INTERNATIONAL TELECOMMUNICATION UNION
TELECOMMUNICATION STANDARDIZATION SECTOR
STUDY PERIOD 2025-2028

SG17-LS160
STUDY GROUP 17
Original: English

Question(s): All/17 Virtual, 26 June 2026

LS

Source: ITU-T Study Group 17

Title: LS update of SG17 on Agentic AI, prior and after the SG17 5th plenary meeting

LIAISON STATEMENT

For action to: IETF IAB, IESG, SEC, ART; IRTF HRPC, DINRG

For information to: -

Approval: ITU-T Study Group 17 Management team meeting (by correspondence, virtual, 26 June 2026)

Deadline: 31 August 2026

Contact: Arnaud TADDEI
Chair, ITU-T SG17
E-mail: Arnaud.Taddei.SDO@gmail.com

Abstract: ITU-T Study Group 17 is pleased to share with the IETF and the IRTF an update on its work on the security and trust of Agentic AI, covering both the period prior to and the outcomes of the SG17 5th plenary meeting (Geneva, 1-10 June 2026): the restructuring of SG17 (Questions and the new Focus Group), the related activities (workshops, the X.509 Day and their reports, and forthcoming events), and the work programme (texts determined and the new work items established). SG17 invites the IETF and the IRTF to share, for action, their own updates and future vision in this area.

ITU-T Study Group 17 (Security) is pleased to share with the IETF and the IRTF a consolidated update on its work related to the security and trust of Agentic AI, covering the period prior to, and the outcomes of, the SG17 5th plenary meeting (Geneva, 1-10 June 2026). SG17 treats Agentic AI as a first-order security and trust topic and welcomes close coordination with the IETF and the IRTF.

For the purposes of this liaison, and following the Terms of Reference of the new SG17 Focus Group (see [1]), an Agentic AI (or AI Agent, or Agent) is "an autonomous software entity that interprets goals, formulates intent, and executes actions, often by invoking external tools, APIs, or other agents to achieve outcomes on behalf of users; its behaviour is driven by model-based reasoning, typically relying on Large Language Models (LLMs) or other Foundation Models (FMs)."

1 Executive summary

This liaison statement provides the IETF and the IRTF with a consolidated update on the work of ITU-T SG17 on the security and trust of Agentic AI, before and after the SG17 5th plenary meeting (Geneva, 1-10 June 2026). SG17 treats Agentic AI as a first-order security and trust topic and seeks close coordination with the IETF and the IRTF.

SG17 has restructured itself around AI and Agentic AI: a new Question 16/17 on the security of AI was created (at the 9 April 2026 virtual plenary), Question 3/17 was merged into Question 10/17, Question 7/17 was amended, and SG17 agreed to amend Question 8/17 (provisionally, pending endorsement by TSAG in January 2027); a new Focus Group on Trust and Identity for Humans and Agentic AI (FG-TIDA) was approved and will be launched on 9 July 2026 at the AI for Good Global Summit in Geneva.

The Agentic-AI work programme now comprises 25 work items across five Questions (Q1, Q2, Q7, Q10, Q16): 8 were already in development before the plenary and 17 were newly established at it; one foundational text (X.sr-AI) was determined.

SG17 invites the IETF and the IRTF, for action by 31 August 2026, to share their own updates and future vision; to develop (IESG) a mapping table against the SG17 work items; to engage on a common identity stack for agents and humans (with X.dpidm-aAI as a first candidate); to explore ways to align the leadership of the two organizations; and to host SG17 presentations at IETF 126 in Vienna. The relevant SG17 Temporary Documents are attached.

2 Restructuring of SG17

In preparation for the next study period and as part of its ongoing modernization, SG17 has restructured several of its Questions and created a new Focus Group, both directly relevant to AI and Agentic AI.

