Draft AI Agent Standards Coverage Matrix

Full coverage landscape across W3C ANP, NIST, IEEE P3945, MCP, A2A, and ACP — what is defined, what is partial, and what no standard yet addresses. See Capability Definitions and References at the end of the Matrix.

Capability	W3C ANP	NIST	IEEE P3945	MCP / A2A / ACP
Identity & Authentication				
Decentralized Identifiers (DIDs / DID:WBA)	Yes	No	N/A	No
Verifiable Credentials (VCs)	Yes	No	N/A	No
DID Mutual Authentication	Yes	No	N/A	No
OAuth 2.0 / 2.1 + OpenID Connect	Partial	Yes	N/A	Yes
Agent-Specific OAuth Grant Types	No	Yes	N/A	Partial
SPIFFE / SPIRE Workload Identity	No	Yes	N/A	No
Mutual TLS (mTLS)	No	Partial	N/A	Yes
Agent as Named Enterprise Identity (IAM)	No	Yes	No	No
Discovery & Capability Description				
Agent Description Document (JSON-LD / /.well-known)	Yes	No	Drafted	Yes
Open Agent Schema Framework (OASF)	No	No	No	Yes
Registry / Index-Based Discovery	Partial	No	Drafted	No
Dynamic Discovery (unknown agent)	Partial	No	Drafted	No
Presence / Availability Signaling	No	No	No	No
Capability Versioning / Schema Evolution	No	No	No	No
Architecture & Agent Models				

Capability	W3C ANP	NIST	IEEE P3945	MCP / A2A / ACP
Industrial Agent Architecture (IEC 62264 Mapping)	N/A	N/A	Yes	N/A
Generic Agent Role Categories	N/A	N/A	Yes	N/A
Agent Interaction Patterns / Data-Flow Models	N/A	N/A	Yes	N/A
End / Edge / Cloud Deployment Topology	N/A	Partial	Yes	N/A
Conformance / Certification Testing	No	No	No	No
Communication & Protocols				
Meta-Protocol Negotiation	Yes	N/A	N/A	No
Messaging Profiles (P1–P9)	Yes	N/A	N/A	N/A
Agent Communication Meta-Protocol (Profile 06)	Drafted	N/A	N/A	N/A
Agent Payment Protocol (Profile 10)	Drafted	N/A	N/A	No
Agent-to-Tool Protocol	N/A	N/A	Drafted	Yes
Agent-to-Agent Protocol	Partial	N/A	Drafted	Yes
Task Lifecycle State Machine	No	N/A	Drafted	Yes
Streaming / Async Messaging (SSE / Webhooks)	No	N/A	No	Yes
Multipart / Multimodal Payloads (MIME)	N/A	N/A	N/A	Yes
Cross-Protocol Interoperability	Partial	N/A	No	No
Error Handling / Fault Recovery Protocols	No	No	No	Partial
Agent Suspension / Resume	No	No	No	No
Multi-Party Coordination (3+ agents, fan-out)	No	N/A	No	No
Access Control & Security				

Capability	W3C ANP	NIST	IEEE P3945	MCP / A2A / ACP
Service-Level Authorization (OAuth Scopes)	No	Yes	N/A	Yes
Least-Privilege / Just-in-Time Access	No	Yes	N/A	No
Human Approval Gates	No	Yes	N/A	No
Tool Behavior Annotations (readOnly / destructive)	N/A	No	N/A	Yes
Excessive Agency Taxonomy	N/A	Yes	N/A	N/A
Required Audit Trail (6 mandatory fields)	No	Yes	N/A	No
Agentic Attack Surface Classification	N/A	Yes	N/A	N/A
Prompt Injection Defense (protocol-level)	No	Partial	N/A	No
Real-Time Inference Monitoring	N/A	Partial	N/A	N/A
Agent Lifecycle / Decommissioning	No	No	No	No
Supply Chain Security (agent provenance)	No	Partial	N/A	No
Privacy / Data Retention at Handoff	No	No	N/A	No
Interoperability & Governance				
Benchmarking Framework	N/A	N/A	Drafted	N/A
Memory Interoperability	Partial	No	No	No
Sector-Specific Controls (HIPAA, SOC2, etc.)	N/A	Partial	Drafted	N/A
Regulatory Compliance Attestation	No	No	No	No
Per-Handoff Envelope — No Standard Addresses These				
Conditions / Prerequisites at Handoff	No	No	No	No
Per-Handoff Constraint Declaration	No	No	No	No

