Futurum
RESEARCH

The Hard(er) Challenge in Agent Governance Is Authorization

How the Economics of Agent Platforms Make Universal Runtime Governance Structurally Unlikely, and What Organizations Should Optimize for Instead

Analyst(s): Fernando Montenegro
Publication Date: June 25, 2026
Document #: AIOFM202606

Key Points

- The Governance Gap Beyond the Control Layer: The launch of Agent Control Standard (ACS) in May 2026 marks an interesting development. What ACS is not designed to address is the accountability chain break in goal-directed agents: authorization exists at the goal level, not the action level, and no runtime enforcement layer can reconstruct what was never recorded.
- Platform Economics as the Fragmentation Driver: Agent catalogs, lifecycle policies, and registry ownership are mechanisms that drive platform stickiness. The adoption dynamic that enabled MCP and A2A, where communication-layer interoperability threatens no one's competitive position, does not apply to governance-layer standards. The odds are structurally against a single control layer achieving universal adoption.
- Shrinkage Over Coverage as the Target: Organizations designing governance programs around universal coverage are optimizing for a destination that platform economics predict will not arrive. The correct target is to reduce ungoverned exposure below an existential threshold. Organizations that internalize shrinkage as success now will govern methodically; those waiting for ecosystem convergence will govern reactively, under regulatory pressure.

Recommendations

1. Reframe Win Criteria Around Shrinkage: Vendors that set accurate expectations, framing success as shrinkage rather than universal coverage, will build stronger, more defensible customer relationships than those that imply complete coverage. Runtime enforcement continuously reduces ungoverned exposure; it does not eliminate it. The buyer who understands shrinkage as success measures correctly and renews; the buyer

sold on complete coverage will be disappointed and will attribute that disappointment to the vendor.

2. Invest in the Accountability Chain: Runtime enforcement at execution checkpoints is necessary and insufficient; the accountability chain problem, not just the control layer, is where a more durable advantage lies. The structural gap between goal-level authorization and action-level auditability in goal-directed agents is what will drive regulatory liability and incident response failures at scale. Vendors building toward prospective, goal-scoped authorization architectures, not exclusively toward retrospective audit trails, are solving the harder, more durable problem and establishing a defensible position ahead of the regulatory wave.

3. Position as the Aggregation Layer: Position above platform registries rather than as a competitor to them. Hyperscaler platform economics guarantees that governance-layer fragmentation across major cloud and SaaS platforms is a permanent condition, not a transitional gap. The enterprise-level cross-platform visibility layer will not be built by a hyperscaler. Vendors that position themselves as the neutral aggregation layer above the platform stack are entering a space where competitive dynamics will ensure that no platform incumbent is structurally positioned to fill it.

What You Need to Know

The Arrival of an Agent Control Standard

The Agent Control Standard (ACS), launched at the AI Agent Security Summit conference in San Francisco on May 27, 2026, proposes standardized middleware hooks that fire at specific agent execution checkpoints: input received, tool call initiated, planning-to-execution transition, memory store, code execution, and sub-agent invocation. At each checkpoint, a policy enforcement layer returns a verdict of allow, deny, or modify before the action reaches production systems.

ACS originated inside Zenity, a startup focusing on AI Security. Michael Bargury, Zenity's co-founder and CTO, is the initiative's co-creator, and ACS was launched at Zenity's event. The project positions itself as vendor-neutral and community-governed and is recruiting outside participation. Still, the project is in its early stage, and its center of gravity today is a single security vendor.

A naming note: Microsoft released its own Agent Control Specification at Build on June 2, 2026, six days after ACS launched, using the same acronym for a module within its Agent Governance Toolkit. Throughout this paper, ACS refers exclusively to the Agent Control Standard coordinated by the agentcontrolstandard.ai community.

Prior to its launch, agent governance relied primarily on platform-native controls and system prompts, mechanisms that operate at the configuration layer rather than at execution. ACS

introduces an inline, pre-action enforcement model that operates across frameworks rather than within a single platform. As of mid-2026, ACS is in public preview, with its reference implementation still roadmapped and the specification under active development. Major agent frameworks, including LangGraph, AutoGen, and the Claude Agent SDK, have not yet shipped native integrations.

