

# AI for Good Global Summit

## AI Standards Exchange

*Challenging the status quo of  
AI security*

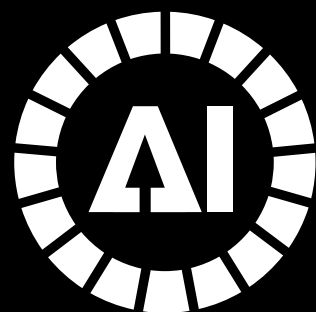
Session 2

Agentic AI identity management

**11 July 2025**

Geneva, Switzerland





**AI for Good**  
Global Summit

## Panelists

**11 July 2025**  
Geneva, Switzerland

- **Moderator: Mr Abbie Barbir**, Q10/17 Co-rapporteur
- **Ms Debora Comparin**, Thales Group | WP1/17 Chair
- **Mr Dr. Alan Chan**, Centre for the Governance of AI
- **Dr. Tobin South**, Stanford & WorkOS
- **Panel discussion covering:**
  - Principles for identity management in the context of AI agent-human interactions to define what values and goals the idea and technology should pursue.
  - Brainstorm on the open problems (e.g. AI system card, machine protocol, like A2A MCP etc.)
  - Identify gaps in standardization

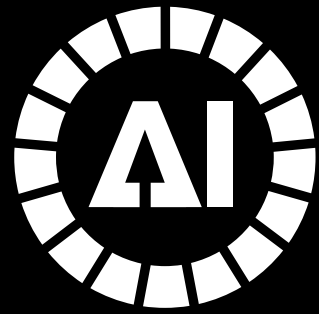


**Goal:** Understand how we can technically and conceptually assign identity to AI agents.

- . **Q1:** How do we define identity for an AI agent? Should it mirror human identity systems, or do we need something new?
- . **Q2:** What unique identifiers or credentials might AI agents require in order to be trusted in online transactions?
- . **Q3:** How can we bind an AI agent to its origin (e.g., the person or organization that created or authorized it)?
- . **Q4:** Should AI agents have persistent identities across ecosystems, or context-specific, ephemeral ones?

**Goal:** Understand how agents can act on behalf of users or organizations.

- . **Q5:** What are the best ways to represent delegation of authority from a human or entity to an AI agent?
- . **Q6:** What role can Verifiable Credentials, signed attestations, or delegation tokens play in this context?
- . **Q7:** How do we prevent unauthorized escalation—agents doing more than they were allowed?



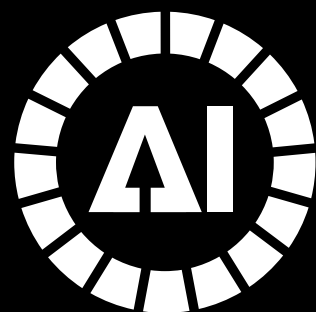
**AI for Good**  
Global Summit

## Authorization and Policy Enforcement

**11 July 2025**  
Geneva, Switzerland

- **Goal:** Explore how policies are enforced for AI actions.
- **Q8:** How can we ensure that an agent's actions stay within the bounds of its assigned policies or scope?
- **Q9:** Can existing access control models (RBAC, ABAC, ReBAC) be extended to AI agents effectively?
- **Q10:** Who is accountable if an AI agent takes an unauthorized or harmful action? The developer, the delegator, the agent itself?





**AI for Good**  
Global Summit

## Security and Trust Infrastructure

**11 July 2025**  
Geneva, Switzerland

- **Goal:** Examine how security primitives apply to AI agents.
- **Q11:** How can cryptographic primitives like public key infrastructure (PKI) or FIDO2 be extended to AI agents?
- **Q12:** Can agent identity and activity be logged and audited in a tamper-evident way? How important is this?



- **Goal:** Consider broader implications, including standards and regulation.
- **Q13:** What role should identity standards bodies (like W3C, FIDO Alliance, ISO) play in defining AI agent identity?
- **Q14:** How do we prevent identity misuse or impersonation by rogue AI agents?
- **Q15:** Should we require registration or certification of AI agents in certain domains (e.g., healthcare, finance)?