Public Comment on NIST NCCoE Concept Paper

"Accelerating the Adoption of Software and AI Agent Identity and Authorization"

Submitted by: Jamie Thompson, CEO, Sprinklenet **Date:** March 28, 2026

Dear NIST NCCoE Team,

Thank you for the opportunity to comment on this concept paper. The following observations are informed by our experience building governed AI systems for enterprise and government use cases, where identity, authorization, and auditability are foundational operational requirements.

1. The Governance Configuration Layer

The concept paper thoroughly addresses identity and authorization at the agent level: how agents authenticate, how they prove authority to act, and how their actions are logged. However, we believe the paper would benefit from addressing an additional architectural concept: the **governance configuration layer** that sits between human operators and the AI agents they deploy.

Organizations do not deploy AI agents in isolation. They deploy agents through platforms that manage which foundation models an agent can access, what data sources it can retrieve from, what guardrails constrain its outputs, and what audit events are captured. This configuration layer is where identity and authorization policies are *defined and enforced* before an agent ever takes an action.

We recommend the NCCoE consider how identity and authorization standards apply not only to the agents themselves but to the **orchestration and governance infrastructure** through which agents are configured, deployed, and monitored. This includes:

- **Model routing governance.** When an agent can access multiple foundation models (e.g., for different tasks or data sensitivity levels), the configuration layer must enforce which models are authorized for which use cases, and those routing decisions must be auditable.
- **Retrieval authorization.** In RAG architectures, agents access organizational knowledge bases. The authorization question is not only "can this agent access this tool?" but "can this agent retrieve from this specific document collection, and can the end user see the retrieved sources?" Multi-tenant environments require retrieval authorization at the knowledge-domain level, not just the agent level.
- **Guardrail policy as an authorization construct.** Output constraints (content moderation, PII detection, allowed topics, disallowed responses) function as a form of authorization. They define what an agent is *permitted to produce*, distinct from what data it is permitted to access.

2. LLM Agnosticism and Identity Standards

The concept paper references specific protocols (MCP, OAuth 2.0, OIDC, SPIFFE) for agent identity and authorization. We support this direction and would add that identity standards for AI agents must account for **model-agnostic architectures** and for the broadening range of **agent interaction patterns**.

On interaction patterns: agent integration is expanding beyond API-based protocols to include CLI-based workflows, where AI agents operate directly within a user's terminal environment and inherit the user's shell credentials, file system access, and API keys. These patterns are in active production use today and introduce identity and authorization challenges distinct from API-driven architectures. Standards should account for both.

On model agnosticism: organizations increasingly route agent workloads across multiple foundation models, selecting models based on task requirements, cost, data residency constraints, or performance characteristics. An agent's identity and authorization profile should persist regardless of which underlying model executes a given request. This introduces several challenges:

- **Model-level authorization.** Different models may have different security postures (commercial API vs. on-premises vs. air-gapped). The authorization framework should support policies that restrict which models an agent can invoke based on data or task sensitivity.
- **Cross-model audit continuity.** When an agent routes a request from one model to another (e.g., for multi-step reasoning), the audit trail must maintain a consistent thread of identity and authorization decisions across model boundaries.
- **Bring Your Own Key (BYOK) implications.** In enterprise deployments where organizations use their own API keys, the identity framework should distinguish between the agent's identity, the operator's identity, and the model access credential, as these are three distinct layers that converge in a single request.

3. Role-Based Access Control for Agent-Mediated Knowledge

The concept paper discusses authorization in terms of what actions agents can take and what resources they can access. We recommend expanding this to include **role-based access control (RBAC) applied to knowledge retrieval and response generation**.

In governed AI deployments, access control is not binary. It is granular:

- **Document-level RBAC.** Different users interacting with the same agent may be authorized to retrieve from different document collections. An agent serving both internal staff and external partners must enforce retrieval boundaries based on the requesting user's role and clearance.
- **Response-level governance.** Beyond controlling what data an agent can access, organizations need to control what an agent is authorized to *say*. This includes topic restrictions, disclaimer requirements, and output format constraints that vary by user role or organizational context.
- **Sharing controls with non-repudiation.** When an organization shares an AI-powered experience with external parties (e.g., a public-facing AI assistant), the authorization framework must support read-only access patterns where external users can interact with the agent but cannot access underlying source documents. Every interaction must be logged with full attribution.

4. Auditing at the Governance Layer

The concept paper's discussion of auditing and non-repudiation (Section 5) focuses appropriately on logging agent actions. We recommend extending this to include

auditing of **governance configuration changes**: the decisions made by human operators about how agents are configured.

The most consequential authorization decisions are not made at runtime by the agent. They are made at configuration time by the humans who define the agent's permissions, guardrails, and model routing rules. A complete audit framework should capture:

- Who configured the agent's access permissions and when
- What guardrail rules were active at the time of each interaction
- Which model routing policies were in effect
- Changes to any of the above over time, with attribution to the human operator who made the change

This "governance audit trail" is distinct from the "agent action audit trail" and is equally important for non-repudiation, compliance reviews, and incident investigation.

5. Prompt Injection in the Context of Governed Platforms

The concept paper addresses prompt injection prevention (Section 6). We would note that governed AI platforms provide an additional layer of defense not currently addressed: **platform-level input and output filtering that operates independently of the foundation model**.

When guardrail engines (PII detectors, content classifiers, prompt injection detectors) operate at the platform layer rather than the model layer, they provide defense-in-depth that persists regardless of which foundation model processes the request. This is relevant because:

- Guardrail policies are configured and authorized by human operators through the governance layer
- Guardrail enforcement generates its own audit events (e.g., "PII detected and redacted," "prompt injection attempt blocked")
- The guardrail layer's identity and authorization are separate from the agent's identity: it is infrastructure, not an agent, but it mediates agent behavior

We suggest the NCCoE consider how platform-level guardrails fit into the identity and authorization framework, particularly in the context of the "trust domain" concept depicted in Figure 1.

Our primary recommendation is that the NCCoE expand the scope to include the **governance and orchestration layer** through which organizations configure, deploy, and monitor AI agents. Identity, authorization, and auditing at this layer are essential for the practical adoption of agentic AI in federal and enterprise environments. We welcome the opportunity to participate in any subsequent NCCoE demonstration project or collaborative activity.

Respectfully,

Jamie Thompson
Chief Executive Officer, Sprinklenet
jamie@sprinklenetlabs.com https://sprinklenet.com