- Prior to this plenary, at its fully-virtual plenary meeting of 9 April 2026, SG17 created a new Question 16/17 on the security for AI and Agentic AI, merged the former Question 3/17 into Question 10/17, and amended Question 7/17 to explicitly address AI and Agentic AI.
- In addition, Question 8/17 has been provisionally amended, pending endorsement by TSAG in January 2027, and the revision of the Terms of Reference of Question 1/17 was initiated at the 5th plenary.
- The Correspondence Group on restructuring and modernization was reorganized: CG-RES-MODERN became CG-NSP-MODERN (preparation for the next study period and modernization).
- SG17 approved the creation of the Focus Group on Trust and Identity for Humans and Agentic AI (FG-TIDA), with its Terms of Reference (see [1]). FG-TIDA covers terminology and working definitions, trust models and the trust control plane, the identity stack for agentic AI, and interoperability for environments involving both humans and agentic AI. FG-TIDA will be officially announced and launched on 9 July 2026 during the AI for Good Global Summit in Geneva.

3 Activities (workshops, X.509 Day, and forthcoming events)

3.1 Workshops and events held

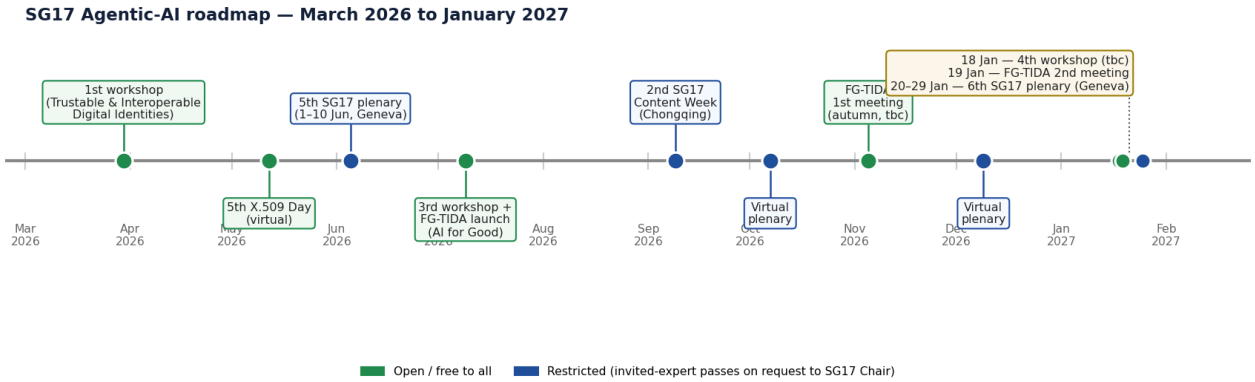
Date	Event	Reference	Participation
30–31 March 2026	1st workshop — "Trustable and Interoperable Digital Identities for Human and Agentic AI" (Geneva, 1st SG17 Content Week)	Executive summary (see [2]); report (see [3])	Open
12 May 2026	5th ITU-T X.509 Day (virtual)	Report (see [4])	Open
2 June 2026	2nd workshop — "Global interoperability for trust management of digital identity for humans and agents" (Geneva, alongside the plenary)	Report (see [5])	Open

3.2 Forthcoming events and key meetings

Date	Event	Reference	Participation
9 July 2026	3rd workshop — "Designing the trust management for Agentic AI" (Geneva, AI for Good Global Summit)	Draft programme (see [6])	Open (free for everyone)
7–11 September 2026	2nd SG17 Content Week (Chongqing, China)	TBA (to be announced)	Restricted[1]
7 October 2026	Virtual SG17 plenary meeting	TBA (to be announced)	Restricted[1]
Autumn 2026 (tbc)	FG-TIDA 1st meeting (Paris, to be confirmed)	TBA (to be announced)	Open (free for everyone)
9 December 2026	Virtual SG17 plenary meeting	TBA (to be announced)	Restricted[1]
18 January 2027	4th workshop (to be confirmed)	TBA (to be announced)	Open (free for everyone)
19 January 2027	FG-TIDA 2nd meeting	TBA (to be announced)	Open (free for everyone)
20–29 January 2027	6th SG17 plenary meeting (Geneva)	TBA (to be announced)	Restricted[1]

[1] *Restricted — invited-expert passes are available on request to the SG17 Chair.*

3.3 Roadmap to the 6th SG17 plenary (January 2027)



Note: in addition to the events above, a number of additional SG17 Rapporteur Group meetings will be held; for more information, please contact tsbsg17@itu.int.

