

Agent ID: Declarative Ontological Layer for Artificial Agents

Version 3.0 — Dense Edition

Abstract

Agent ID defines a minimal declarative identity layer for artificial agents. It does not implement authentication, authorization, governance, or enforcement. Instead, it defines the ontological conditions under which an artificial agent may be treated as a persistent socio-technical actor. This paper formalizes the architectural need, conceptual model, structural separation principles, and interoperability implications for large-scale autonomous systems.

1. The Structural Problem of Agent Identity

As AI agents evolve from simple assistants to autonomous systems capable of generating sub-agents, acting across trust domains, and executing high-frequency decisions, identity becomes structurally ambiguous. Existing IAM systems answer session-level questions (Who is authenticated? What permissions are granted?) but fail to address the ontological question of persistent agent existence.

2. Ontology vs Authentication

Authentication verifies a session. Authorization grants access. Governance regulates behavior. Agent ID precedes all three. It defines what exists before any session or permission is evaluated. Without a stable referent, delegation chains collapse into implementation artifacts.

Layer Comparison

Layer	Function
Agent ID (Ontology)	Defines existence
Authentication	Verifies session
Authorization	Grants permissions
Governance	Regulates behavior

3. Delegated Responsibility Graph Model

Each agent A is defined by a tuple: $A = (ID, Principal, Scope)$. Delegation creates a directed graph G where edges represent authority transfer. Agent ID does not evaluate moral validity or correctness; it records declared edges within a responsibility graph. This enables traceability without enforcement.

4. Threat Model Implications

Without a declarative layer, recursive agent chains obscure attribution. Autonomous spawning, cross-domain interactions, and ephemeral instances require persistent identifiers independent of implementation lifecycle. Agent ID separates meaning from proof and identity from execution.

5. Interoperability Considerations

Agent ID complements OpenID, OAuth, Entra Agent Identity, and enterprise IAM systems. Where IAM secures sessions, Agent ID secures referential continuity. It enables interoperability by stabilizing identity semantics across domains.

6. Design Principles

- 1 Minimalism — identity only.
- 2 Non-enforcement — no governance logic.
- 3 Separation of Meaning and Proof.
- 4 Principal Attribution Required.
- 5 Persistence Across Contexts.

7. Strategic Positioning

Agent ID is infrastructure, not a platform. It provides a neutral declarative substrate enabling future autonomous systems to scale without dissolving accountability structures.

Agent ID v3.0 Dense Edition — Maintained by aiagentid.org