Capability	W3C ANP	NIST	IEEE P3945	MCP / A2A / ACP
Boundary Declarations (what receiver may NOT do)	No	No	No	No
Delegation Chain at Message Level	Partial	No	No	Partial
Sealed Provenance Chain (hop-by-hop)	No	No	No	No
On-Violation Behavior Declaration	No	No	No	No
Mandatory Receiver Inspection	No	No	No	No
Rate Limiting / Resource Quotas per Handoff	No	No	N/A	No
Content / Harm Scoring Layer	No	No	N/A	No

Key: **Yes** = fully defined **Drafted** = formal draft / PAR-approved WG **Partial** = deployed but incomplete **No** = not addressed **N/A** = outside scope by design

Capability Definitions

Industrial Agent Architecture (IEC 62264 Mapping) — The IEEE P3945 parent standard's framework for mapping IEC 62264-compliant industrial devices (manufacturing systems, PLCs, SCADA) to interoperable intelligent agents. Covers end-node, edge, and cloud deployment tiers. The architectural foundation on which P3945.1 and P3945.2 are built.

Generic Agent Role Categories — IEEE P3945's taxonomy of standardized agent roles within an industrial system — producer, consumer, coordinator, broker, and monitor — enabling unambiguous assignment of responsibility and authority during multi-agent task execution.

Agent Interaction Patterns / Data-Flow Models — IEEE P3945's specification of canonical interaction patterns (request/response, publish/subscribe, delegation, negotiation) and the data-flow models governing how information moves between agents, tools, and devices in an industrial deployment.

End / Edge / Cloud Deployment Topology — A defined architecture specifying how agents are distributed across end-node devices, edge compute, and cloud tiers — with clear rules about which functions may run where, what latency and connectivity assumptions apply, and how agents in different tiers communicate.

Conformance / Certification Testing — A standardized test suite and process for certifying that an implementation correctly implements a protocol or standard. Distinct from benchmarking (measuring capability): conformance testing gives a binary pass/fail against a specification, enabling interoperability guarantees between independently built agents.

Decentralized Identifiers (DIDs / DID:WBA) — A W3C standard for globally unique identifiers that require no centralized registration authority. The DID:WBA method anchors agent identity to HTTPS URLs, allowing any agent to publish a cryptographically verifiable identity document without relying on an OAuth server or third-party provider.

Verifiable Credentials (VCs) — W3C VC Data Model v2.0 standard for cryptographically signed digital attestations. Enables agents to present tamper-evident credentials (e.g., "this agent is authorized to act on behalf of Acme Corp") that any party can independently verify without contacting the issuer.

DID Mutual Authentication — A handshake in which two agents each verify the other's DID-based identity by exchanging and validating cryptographic signatures, eliminating the need for an OAuth intermediary between peer agents.

OAuth 2.0 / 2.1 + OpenID Connect — Industry-standard authorization and authentication frameworks adapted for agents. OAuth 2.1 consolidates security best practices from RFC 6749; OpenID Connect adds identity assertion. NIST and all three industry protocols (MCP, A2A, ACP) rely on this stack.

Agent-Specific OAuth Grant Types — Extensions to OAuth 2.0 for AI agent authorization flows — including delegated authority from a human principal, task-scoped tokens, and machine-to-machine flows without an interactive user session.

SPIFFE / SPIRE Workload Identity — Secure Production Identity Framework for Everyone (SPIFFE) defines a standard for workload identity; SPIRE is its reference implementation. NIST maps these to AI agents as the preferred cryptographic identity mechanism for enterprise cloud and on-premises workloads.

Mutual TLS (mTLS) — Transport Layer Security with certificate authentication in both directions. Each party proves its identity via X.509 certificate before the connection is established. Used by A2A as a transport-level authentication option.

Agent Description Document (JSON-LD / /.well-known) — A structured document published at a well-known HTTPS URL (RFC 8615) declaring an agent's capabilities, interfaces, and interaction requirements. W3C uses JSON-LD + Schema.org; A2A uses JSON Agent Cards at /.well-known/agent.json.