4 Work programme on Agentic AI

The following summarises the Agentic-AI work items, before and after the 5th plenary. Each work item is linked to its entry in the ITU-T work programme. The table below summarises the figures.

	Agentic-AI work items	Questions impacted
In development prior to the 5th plenary	8	3 (Q7, Q10, Q16)
New work items established at the 5th plenary	17	5 (Q1, Q2, Q7, Q10, Q16)
Total (distinct)	25	5 (Q1, Q2, Q7, Q10, Q16)

In addition, one of these items — X.sr-AI (Q16/17) — was determined (TAP) at the 5th plenary.

4.1 Items in development prior to the 5th plenary

- Q10/17 — X.dpidm-aAI: Terminology and design guidelines for Agentic AI identity management.
- Q10/17 — XSTR.gidi: Globally interoperable digital identity (including humans/enterprise/non-humans i.e. Agentic AI).
- Q16/17 — X.S-AIA: Security requirements and guidelines for AI agent.
- Q16/17 — X.sr-taimas: Security requirements for terminal-based AI multi-agent system.
- Q16/17 — XSTR.ltf-AAI: Landscape of trust framework for Agentic AI (Technical Report).
- Q16/17 — XSTR.sem-AIA: Security evaluation methods for AI agent.
- Q16/17 — X.sr-AI: Security requirements for AI systems (foundational).
- Q7/17 — X.gavd-mas: Guidelines for application vulnerability detection using multi-agent system.

4.2 Actions at the 5th plenary (June 2026)

TAP approval (Agentic-AI scope): none.

TAP determination (Agentic-AI scope): Q16/17 — X.sr-AI, “Security requirements for AI systems” (foundational AI-security text underpinning agentic systems).

AAP consent (Agentic-AI scope): none.

Agreement (Agentic-AI scope): none.

4.3 New work items established at the 5th plenary (June 2026)

4.3.1 Question 1/17

X.dmst **— Design methodology for security and trust**

Scope: Establishes a design methodology applicable to the design of secure and trustworthy ICT systems, re-articulating the design-theoretic discussion currently in clause 16 of X.arch-design and addressing novel technologies — including Agentic AI — without precedent human experience.

Summary: The proposed Recommendation establishes a Design methodology applicable to the design of secure and trustworthy ICT systems. It re-articulates the design-theoretic discussion currently found in clause 16 of X.arch-design (out of scope of that Recommendation), formalizes the four-dimension model (anthropology/ethics/law/technology) and its tree-based reverse traversal for novel technologies without precedent human experience, and operationalises the trustworthiness aggregation function. The Recommendation is intended to be foundational to the CRAMMS Series and complementary to X.arch-design, X.SECADEF, X.805, X.1205, X.1051 and X.1060.

Note: Beyond X.dmst, Question 1/17 has also successfully progressed several work items on the foundations of security and trust that underpin this Agentic-AI work. In particular, Q1/17 is advancing the CRAMMS series — X.cramms (Framework for Cyber Security Reference Architectures, Models and Methodologies Strategy) and its companion Technical Report XSTR.CRAMMS, which expands the CRAMMS work — together with related foundational items including X.arch-design (Security Design Principles and Concepts), X.secadef (Security capabilities definitions), X.cs-ra (Cyber Security Reference Architecture), XSTR.psam (Practical Security Architecture Methodology) and X.crta (Framework for Cyber Resilience Testing and Assurance). On trust, Q1/17 is progressing XSTR.trust (Trust framework for Telecommunications/ICT). These foundational and trust work items provide the conceptual scaffolding on which the Agentic-AI security and trust work items build.

4.3.2 Question 2/17

XSTR.sg.aai **— Security Guidelines for Agentic AI-Driven IMT-2020 and beyond based Heterogeneous Networks**

Scope: Security guidelines for the design, deployment and operation of agentic AI systems that perceive, reason, collaborate and autonomously act across IMT-2020/5G, IMT-2020-Advanced and IMT-2030/6G heterogeneous networks (core, RAN, edge, IoT, WSN, MANET, SDN), defining a reference framework across the IMT infrastructure/network layers, an AI-components plane, and agent-to-agent, adversarial-defence and governance control planes.

