The 2026 Agent Authority Gap Report

Identity, Authorization, and Evidence for Federal AI Agents

Prepared by	NeoXFortress
Edition	2026 Edition (v1.0)
Date	May 2026
Position	Agent Authority Infrastructure™ for regulated and federal AI-agent workflows
Patent posture	U.S. patent pending

NeoXFortress is pioneering Agent Authority Infrastructure™: the identity-bound, mission-scoped authority and evidence layer for AI agents in regulated and federal environments.

CORE THESIS

Identity governance answers who the agent is. Agent Authority Infrastructure™ answers whether a specific agent action should happen right now, under a bounded mission, with required approval and reviewer-ready evidence afterward.

Executive Summary

AI agents are moving from experiments into operational workflows faster than security, identity, and authorization systems are adapting. The market is no longer debating whether agents will matter. Gartner projects that 40% of enterprise applications will include task-specific AI agents by the end of 2026, up from less than 5% in 2025. NIST, NCCoE, CISA, NSA, and allied cyber agencies are now publishing agent-specific work around identity, authorization, secure adoption, least privilege, monitoring, and accountability.

The emerging problem is not only that organizations need to know what AI agents exist. That is the identity problem, and it is real. The deeper production blocker is that security reviewers, federal authorizing officials, CISOs, and prime integrators need to know what each agent was authorized to do, under what mission or business purpose, with what approval posture, what it attempted, what was denied or escalated, and what evidence exists afterward.

This is the Agent Authority Gap.

By "authority" here we mean action-level authorization: whether a specific action should execute right now, under a bounded mission, with the human approval and reviewer-ready evidence that approval bodies will later demand. This is deliberately distinct from an Authorization to Operate (ATO); the two are complementary, not interchangeable. The practical consequence is blunt. In 2026, the constraint on agentic AI is no longer capability. It is approvability.

For federal and regulated environments, this gap matters because AI-agent workflows increasingly touch sensitive data, production systems, internal APIs, operational tooling, and compliance-relevant processes. Reviewers cannot approve what they cannot reconstruct. Security teams cannot defend what they cannot constrain. Program leaders cannot scale what they cannot explain to an authorizing official or risk owner.

NeoXFortress defines Agent Authority Infrastructure™ as the identity-bound, mission-scoped authority and evidence layer for AI agents in regulated and federal environments. The purpose of this report is to frame the market shift, translate current standards and guidance into buyer language, and provide a practical checklist for teams preparing AI-agent workflows for production review.

2026 Public Signals

Signal	What it says	Why it matters
Gartner enterprise-app projection	Task-specific AI agents are moving into mainstream enterprise applications in 2026.	The agent-control problem is moving from theory to production planning.

Signal	What it says	Why it matters
NIST AI Agent Standards Initiative	Secure, interoperable AI agents require standards, identity work, and confidence for adoption.	Federal standards bodies are treating agents as a distinct adoption and security problem.
NCCoE Software and AI Agent Identity and Authorization	AI agents increase the scale and range of actions taken by software systems.	Identity and authorization for software actors is becoming a named federal workstream.
NIST / CAISI AI Agent Security RFI	Agent systems can take actions that affect external state and require security practices around tool use, monitoring, oversight, and constraints.	Runtime action governance is becoming part of the federal policy conversation.
CISA / NSA / Five Eyes allied guidance	Names accountability gaps as a core agentic-AI risk and calls for cryptographically anchored agent identity, ephemeral credentials, and least privilege.	Allied governments are describing the exact authority-and-evidence gap this report names.

