: 2026
07. Cisco Talos' 2025 Year in Review
08. IBM, Cost of a Data Breach Report 2026
09. Ponemon Institute, 2026 Cost of Insider Risks Global Report (2026, 7th edition)
10. NIST AI Risk Management Framework 1.0 (nist.gov/itl/ai-risk-management- framework)
11. NIST AI 600-1 Generative AI Profile; NIST SP 800-53 Rev. 5 (csrc.nist.gov)
12. NIST SP 800-207 Zero Trust Architecture
13. NIST SP 1800-18 PAM for Financial Services; ISO/IEC 27001:2022 and ISO/IEC 27002:2022 (iso.org)
14. ISO/IEC 42001:2023; Regulation (EU) 2024/1689 (eur-lex.europa.eu)
15. OWASP Top 10 for LLM Applications 2025 (genai.owasp.org/llm-top-10)
16. OWASP Agent Observability Standard
17. MITRE ATLAS (atlas.mitre.org)
18. OCSF 1.3 (schema.ocsf.io)
19. OpenID Shared Signals Framework and CAEP 1.0 (openid.net)
20. IETF RFC 9700 OAuth 2.0 BCP
21. IETF RFC 7643/7644 SCIM 2.0
22. Cloud Security Alliance AI Controls Matrix and MAESTRO
23. MCP Specification 2025-06-18 Security Best Practices (modelcontextprotocol.io)
24. Google Threat Intelligence Group reporting on the Salesloft/Drift OAuth compromise (cloud.google.com)
25. Publicly reported analysis of the Amazon Q Developer, Replit, Lenovo Lena, EchoLeak, and CamoLeak incidents

AppViewX

165 Broadway, 23rd Floor,
New York, NY 10006

info@appviewx.com
www.appviewx.com

+1 (206) 207-7541
+44 (0) 203-514-2226