ITU AI FOR GOOD GLOBAL SUMMIT 2026

Designing Trust Management for Agentic AI

Why trust must become explicit, measurable, and enforceable for autonomous AI systems

Samuel Waymouth
Senior Compliance
Solution Architect

© 2026, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The Trust Problem

- Agentic AI has shifted from advisory to autonomous — agents perceive, decide, and act
- Trust was implicit when humans decided; now it must be explicit and enforceable
- Four trust dimensions: Identity (who?), Intent (what goal?), Capability (safely?), Accountability (auditable?)
- The trust gap: governance frameworks (NIST, ISO, EU AI Act) were designed for predictive models, not autonomous actors
- No standard for agent identity, cross-org trust federation, or dynamic risk classification

aws
© 2026, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Trust Requires Structure, Not Approval Queues

Identity Trust

Who is this agent? Who deployed it? Who is accountable when it acts across organizational boundaries?

Intent Trust

Is the agent's goal aligned with its principal's intent? Can we verify alignment in multi-agent delegation chains?

Capability Trust

Are boundaries enforced mechanistically — not just declared? Graduated autonomy: narrow first, expand with demonstrated trust.

Accountability Trust

End-to-end traceability. Every action attributed. Trust revocable. Agents more auditable than human processes, not less.

© 2026, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Call to Action: Research Priorities for ITU-T SG17

- 1. Standardized agent identity protocols — analogous to X.509 certificates for agents
- 2. Cross-organization trust federation framework — SAML/OIDC equivalent for agent-to-agent trust
- 3. Automated, continuous compliance attestation — agents that prove they complied, not just claim it
- 4. Dynamic autonomy classification — risk labels that reflect what the agent is doing, not just what it was designed to do
- 5. Policy-as-code enforcement — governance enforced computationally at execution time, not documented in wikis

aws
© 2026, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Samuel Waymouth
swaym@amazon.co.uk

© 2026, Amazon Web Services, Inc. or its affiliates. All rights reserved.