1. The Market Signal: Agents Are Becoming Operational Actors

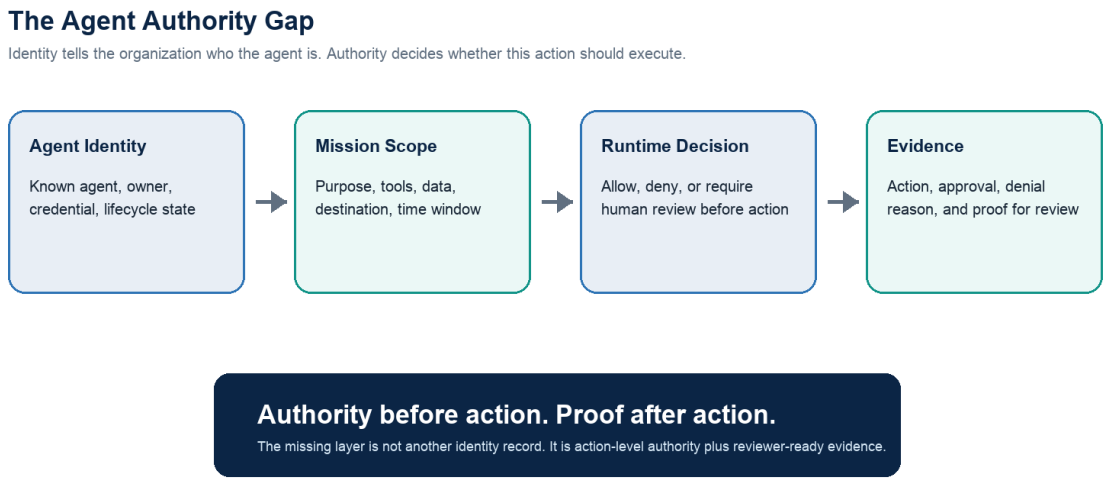

Figure 1. Public conceptual view of the Agent Authority Gap.

The agentic AI conversation has shifted. In 2023 and 2024, most enterprise AI work centered on copilots, chat interfaces, retrieval-augmented generation, and productivity assistants. By 2026, the center of gravity is moving toward systems that can plan, select tools, call APIs, update records, retrieve data, trigger workflows, coordinate with other agents, and operate with varying levels of autonomy.

That shift changes the unit of security risk.

Traditional software access models ask whether a user, service account, or workload can access a system. AI-agent workflows raise a more specific question: should this agent take this action, under this mission, at this moment, with this data, and with this approval posture?

The difference is not semantic. An AI agent may have a valid identity, an active credential, and reviewed access to a system. It may still be unauthorized to close an incident ticket, modify a deployment file, summarize restricted data for an external destination, disable a control, call an administrative API, or continue operating after the mission has expired.

Recent market data reinforces the urgency. Gartner expects task-specific AI agents to become common inside enterprise applications by the end of 2026. In an April 2026 survey of 418 security professionals, the Cloud Security Alliance reported that 82% of enterprises had discovered unknown AI agents in their environments and 65% had experienced AI-agent-related incidents in the prior year. These figures are self-reported and point to a basic reality: AI agents are already appearing inside organizations faster than formal governance systems can account for them.

For regulated and federal environments, the risk is sharper. Agents are not merely generating text. They can affect external state. They can create durable records. They can change system configuration. They can touch sensitive data. They can operate through credentials that look legitimate to downstream systems. Without a dedicated action-authority layer, organizations are left reconstructing intent, approval, and evidence after the fact.

That is not a scalable control model.

2. The Federal Signal: Standards Bodies Are Naming The Problem

Federal and allied institutions are now treating AI agents as a distinct security and governance problem.

In February 2026, NIST launched the AI Agent Standards Initiative to support interoperable and secure AI agents. NIST emphasized that agents capable of autonomous actions need confidence, security, and interoperability to be widely adopted. NIST also identified research in AI-agent security and identity as one of the initiative's pillars.

NCCoE is separately working on Software and AI Agent Identity and Authorization. That project frames the shift clearly: agents can take actions with limited human supervision, increasing the scale and range of actions taken by software systems. The stated focus is applying identity standards and best practices to this new class of digital actor.

In January 2026, NIST's Center for AI Standards and Innovation issued a Federal Register request for information on security considerations for AI agents. The RFI focused on AI agent systems that can take actions affecting external state. It asked about tool use, deployment environments,

multi-agent systems, human oversight controls, least privilege, monitoring, and constraining the environments in which agent actions occur.

In April 2026, CISA, NSA, and their Five Eyes counterparts released joint guidance titled "Careful Adoption of Agentic AI Services," aimed at government, critical infrastructure, defense, and industry stakeholders. Alongside the expected warnings against broad or unrestricted access to sensitive data and critical systems, the guidance is notable for what it names as core risk. It identifies accountability gaps as one of its primary categories of agentic-AI risk, and it calls for agentic systems to use cryptographically anchored identity, ephemeral credentials, least privilege, and existing security models. In other words, allied governments are no longer only asking who the agent is. They are asking whether its actions can be constrained, attributed, and accounted for.

NIST's draft Cybersecurity AI Profile adds another important signal. It recommends that AI systems have unique and traceable identities and credentials, that agent and service identities be bound to credentials using cryptographic mechanisms, and that AI agents be treated with the security precautions normally applied to privileged users.

Read together, the message is plain:

AI agents are no longer only a model governance topic. They are an identity, authorization, runtime control, monitoring, evidence, and accountability topic.