Open Agent Schema Framework (OASF) — An OCI-based, protocol-agnostic data model developed by IBM/AGNTCY for describing agent attributes, capabilities, and interfaces. Designed to bridge MCP, A2A, and ACP so agents can be described once and addressed by any protocol.

Registry / Index-Based Discovery — A mechanism by which agents register descriptions in a shared registry or index, enabling other agents to find them without prior knowledge of their URL. Analogous to DNS or a search engine for agents. IEEE P3931 addresses this at a standards level.

Dynamic Discovery (unknown agent) — The ability for an agent to find and contact another agent it has no prior knowledge of, by querying a registry or network index. Distinct from URL-based discovery, where the address must already be known. No current production protocol supports this.

Presence / Availability Signaling — Real-time notification of whether an agent is online, busy, or accepting requests — analogous to presence indicators in messaging systems. Not addressed by any current standard or deployed protocol.

Capability Versioning / Schema Evolution — A mechanism for agents to signal changes to their capabilities over time — including backward-compatible additions, breaking changes, and deprecations — and for calling agents to negotiate which version to use. Without this, a capability description valid when a service relationship was established may silently become incorrect as the agent evolves.

Meta-Protocol Negotiation — W3C ANP's mechanism allowing two agents to agree dynamically on which application-layer protocol (A2A, ACP, or custom) they will use for a specific interaction. Enables self-organization without pre-configuration.

Messaging Profiles (P1–P9) — A set of nine W3C ANP interaction profiles covering common agent communication patterns from simple request/response to multi-party coordination.

Agent Communication Meta-Protocol (Profile 06) — A W3C ANP draft protocol (profile 06) defining how agents negotiate communication frameworks and establish sessions. PAR-approved working group; not yet ratified.

Agent Payment Protocol (Profile 10) — A W3C ANP draft protocol (profile 10) for agent-to-agent payment initiation and settlement, enabling agents to transact autonomously. PAR-approved working group; not yet ratified.

Agent-to-Tool Protocol — A protocol governing how an AI agent invokes external tools, APIs, or data sources. MCP (Model Context Protocol) is the dominant deployed implementation using JSON-RPC 2.0; IEEE P3945.2 addresses this at the standards level.

Agent-to-Agent Protocol — A protocol governing peer-to-peer communication between autonomous AI agents — task delegation, capability negotiation, result exchange. A2A and ACP are deployed implementations; IEEE P3945.1 is the emerging standard.

Task Lifecycle State Machine — A defined set of states (submitted → working → completed / failed / canceled) and transition rules governing the lifecycle of a task delegated from one agent to another. Fully specified only in A2A.

Streaming / Async Messaging (SSE / Webhooks) — Protocols enabling agents to push updates in real time via Server-Sent Events (SSE) or asynchronously via webhooks, without the requesting agent polling for results. Deployed in MCP and A2A.

Multipart / Multimodal Payloads (MIME) — Support for exchanging mixed content types (text, images, audio, video, binary) in a single agent-to-agent message using MIME multipart encoding. Defined in ACP; not addressed by W3C, NIST, or IEEE.

Cross-Protocol Interoperability — The ability for an agent on one protocol (e.g., A2A) to delegate a subtask to an agent on a different protocol (e.g., ACP) within a unified task lifecycle. Currently a gap; a joint MCP/A2A convergence spec was expected Q3 2026.

Error Handling / Fault Recovery Protocols — A standardized vocabulary of error codes and recovery behaviors for agent-to-agent interactions — covering rejection, timeout, partial completion, and escalation. A2A defines task failure states but no shared error taxonomy; agents from different vendors cannot reliably distinguish a transient network error from a policy refusal.

Agent Suspension / Resume — The ability to pause a running agent task — preserving full execution state — and resume it later, potentially on a different agent instance or after a system restart. Required for long-running agentic workflows but not addressed by any current standard.

Multi-Party Coordination (3+ agents, fan-out) — Coordination of a single task across three or more agents executing in parallel — including fan-out (one orchestrator, many subagents), fan-in (aggregating results), and conflict resolution when parallel agents reach contradictory conclusions. All current protocols assume bilateral (one-to-one) delegation chains.

Agent as Identifiable Enterprise Entity — The NIST principle that AI agents must be treated as distinct, named, auditable entities within enterprise identity management — not as anonymous automated processes or extensions of the user who created them.