Authorization in Goal-Directed Agents

Goal-directed agents sever the clean authorization chain that traditional IAM assumes. A human authorizes a goal; the agent infers what actions are necessary to accomplish it; some of those actions are unexpected. The question "who authorized this specific action?" frequently has no clean answer, because authorization existed only at the goal level. This distinguishes goal-directed agents categorically from RPA platforms, in which every action is known and implicitly authorized at workflow design time. Modern observability tools, including LangSmith, LangFuse, and AgentCore Observability, capture agent reasoning traces alongside actions. What these traces cannot provide is validation against an external authorization record, because no such record exists at the action level.

The Hyperscaler Registry Landscape

The three dominant hyperscalers have each launched agent identity infrastructure in 2026. Microsoft Entra Agent ID, in preview for Copilot Studio agents and enabled per environment, introduces agent-specific Conditional Access policies and identity governance templates that propagate settings across agents of a given class. AWS Agent Registry, part of Amazon Bedrock AgentCore and in preview since April 2026, provides a centralized catalog for agents, tools, and MCP servers with full CloudTrail audit integration. Google's A2A protocol introduces Agent Cards for cross-platform discovery, backed by Agent Engine in Google's Gemini Enterprise Agent Platform (the rebranded Vertex AI) for lifecycle management. Each initiative provides genuine visibility within its platform boundary. The authentication layer is converging: AWS AgentCore supports Entra ID as an authentication provider, and SPIFFE/SPIRE has integration support across the major hyperscalers. The governance layer, covering catalog ownership, lifecycle policy, and audit log location, is following a different trajectory.

MCP and A2A in the Protocol Layer

MCP, with support from Anthropic, OpenAI, Google, and Microsoft, has achieved widespread adoption as the standard for agent-to-tool communication. A2A v1.0 addresses agent-to-agent communication, with Signed Agent Cards providing cryptographically verifiable capability declarations. Both standards address the communication and discovery layers. MCP's architecture creates both an enforcement insertion point for gateway policy and an amplified attack surface for prompt injection through tool responses. A2A's Signed Agent Cards are an integrity mechanism, not an authorization mechanism. They confirm a capability declaration

came from a legitimate source, not that the sub-agent is authorized to exercise those capabilities within a specific trust chain.

The Agent Identity and Governance Market

The agent identity and governance market is active, crowded, and early. It is useful to think about it in three tiers, each approaching the problem from a different starting point.

Numerous specialist vendors are building purpose-built platforms with agent identity and governance as their core problem. Zenity, as a founding contributor to ACS, sits in this tier alongside Aembit, which launched its IAM for Agentic AI platform into general availability in April 2026, focusing on workload identity and secretless credential management for agents; Pillar Security, focused on runtime behavioral protection and guardian agent capabilities; Knostic, focused on knowledge-layer governance and preventing AI oversharing across copilots and agents; and Opal Security, which launched an AI-native identity governance platform in early 2026 anchored by an access evaluation agent.

Established identity providers are extending their platforms to cover agent identity as a first-class identity problem. Okta launched Okta for AI Agents into general availability in April 2026, positioning itself as a provider-neutral governance layer that works across agent ecosystems and identity providers. Saviynt launched an Identity Control Plane for AI Agents in March 2026, extending its identity governance heritage into agentic deployments, following a $700 million round at an approximately $3 billion valuation closed in December 2025. Ping Identity launched Identity for AI into general availability, framing agent governance as a runtime enforcement and continuous verification challenge. SailPoint is extending its identity governance platform to cover non-human identities alongside its existing human identity governance capabilities.

Large security platforms are integrating agent identity governance through existing positions and targeted acquisitions. Palo Alto Networks completed its acquisition of CyberArk in February 2026 for approximately $25 billion, establishing identity security as a core pillar of its platform strategy with agentic identity explicitly in scope. CrowdStrike agreed to acquire SGNL to bring zero-standing-privilege access control into Falcon, with agent identity as an explicit use case. Cisco extended its Duo IAM capabilities at RSA 2026 to cover verified agent identities with time-bound permissions. SentinelOne launched its Singularity Identity portfolio in February 2026, extending its behavioral detection approach to non-human identities and autonomous AI agents. Delinea completed its acquisition of StrongDM in March 2026, combining enterprise PAM with just-in-time runtime authorization to govern privileged access for both human and non-human identities, with agentic AI explicitly in scope.