That is exactly where the Agent Authority Gap appears.

3. Why Identity Is Necessary But Not Sufficient

Agent identity is an essential control layer. Organizations need to know what agents exist, who owns them, what credentials they hold, what systems they can reach, whether they are stale or orphaned, and whether their access has been reviewed.

Identity governance will be a major part of the agentic enterprise. Existing identity, non-human identity, privileged access, and lifecycle management vendors are well positioned to help organizations discover and govern agent identities.

But identity alone does not answer the production approval question.

Knowing who the agent is does not prove that a particular action was inside mission scope. Knowing that an agent has a valid credential does not prove that a human approved a sensitive action. Knowing that access was reviewed last quarter does not prove that an action attempted today was allowed under the current business purpose, data boundary, tool boundary, environment, or time window.

Durable access is not runtime authority.

This distinction becomes especially important for agents because they can select tools dynamically, chain actions, operate across multiple systems, spawn sub-agents, and produce outcomes that were not fully enumerated during design review. A conventional identity record can tell the organization that an agent is known. It does not, by itself, answer whether a specific action should happen now.

The production review question is more demanding:

Who or what was the agent? Who owned it? Why was it being used? What mission or business purpose did it support? What was it allowed to do? What did it actually attempt? What was denied or escalated? Who approved sensitive actions? What evidence exists afterward? Can that evidence be reviewed without trusting the agent's own summary?

Those are authority and evidence questions.

Two Adjacent Layers

Identity Governance	Agent Authority Infrastructure™
Establishes who or what the agent is.	Determines whether a specific action should execute.
Manages ownership, lifecycle, credentials, and access posture.	Evaluates mission scope, policy context, approval requirements, and revocation state before action.
Supports discovery, inventory, access review, and lifecycle governance.	Supports allow, deny, or human-review decisions at runtime.
Answers whether the agent is known and governed as an identity.	Answers whether the agent is authorized to act right now under this mission.
Produces identity and access governance records.	Produces reviewer-ready evidence of actions, denials, approvals, and authority decisions.

4. The Agent Authority Gap

The Agent Authority Gap is the space between agent identity and action accountability.

It appears when an organization can identify an agent but cannot prove, at the action level, whether that agent was authorized to do what it attempted.

It appears when the security team can see API calls in logs but cannot reconstruct the mission context, policy evaluation, human approval state, denial reason, or evidence chain that explains why the action was allowed or blocked.

It appears when a prime integrator has a working agentic workflow but cannot hand the ISSO, ISSM, AO, CISO, or risk owner a reviewer-ready evidence pack that makes the control story defensible.

It appears when an agent's credential is valid but the action is outside the mission, or when access is durable but the approval should be temporary.

It appears when the model's reasoning says "do this," but the organization has no separate runtime authority layer deciding whether the action may actually execute.

The Agent Authority Gap is not solved by another dashboard, another chat interface, or another model risk statement. It requires a runtime boundary between agent intent and tool execution. That

boundary must evaluate identity, mission scope, policy, context, human-review requirements, and revocation state before action. It must also preserve evidence after the decision.

The core principle is simple:

PRINCIPLE

Authority before action. Proof after action.

5. Agent Authority Infrastructure™

Agent Authority Infrastructure™ is the runtime authority and evidence layer for AI agents.

It sits between the agent runtime and the tools, APIs, data systems, and workflows the agent touches. It does not replace the identity provider. It does not replace RMF, GRC, SIEM, model monitoring, or AI risk management. It consumes signals from those systems and answers a narrower operational question:

Should this agent action execute right now?

At a public architectural level, Agent Authority Infrastructure should provide five capabilities:

1. **Identity-bound action context.** The system knows which agent is acting, who owns or sponsors it, and what identity or credential state applies.
2. **Mission-scoped authority.** The system evaluates whether the attempted action is inside the approved mission, business purpose, data boundary, tool boundary, environment, and time window.
3. **Runtime decisioning.** The system resolves each governed action as allow, deny, or require human review before execution.
4. **Human approval where needed.** Sensitive or consequential actions can be held for named review instead of being trusted to autonomous execution.
5. **Reviewer-ready evidence.** The system records what was attempted, what decision occurred, what was denied or escalated, who approved, and what evidence exists afterward.

This is not generic "AI governance." It is narrower and more operational. The function is action-level authority and evidence for AI agents in environments where approval, accountability, and auditability matter.

6. The Reviewer Questions That Matter

Federal and regulated buyers do not only ask whether the agent works. They ask whether the agent can be approved, monitored, constrained, and defended after something goes wrong.