Service-Level Authorization (OAuth Scopes) — The use of OAuth access token scopes to define what actions an agent is permitted to take at the service or API level. Standard practice in MCP, A2A, and ACP; specified by NIST as a baseline control.

Least-Privilege / Just-in-Time Access — The security principle that agents should receive only the permissions needed for the specific task, granted only for the duration of that task. NIST specifies this as a required control for enterprise agentic deployments.

Human Approval Gates — Mandatory checkpoints at which a human must review and approve an agent's proposed action before execution, particularly for irreversible or high-impact operations. Required by NIST for high-risk agentic actions.

Tool Behavior Annotations (readOnly / destructive) — MCP metadata tags on tool definitions indicating whether a tool is read-only, potentially destructive, idempotent, or has open-world side effects. Allows calling agents to decide whether to invoke a tool and whether to require human approval.

Excessive Agency Taxonomy — NIST's classification (drawing on OWASP) of three AI agent failure modes: excessive functionality (more tools than needed), excessive permissions (broader access than needed), and excessive autonomy (acting beyond sanctioned scope).

Required Audit Trail (6 mandatory fields) — NIST's specification that agentic systems must capture: initiator identity, action type, tool accessed, timestamp, result, and approval chain — for every agent action. Enables forensic review and accountability.

Agentic Attack Surface Classification — NIST's taxonomy of attack vectors specific to AI agent systems: direct prompt injection, indirect prompt injection (malicious instructions in retrieved content), tool misuse, and multi-agent lateral movement.

Prompt Injection Defense (protocol-level) — Architectural controls preventing malicious instructions embedded in documents, emails, or web content from hijacking an agent's behavior. NIST classifies this as an architectural requirement but has not defined specific protocol-level controls.

Real-Time Inference Monitoring — The ability to observe and flag anomalous agent behavior during task execution — not merely in post-hoc audit logs. NIST acknowledges this as an important control but explicitly notes its guidance here is less mature.

Agent as Named Enterprise Identity (IAM) — The requirement that every AI agent be provisioned as a distinct, named principal within the enterprise's Identity and Access Management system — with its own credentials, policy bindings, and audit identity — rather than inheriting the identity of the human or service account that created it. NIST specifies this as a foundational control for agentic deployments.

Agent Lifecycle / Decommissioning — Governance processes for retiring an AI agent: revoking credentials, removing registry entries, cascading revocation to sub-agents, and ensuring no orphaned access persists. Not addressed by any current standard.

Supply Chain Security (agent provenance) — Controls ensuring the integrity of an agent's software, training data, and configuration — verifying that the agent running in production is the one that was reviewed and approved, and has not been tampered with or substituted. Analogous to software supply chain security (SLSA) but applied to AI agents.

Privacy / Data Retention at Handoff — Explicit rules governing what data the receiving agent may retain from a handoff payload after task completion — including personal data, intermediate results, and context. Relevant to GDPR Article 5 (purpose limitation, data minimisation). No current standard or protocol specifies post-handoff data handling obligations.

Regulatory Compliance Attestation — A mechanism by which an agent declares, in a verifiable way, the regulatory frameworks under which it operates (e.g., HIPAA, SOC 2, GDPR) — allowing the delegating agent or human principal to confirm compliance requirements are met before a task is assigned.

Benchmarking Framework — Standardized metrics and test suites for measuring AI agent robustness, capability, and interoperability in a comparable, reproducible way. IEEE P3777 is developing this; no ratified standard exists.

Memory Interoperability — The ability for an agent to persist, retrieve, and share contextual memory across sessions and platforms in a standardized way, with defined controls for access and erasure. W3C has an active Community Group; no ratified standard.

Sector-Specific Controls — Governance requirements tailored to a specific industry (healthcare, finance, education) — beyond general-purpose agent security to address domain-specific risk, privacy, and compliance obligations.

Rate Limiting / Resource Quotas per Handoff — A declaration in the handoff envelope specifying the maximum resources the receiving agent may consume in response to this specific request — CPU time, API calls, token budget, wall-clock duration, or monetary cost. Prevents a delegated agent from consuming unlimited resources of the delegating party, which is not addressable at the OAuth scope level since scopes operate at the service level, not per-invocation.

Conditions / Prerequisites at Handoff — A structured declaration in the agent-to-agent message envelope of conditions that must be confirmed true before the receiving agent may act (e.g., patient consent verified, insurance authorized, triage complete). A universal gap across all current standards.