Summary: Agentic AI dramatically expands network complexity and attack surface across multi-tier heterogeneous networks, which traditional static security cannot defend at modern scale and speed. The item provides the A.1 justification and initial text for security guidelines covering autonomous AI agents acting as network self-defence functions.

4.3.3 Question 7/17

X.srm-AIphs **— Security requirements and measures for AI agent-enabled public health services**

Scope: Security requirements and measures for AI agent-enabled public health services, including the security threats, the security requirements, and the measures for security protection of such services.

Summary: Addresses the security of AI agents deployed in public-health service contexts. Approval process is TAP, with liaisons to ITU-T SG21 and ISO/IEC JTC 1/SC 27.

4.3.4 Question 10/17

X.AIA-dii-req **— Security requirements for DLT-based decentralized identity interoperability across artificial intelligence agents**

Scope: Security requirements for DLT-based decentralized identity interoperability (DII) across AI agents with heterogeneous identity solutions: an end-to-end DII framework, the relevant security risks and threats, and the security-enhanced framework and requirements.

Summary: AI agents are autonomous, networked software entities acting on behalf of users, services or organizations; their digital identity (attributes, credentials, attestations and provenance bindings) defines who an agent is, what it is authorized to do and whom it represents. Identity interoperability requires standardized data models, minimal attribute sets and supporting infrastructure (federation, trust anchors, credential-exchange protocols, onboarding and lifecycle management) so that agents from different vendors, domains and jurisdictions can authenticate, negotiate trust and enforce policy without bespoke integration, with end-to-end identification, authentication and authorization for accountability. Security requirements for DII across AI agents must guarantee verifiable, tamper-resistant identity assertions and bindings; secure provisioning and key/credential management; consistent revocation and recovery; interoperable trust and federation; and end-to-end authorization enforcing least privilege and safe delegation, with auditability, provenance, privacy-preserving disclosure and a threat/assurance framework against impersonation, supply-chain compromise and cascading cross-domain failures. The draft Recommendation introduces an end-to-end DII across AI agents with heterogeneous identity solutions, identifies the relevant security risks and threats, and specifies the security-enhanced framework and requirements.

X.aiidm **— Canonical Agentic AI Identity Data Model**

Scope: Defines an abstract information model for agentic identity: core data structures, attributes and relationships; claim vocabularies and ontologies specific to agentic systems; and binding mechanisms to controllers, models and governance policies.

Summary: A canonical, technology-neutral data model for agentic AI identity addressing gaps in representation, interoperability and governance of autonomous agents across heterogeneous trust environments, providing a semantic foundation for consistent identity expression, verifiable accountability and secure multi-hop interactions.

XSTR.id-AA-dis **— Identity management framework for AI agents using decentralized identity systems**

Scope: A framework for identity management of AI agents — including authentication and authorization using decentralized identity systems (DIDs / verifiable credentials) — with a reference architecture and identity-management workflows, agentic-AI-specific risks and mitigations, and accountability and privacy considerations.

Summary: Supports secure, interoperable operation of AI agents that independently plan, decide and act across platforms and trust domains, using decentralized identity to provide scoped, task-specific access controls and explicit, auditable agent workflows.

XSTR.athena **— Autonomous Trusted Hierarchical ENtity Architecture (Technical Report)**

Scope: A Technical Report analysing requirements for telco-grade identity for Agentic AI and introducing ATHENA as a framework for secure identification, delegation, lifecycle management and accountability of AI agents across networks and services.

Summary: Existing identity/authentication/authorization mechanisms were designed for human users, devices or applications and are insufficient for autonomous agents. ATHENA proposes a telco-grade trust and identity framework enabling trusted, controlled and accountable operation of AI agents at global scale.

4.3.5 Question 16/17

X.Max-trust **— A Multi-party, Agentic, and eXtensible Trust Framework for next-generation Agentic-AI-enabled telecommunication networks**

Scope: A trust framework for next-generation agentic networks where AI agents operate with increasing autonomy in network management, resource orchestration and service delivery; covers core trust principles, a high-level functional architecture and components, implementation requirements, and enabling technologies.

Summary: Addresses trust challenges from the shift to intelligent, autonomous communication systems. The current baseline incorporates audit findings (alignment with X.arch-design design principles and a definition of trust).