The practical review questions are:

1. What agent performed or attempted the action?
2. Who owned, sponsored, or delegated authority to that agent?

3. What mission or business purpose governed the workflow?
4. What tools, systems, data classes, and environments were in scope?
5. What actions were explicitly out of scope?
6. What policy or authority boundary was evaluated before execution?
7. Was the action allowed, denied, or held for human review?
8. If human review was required, who approved or rejected the action?
9. If the action was denied, what was the denial reason?
10. What evidence can a reviewer inspect later without trusting the agent's own summary?

These questions are especially relevant for ISSOs, ISSMs, AOs, RMF practitioners, federal CISOs, CAIO offices, prime integrators, and regulated-enterprise security leaders.

If an AI-agent architecture cannot answer these questions, the organization has an approval problem even if the demo is impressive.

7. Readiness Checklist For Federal AI-Agent Workflows

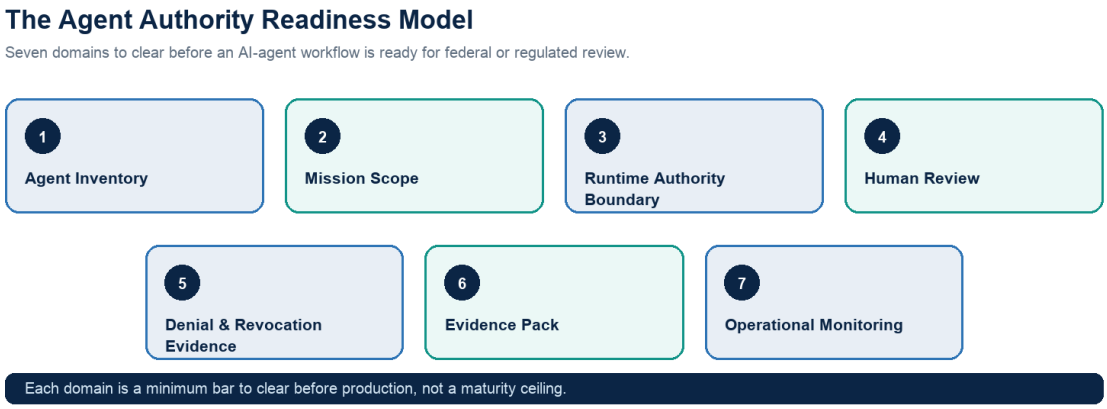

Figure 2. The Agent Authority Readiness Model.

The seven domains below form the Agent Authority Readiness Model: a practical self-assessment for teams preparing an AI-agent workflow for federal or regulated review. It is not a substitute for agency-specific policy, RMF artifacts, legal review, or formal authorization. Treat each domain as a minimum bar to clear before production rather than a maturity ceiling. A workflow that cannot satisfy these has an approval problem regardless of how well it demonstrates.

1. Agent Inventory

- Each agent has a unique identity or stable identifier.
- Each agent has an owner, sponsor, or accountable human authority.
- Each agent's lifecycle state is known: active, stale, revoked, orphaned, experimental, or production-bound.
- Agent credentials are not shared across unrelated workflows.

2. Mission Scope

- Each agent workflow has a defined purpose, scope, owner, and time window.
- The mission identifies allowed tools, allowed data classes, allowed destinations, and allowed environments.
- Out-of-scope actions are documented explicitly.
- Mission expiry or revocation can change runtime behavior.

3. Runtime Authority Boundary

- Agent tool calls or action requests pass through a control point before execution.
- The control point can evaluate identity, mission scope, policy state, context, and approval requirements.
- The control point can allow, deny, or hold for human review.
- The agent runtime cannot silently bypass the authority boundary.

4. Human Review

- Sensitive actions can be paused before execution.
- A named reviewer can approve or reject the held action.
- The approval decision is recorded as part of the evidence record.
- The system distinguishes between autonomous low-risk activity and consequential actions requiring review.

5. Denial And Revocation Evidence

- Denied actions are recorded, not only successful actions.
- Denial reasons are explicit enough for a reviewer to understand the control boundary.
- Revoked, expired, stale, or ownerless agents can be constrained or blocked.
- Subsequent action attempts after revocation generate evidence.

6. Evidence Pack

- The workflow can produce a portable evidence package for review.
- The evidence package includes agent identity, mission scope, runtime decisions, approvals, denials, and verification status.
- Reviewers can inspect evidence without relying only on model-generated summaries.
- Evidence can support RMF, audit, security review, or incident reconstruction conversations.