Per-Handoff Constraint Declaration — A declaration in the message envelope specifying the scope limits of this specific delegation — which patient, which procedure, how many slots, within what time window. Distinct from OAuth scopes, which operate at service level rather than per-message level.

Boundary Declarations (what receiver may NOT do) — Explicit prohibitions in the handoff envelope listing actions the receiving agent is forbidden from taking in response to this specific request, regardless of its general capabilities or permissions.

Delegation Chain at Message Level — A record embedded in the message envelope of the sequence of agents and principals that authorized this specific request — from the original human or system through each intermediate agent to the current sender. Distinct from OAuth's service-level delegation.

Sealed Provenance Chain (hop-by-hop) — A cryptographically linked, append-only record of every agent that has handled a task, with each hop adding its identity, timestamp, and a hash of the prior chain state — creating a tamper-evident audit trail that travels with the task itself.

On-Violation Behavior Declaration — A field in the handoff envelope specifying what the receiving agent must do if it detects a violation of stated conditions, constraints, or boundaries — e.g., reject and escalate, reject and log, or reject silently.

Mandatory Receiver Inspection — A protocol requirement that the receiving agent must explicitly verify conditions, constraints, and boundaries in the handoff envelope before acting — not merely receive the message. Not specified by any current standard.

Content / Harm Scoring Layer — An optional, domain-specific gate at the agent exchange point that evaluates the content or intent of a request against an ethical or harm-risk rubric before the receiving agent acts. Relevant for content-facing agents; not universal. Not addressed by any current standard.

References

1. IEEE P3945 Standard Page — https://standards.ieee.org/ieee/3945/12436/
2. IEEE P3945.1 — Agent-to-Agent Interoperability — https://standards.ieee.org/ieee/P3945.1/
3. IEEE P3945.2 — Agent-to-Tool Interfaces — https://standards.ieee.org/ieee/P3945.2/
4. IEEE P3931 — Agent Description, Discovery, Registry — https://standards.ieee.org/ieee/P3931/
5. IEEE P3777 — Benchmarking Framework — https://standards.ieee.org/ieee/P3777/
6. IEEE P3428 — LLM Agents for Education — https://standards.ieee.org/ieee/P3428/
7. W3C ANP White Paper — https://w3c-cg.github.io/ai-agent-protocol/
8. W3C DID Core Specification — https://www.w3.org/TR/did-core/
9. W3C Verifiable Credentials Data Model v2.0 — https://www.w3.org/TR/vc-data-model-2.0/
10. W3C AI Agent Protocol Community Group — https://www.w3.org/community/aiagentprotocol/
11. W3C ANP Community Group Progress (June 2025) — https://agent-network-protocol.com/blogs/posts/w3c-agent-protocol-progress-202506
12. NIST AI Agent Standards Initiative (CAISI) Announcement — https://www.nist.gov/news-events/news/2026/02/announcing-ai-agent-standards-initiative-interoperable-and-secure
13. CSA Research Note: NIST AI Agent Standards Initiative — https://labs.cloudsecurityalliance.org/research/csa-research-note-nist-ai-agent-standards-initiative-2026032/
14. Zylos: Agent-to-Agent Communication Protocols (Feb 2026) — https://zylos.ai/research/2026-02-15-agent-to-agent-communication-protocols/
15. Zylos: MCP / A2A / ACP Convergence Report (Mar 2026) — https://zylos.ai/research/2026-03-26-agent-interoperability-protocols-mcp-a2a-acp-convergence/
16. Model Context Protocol Specification — https://modelcontextprotocol.io/specification
17. A2A Protocol — Google / Linux Foundation — https://google.github.io/A2A/
18. ACP / AGNTCY — IBM / Linux Foundation — https://agntcy.org/
19. Open Agent Schema Framework (OASF) — https://oasf.agntcy.org/

20. SPIFFE / SPIRE Workload Identity Project — https://spiffe.io/

This piece is part of an ongoing exploration of exchange-point governance for AI agent systems. The standards landscape referenced here, covering W3C ANP, NIST CAISI, and the IEEE P3945 family, is documented in full in AI Agent Interoperability Standards: What's Defined. What's in Draft. What's Missing. (© 2026 Thomas D. Holt).