The market structure reflects the layered nature of the problem. Identity providers are approaching it from the credential and lifecycle layer. Security platforms are approaching it from the detection and enforcement layer. Specialist startups are approaching it from

purpose-built agentic governance architectures. No vendor has credibly claimed the full stack, and the fragmentation of the problem domain maps directly onto the fragmentation of the vendor response. The vendors listed here are representative rather than exhaustive; the specialist tier, in particular, includes dozens of early-stage companies addressing specific sub-problems within the broader agent identity stack.

Analysis

The arrival of ACS marks a notable effort. The industry has identified the control-layer problem and developed an architecturally correct response.

We argue two independent claims, either of which is sufficient to reach the paper's central recommendation. First, even when something such as ACS is fully deployed, it does not close the accountability gap between structural and goal-directed agents. Second, platform economics predict that the governance layer will fragment rather than converge, limiting the deployment of initiatives such as ACS.

Shrinkage follows from either claim on its own: if the accountability gap is real, full coverage still leaves a residual tail; if fragmentation is real, full coverage is structurally unreachable. A skeptic must defeat both claims to undermine the recommendation.

The two claims also anchor different vendor recommendations: the accountability gap argument supports investment in prospective authorization architecture; the fragmentation argument supports positioning as the neutral aggregation layer above platform registries.

ACS Is Architecturally Necessary and Analytically Insufficient

We argue that ACS represents a notable structural advance in agent governance to date. Its emergence as an open, vendor-neutral standard speaks to the kind of industry collaboration the governance problem demands. Its inline, pre-action enforcement model is architecturally correct: it operates at the right layer, fires at the right checkpoints, and returns verdicts before actions reach production systems. The analytical question is not whether ACS is well-designed. It is whether a well-designed control layer resolves the governance problem it was built to address.

It does not, for a reason that the ACS initiative cannot fix as it matures. Runtime enforcement operates on actions the agent is about to take. The accountability chain problem operates one layer upstream: goal-directed agents infer actions from goals, and authorization at the goal level does not propagate to the action level in any artifact ACS can inspect.

Not all agents are equal. Many agentic workflows in production today are closer to sophisticated RPA than to truly goal-directed agents: bounded, largely deterministic, with action graphs that are knowable at design time. For these deployments, existing accountability

models largely hold. The governance challenge concerns genuinely goal-directed agents, where the same goal grant can yield different action sequences depending on what the agent encounters at runtime.

Goal-scoped authorization reduces this gap meaningfully. If a goal is defined precisely enough, a complete action log allows retrospective assessment of whether individual actions fell within scope. The limit is that genuinely open-ended goals cannot be scoped precisely enough for that assessment to be deterministic, and the evaluation remains retrospective and judgmental rather than prospective and structural. ACS governs execution within whatever authorization scope exists; it cannot substitute for the authorization record that goal-directed architecture never creates.

Initiatives such as ACS cannot determine whether a given action was within the scope of the original human authorization, because that authorization was never recorded at the action level. This is why shrinkage, reducing ungoverned exposure below an existential threshold rather than eliminating it, is the correct organizational target. Governance programs designed for the least capable agents in the estate will be wrong when the most capable ones cause harm.

Platform Economics Predict Fragmentation: This Is Not a Maturity Gap

The conventional wisdom is that governance standards follow protocol standards with a delay. MCP and A2A achieved broad adoption, and the thinking is that ACS could follow the same trajectory as enterprise pressure builds. We argue this misreads why MCP and A2A succeeded.

Both standards address the communication and discovery layer, where interoperability benefits all participants without any platform surrendering competitive control. The governance layer operates under entirely different economics. Agent catalogs, lifecycle policies, identity blueprints, and audit surfaces are mechanisms that drive platform stickiness, and no platform will voluntarily make a competitor's infrastructure load-bearing in its own stack.

Major SaaS and cloud platforms, of which Salesforce, AWS, and Microsoft are illustrative rather than exhaustive, each have strong structural incentives to maintain proprietary governance control regardless of how authentication converges. The authentication layer converges because it answers a question with no competitive dimension. The governance layer fragments because owning the answer is the business model.

ACS adoption is a live test of this prediction. The leading indicator is whether major agent frameworks ship first-class ACS hooks natively, or whether integration remains a manual wrapper that enterprises must maintain themselves.