7. Operational Monitoring

- Security teams can monitor governed actions and attempted policy violations.
- Abnormal action patterns can be investigated with context.
- Logs alone are not treated as sufficient proof of authorization.
- Monitoring aligns with the organization's existing security model and risk posture.

8. Buyer And Stakeholder Map

The Agent Authority Gap is cross-functional. It sits between AI adoption, identity security, cybersecurity operations, federal authorization, and regulated compliance.

The first buyers and influencers are likely to be:

Prime Integrators And Federal Systems Builders

Primes building agentic solutions for agencies need approval paths. They need architectures that can survive security review, not only demos that impress program teams.

Federal CAIO, CDO, And Mission Innovation Offices

AI adoption leaders want agents in production, but they need security cover. They need language and artifacts that help them explain why a workflow is bounded, reviewable, and controllable.

CISOs, ISSOs, ISSMs, And Authorizing Officials

These stakeholders can block, reshape, or bless adoption. They care about authorization, accountability, evidence, review boundaries, and incident reconstruction.

Regulated Enterprise CISOs

Banks, insurers, healthcare organizations, critical infrastructure operators, and defense industrial base companies face a similar production question: what exactly is this agent allowed to do, and how do we prove it?

Identity And Security Platform Vendors

Identity vendors will help organizations discover and govern agents. The complementary layer is runtime action authority and evidence. That makes Agent Authority Infrastructure a potential integration category as the market matures.

9. Implications For 2026

Agentic AI will not be blocked because organizations cannot build agents. The tooling is moving quickly. Agents can already call APIs, manage tickets, query databases, inspect code, and interact with business systems.

The blocker will be approval.

Security teams will ask for evidence. RMF teams will ask what control story applies. CISOs will ask how the agent is constrained. Identity teams will ask how agent credentials are governed. Legal and compliance teams will ask who approved consequential actions. Program leaders will ask why a working pilot cannot move into production.

Organizations that answer these questions early will move faster. Organizations that treat authority and evidence as afterthoughts will keep rediscovering the pilot-to-production gap.

The 2026 adoption race is not only about who builds the most capable agent. It is about who builds the most approvable agent.

That is the case for Agent Authority Infrastructure™.

10. NeoXFortress Position

NeoXFortress is pioneering Agent Authority Infrastructure™: the identity-bound, mission-scoped authority and evidence layer for AI agents in regulated and federal environments.

The NeoXFortress thesis is narrow by design:

PRINCIPLE

Authority before action. Proof after action.

AI agents need identity. They also need runtime authority, human-review boundaries, denied-action evidence, revocation evidence, and reviewer-ready proof.

NeoXFortress has U.S. patent-pending systems in this area and is building toward design-partner evaluation with federal, regulated, and high-trust AI-agent workflows.

Engage

NeoXFortress is opening a limited design-partner program for federal, regulated, and high-trust teams preparing agentic workflows for production review, beginning with reviewer-evidence walkthroughs for qualifying programs. To request a walkthrough or discuss a design-partner evaluation, contact NeoXFortress at info@neoxfortress.com or visit https://www.neoxfortress.com/demo.

Sources

1. NIST, "Announcing the AI Agent Standards Initiative for Interoperable and Secure Innovation," February 17, 2026. Source link

2. NIST NCCoE, "Software and AI Agent Identity and Authorization." Source link

3. NIST / CAISI, "Request for Information Regarding Security Considerations for Artificial Intelligence Agents," Federal Register, January 8, 2026. Source link

4. CISA, NSA, ASD ACSC, CCCS, NCSC-NZ, and NCSC-UK, "Careful Adoption of Agentic AI Services," April 30, 2026. Source link

5. NIST, "Cybersecurity Framework Profile for Artificial Intelligence," NIST IR 8596 Initial Preliminary Draft, December 2025. Source link

6. Cloud Security Alliance, "New Cloud Security Alliance Survey Reveals 82% of Enterprises Have Unknown AI Agents in Their Environments," April 21, 2026 (survey of 418 security professionals, conducted January 2026). Source link

7. Gartner, "Gartner Predicts 40% of Enterprise Apps Will Feature Task-Specific AI Agents by 2026, Up from Less Than 5% in 2025," August 26, 2025. Source link

Disclaimer: This report is provided for informational purposes and reflects publicly available information as of May 2026. It does not constitute legal, compliance, or accreditation advice. Agent Authority Infrastructure™ and Agent Authority Gap are positioning terms used by NeoXFortress. U.S. patent pending.