Trusted AI Agents Working Group Charter

This Working Group Charter establishes the Scope and intellectual property terms used to develop the materials identified in this Working Group Charter for the Project. Only Project Steering Members, Associates, and Contributors, as applicable, that Joined this Working Group Charter will be bound by its terms and be permitted to participate in this Working Group.

1. Working Group Name. **Trusted AI Agents Working Group**.

2. Working Group Scope.

The Trusted AI Agents Working Group (WG) at the Decentralized Identity Foundation (DIF) focuses on defining an opinionated, interoperable stack to enable trustworthy, privacy-preserving, and secure AI agents. These agents act on behalf of users or systems and require robust mechanisms for identity, authority, and governance.

As autonomous agents gain real-world responsibility, composing tools, making decisions, exchanging verifiable data, the WG will build and maintain specifications, reference implementations, and governance patterns for enabling high-trust agent ecosystems.

This group complements broader initiatives by focusing directly on technical infrastructure and portable mechanisms for agent trust.

The Trusted AI Agents WG will produce:

- Use-Cases & Glossary: A living use-cases document anchoring other work items and their shared delegation terminology, and
- Evaluative Documentation: per-use-case reports including architectural or tooling-specific recommendations.

The Trusted AI Agents WG may also produce any or all of the following deliverables:

- Standards & Profiles: Define core data models and schemas for agent identity, invocation, and delegation—reusing DID and Verifiable Credential vocabularies wherever possible and extending them to meet agent-specific use cases.
- Object Capability Framework: Evaluate existing specifications (UCAN, ZCAP-LD, CapTP, macaroons) for suitability in decentralized agent contexts and, potentially, produce a streamlined, agent-centric capability expression specification bound to DIDs. This may be supersets or subsets of one or more pre-existing delegation DSLs like those listed above.
- Interoperability Libraries: Deliver reference implementations in Python and TypeScript for popular LLM agent frameworks (e.g., LangGraph, Autogen); these may support multiple transports such as DIDComm, HTTP, and/or pub/sub.
- Runtime Trust Enforcement: Specify composable workflows for both pre-execution attestation (manifests, proofs) and runtime capability checks, with built-in revocation, attenuation, and full audit trails.
- Human-Agent Delegation Patterns: Capture and document fine-grained delegation, oversight, and revocation flows between humans and agents (HITL), producing governance profiles and best-practice guidance.
- Risk Assessment: Establish a structured framework for identifying, evaluating, and mitigating risks across the entire agent lifecycle

Throughout, we will leverage existing standards wherever they suffice, extending or adapting them only to address the unique requirements of agentic workflows.

3. Copyright Policy. Each Working Group must specify the copyright mode under which it will operate prior to initiating any work on any Draft Deliverable or Approved Deliverable other than source code. The copyright mode for this Working Group is.

   ■ Creative Commons Attribution 4.0, as set forth in DIF Project Charter 4.0.3 Appendix A

4. Approved Deliverable Patent Licensing. Each Working Group must specify the patent mode under which it will operate prior to initiating any work on any Draft Deliverable or Approved Deliverable other than source code. The patent mode for this Working Group is:

   ■ W3C Mode, as set forth in DIF Project Charter 4.0.3 Appendix A,

   The assurances provided in the selected patent mode are binding on the Working Group Participant's successors-in-interest. In addition, each Working Group Participant will include in any documents transferring ownership of patents subject to the assurance provisions sufficient to ensure that the commitments in the assurance are binding on the transferee, and that the transferee will similarly include appropriate provisions in the event of future transfers with the goal of binding each successor-in-interest.

5. Source Code. Working Group Participants contributing source code to this Working Group agree that those source code contributions are subject to the Developer Certificate of Origin version 1.1, available at http://developercertificate.org/, and the license indicated below. Source code may not be a required element of an Approved Deliverable specification.

   ■ Apache 2.0, available at http://www.apache.org/licenses/LICENSE-2.0.html.

6. Non-Working Group Participant Feedback and Participation. Upon the Approval of the Working Group Participants, the Working Group can request feedback from and/or allow Non-Working Group Participant participation in a Working Group, subject to each Non-Working Group Participant executing the Feedback Agreement set forth in DIF Feedback Agreement

By making a Contribution to this Working Group or adding its name to this Working Group's member list, the member agrees to the terms of this Working Group Charter.