One force could override platform competitive incentives: regulatory mandate. SAML and OAuth achieved universal governance-adjacent adoption precisely because compliance

requirements removed optionality. If financial services regulators require agent governance interoperability as a condition of compliance, the fragmentation prediction weakens. Until that mandate exists and is enforced, the competitive economics holds.

One reading of ACS places it closer to MCP than to a governance registry: a hook-and-wire protocol that defines how policy decisions are requested and returned, rather than what those decisions are or who owns them. Under that reading, our own MCP analogy points toward adoption rather than fragmentation.

We argue the analogy does not hold. MCP defines how an agent communicates with a tool. ACS hooks fire at decision points that determine what an agent is permitted to do: whether a tool call proceeds, whether a sub-agent is invoked, and whether a memory write is allowed.

Those are governance decisions with business consequences, not communication events. Authorization decisions carry context-dependence, liability implications, and organizational ownership questions that communication formats do not. A platform ceding control over those decision points surrenders something competitively meaningful. A platform ceding control over a communication format does not.

The clearest evidence arrived six days after ACS launched. At Build 2026, Microsoft shipped its own Agent Control Specification, a same-named open control spec inside its Agent Governance Toolkit, with a different set of interception points, a different verdict vocabulary, and named enterprise customers and partners behind it. Two open standards for the same layer, sharing an acronym and incompatibility in the details, appeared within a week. That is the fragmentation this section predicts, surfacing before adoption even begins.

The Governance Signal Will Precede the Governance Capability

We expect agent IAM governance programs to proliferate rapidly over the next 18 months as regulatory and peer pressure create institutional signals. We expect the majority to reflect genuine intent, even as they produce uneven security outcomes. This is a structural observation, not a critique of organizational effort.

Governance programs that spread through normative channels are calibrated to satisfy those channels. Regulators and auditors can verify the completeness of documentation, registration counts, and certification labels. They cannot easily verify whether the authorization scope is appropriate or whether oversight frequency is commensurate with the blast radius. The measure and the outcome it was meant to proxy will diverge under that pressure.

The organizations that invest deeply in governance capability may be indistinguishable from those that have met the visible requirements, until an agent incident reveals the difference. That correction will come. The question is how far organizations allow the gap between their governance posture and their governance capability to go before it does.

The Aggregation Layer Is the Durable Market Opportunity

We see a more durable market opportunity that platform economics point to: the enterprise-level governance layer that aggregates across fragmented hyperscaler registries. Most enterprises operate across multiple cloud and SaaS platforms simultaneously. Each delivers genuine governance within its own boundary. None has the structural incentive to aggregate across the others.

The CISO who needs to answer "what agents do we have, what can they do, and are they all governed?" cannot answer that question from any single admin console. That gap will not be filled by a hyperscaler on neutral terms. Microsoft, through Entra Agent ID and Agent 365, is the most credible exception to test, being the one platform with a plausible aggregation play across third-party agents. The structural limit remains: Microsoft's aggregation serves Microsoft's ecosystem. Salesforce agents, agentic flows within applications, and custom open-source deployments are unlikely to be governed by Entra on terms that benefit Microsoft's competitors.

The structural parallel is SSPM. No single SaaS platform would build cross-platform visibility into a competitor's configuration, so a distinct vendor category emerged to fill the gap. That category has faced consolidation pressure as larger platforms absorb SSPM capabilities, which is itself instructive. The aggregation opportunity is real; capturing it as a durable independent category requires moving before platform consolidation closes the window.

Vendors competing within the platform governance layer face incumbents with distribution advantages and native integration. Vendors positioning themselves as the aggregation layer above it are entering a space that incumbents are structurally unable to fill.

Shrinkage Is the Correct Target: Accumulation Is the Correct Urgency Frame

The absence of a hard regulatory deadline makes deferral easy to rationalize. Unlike post-quantum cryptography, there is no NIST standards deadline, no NSA migration guidance creating a forcing function. The urgency is real regardless, and it compounds.

Every month of ungoverned agent deployment adds governance debt. Identities go uninventoried. Authorization scopes go undocumented. Procurement contracts are signed without disclosure requirements. That debt does not retire itself. It accumulates until a governance program must be built retroactively, which is always more expensive than building it prospectively.

We borrow the frame from retail loss prevention: shrinkage. Retailers do not expect to eliminate theft and inventory loss entirely. They invest to reduce it below the threshold where it becomes existential to the business, and they manage it continuously. Agent governance

deserves the same frame. Permanent residual ungoverned exposure is the honest expectation. The goal is reduction below an existential threshold, not elimination. A practical starting point: can you reconstruct the authorization chain for your highest-blast-radius agents, do you have a content trust model for agents processing external content, and do your procurement contracts require disclosure of agent capabilities? Organizations that cannot answer yes to all three have not yet reached the threshold.

The third-party agent market today lacks an Agent Bill of Materials (ABOM), the capability-disclosure artifact analogous to an SBOM. Without it, buyers cannot evaluate an agent's real capabilities at procurement time, so the market exhibits Akerlof information asymmetry: vendors of genuinely well-scoped agents cannot credibly signal that above vendors who merely assert it. Procurement disclosure requirements are the buyer-side forcing function that starts to close that gap. Shrinkage framing is a management discipline, not a license for inaction. The goal is a continuously shrinking ungoverned perimeter, with explicit ownership of what remains ungoverned and the reasons why.

Organizations building governance programs now do so methodically, before external scrutiny compresses the timeline. Those that defer will build under regulatory pressure, competing for a thin pool of expertise, while simultaneously discovering the scope of what should have been inventoried earlier.

The correct investment target is the continuous reduction of ungoverned exposure. Explicit governance design for the highest-risk agent classes, honest accounting for the ungoverned tail that will persist, and vendor relationships built around that honest frame will prove more durable than programs built around promises of completeness that the market structure cannot deliver.

What to Watch

- Will major agent frameworks support ACS? The fragmentation thesis hinges on whether platforms treat ACS as a communication-layer standard worth supporting or a governance-layer standard worth resisting. Native integration in major frameworks within 12 months would be meaningful evidence against it. Continued manual wrappers would confirm it.

- How quickly will ACS evolve? The standard is in public preview with its reference implementation still roadmapped. As implementation experience accumulates, the specification is likely to grow in both the number of hooks covered and policy decisions expressible. That evolution will indicate whether ACS matures into a comprehensive governance instrument or remains a foundational but partial control layer.

- Will a neutral aggregation layer emerge as a distinct vendor category? The SSPM parallel predicts that a specialist category will emerge above the hyperscaler registry

layer. Watch for vendors explicitly positioning themselves as cross-platform governance aggregators. Early category definition will carry disproportionate market influence.

- Which sector regulator moves first on agent governance requirements? Financial services are the highest-probability first mover. Watch for SEC, OCC, or FFIEC guidance extending model risk management frameworks to goal-directed agents. The form that guidance takes, principles-based or prescriptive, will shape governance program design across all sectors.
- Will liability for agent-mediated harm shift toward platform vendors? Autonomous vehicle precedents suggest deployer liability will eventually extend upstream toward platform vendors. Watch for early litigation or enforcement actions that name platform vendors alongside deploying organizations. That shift changes the economics of governance investment at the platform layer.

Other Insights from Futurum

RSAC 2026: The AI 'Tragedy of the Commons' and the Future of Agentic Security

Can Zscaler and Its GSI Partners Govern the Agentic Enterprise? - Futurum

Does Cisco Put an Astrix on the Agentic Identity Race? - Futurum

Cisco Live 2026: Platform, Silicon, and Security for the Agentic Era - Futurum

About Us

About the Authors

Fernando Montenegro is the Vice President and Practice Lead for Cybersecurity at The Futurum Group. In this role, he leads the development and execution of the Cybersecurity research agenda, working closely with the team to drive research for clients. His background is centered on enterprise cybersecurity and includes diverse roles in industry research, customer support, security, IT operations, professional services, and sales engineering. Fernando can be reached at fmontenegro@futurumgroup.com.

About The Futurum Group

Every day, The Futurum Group's analysts, researchers, and advisors help business leaders worldwide anticipate tectonic shifts in their industries and leverage disruptive innovation. Unlike traditional analysts, The Futurum Group works not only in analysis and research but also takes that insight and knowledge even further, engaging all the way through the go-to-market process.

Futurum Research provides in-depth research and insights on global technology markets using advisory services, custom research reports, strategic consulting engagements, digital events, go-to-market planning, and message testing. It also creates, distributes, and amplifies rich media content that all stakeholders read, watch, and listen to.

See more details on The Futurum Group at futurumgroup.com.

Copyright & Use License