X.S-AIAI **— Security framework and requirements for invocation by AI agent**

Scope: Identifies threats across the process of invocation by an AI agent to other AI agents, external tools and Modular Capability Units (MCUs), and establishes security requirements and a framework defining core security capabilities and a standard invocation workflow.

Summary: As AI agents autonomously perceive, decide and act, capability invocation — leveraging other agents, tools and reusable MCUs — is central to completing complex tasks and introduces new security risks that this item addresses.

X.SG-AIAD **— Security Guidelines for Artificial Intelligence Agent Data**

Scope: Studies AI-agent data assets and processing activities and performs systematic threat/risk identification; provides security threats across the AI-agent data-processing loop (perception & cognition, planning, memory, action, external environment & agent groups), security requirements and corresponding security guidelines.

Summary: Focuses on data security across the agentic-AI processing stages, with an AI-agent architecture, a gap analysis of related texts, and analysis of the AI data lifecycle.

X.aaia-sec **— Security framework and technical requirements for autonomous AI agent**

Scope: Risk analysis of autonomous AI agents; a security framework covering initial-phase security, agent runtime-phase security and security observation; and security technical requirements for autonomous AI agents.

Summary: As AI evolves toward agentic AI with autonomous planning, decision-making and execution, autonomous agents extend LLMs into full runtime systems whose functionalities create security challenges beyond traditional protection; the item proposes a systematic security defence framework.

X.sg-AIAof **— Security guidelines for AI agent orchestration framework**

Scope: Systematic security guidelines for a cloud-side AI agent orchestration framework, covering the full-stack architecture (cloud infrastructure, foundation-model, agent, orchestration & runtime, and application layers), and clarifying each layer's security boundaries and control responsibilities (ex X.sg-AIAof).

Summary: Large-scale AI-agent deployment has made the orchestration framework core foundational infrastructure — the hub connecting computing resources, foundation-model services and atomic capability components — motivating security guidelines for this layer.

XSTR.magAI-behavior **— Capabilities for Agentic AI security and trust controls to manage the behaviour of single or multiple agentic multi-objective AI systems**

Scope: Addresses emergent, misaligned and multi-objective behaviours in advanced and multi-agentic AI systems and their cybersecurity implications, at the intersection of AI systems engineering, cybersecurity and telecommunications, with focus on ML/deep-learning systems integrated into critical environments.

Summary: Builds on the bijection between multi-criteria choice optimization and collective-choice theory and on empirical evidence from frontier-model system cards, proposing controls to manage the behaviour of single or multiple agentic multi-objective AI systems.

XSTR.atcp **— Cryptographic primitives for enhancing AI agent trustworthiness**

Scope: Normative cryptographic primitives for the quantification and verifiable communication of trust confidence between interacting AI agents: trust-score structures, confidence-interval representations, trust-assertion atoms cryptographically bound to verified agent identity, and trust-composition procedures for multi-agent scenarios.

Summary: Operationalizes the measurement layer underlying agentic-AI trust frameworks, providing verifiable quantification primitives that complement the broader multi-party agentic trust work.

XSTR.stv-OC **— Security threats and vulnerabilities in the OpenClaw framework**

Scope: A Technical Report on security threats and vulnerabilities in the OpenClaw agentic framework — prompt injection, tool misuse, memory poisoning, privilege escalation, cross-session leakage, data exfiltration, denial-of-service, routing tampering, observability abuse and supply-chain compromise — providing a basis for protected assets, security objectives, attacker models and future threat/vulnerability analysis.

Summary: The shift from passive Q&A systems to autonomous tool-using agents breaks conventional assumptions separating content, execution, identity and authorization; the item documents threats and vulnerabilities in this agentic framework.

X.sr-AIapi **— Security Requirements for AI Service APIs**

Scope: Security requirements for AI service APIs — interface-level controls for AI inference/generation/analysis/action services exposed via APIs — covering identity and access

management, capability authorization and combination control, inference-context isolation, logging/auditing, abuse prevention and multi-tenant deployment.

Summary: Defines verifiable, non-intrusive interface-level controls at the API layer, addressing AI-specific risks such as instruction-driven inputs, non-deterministic outputs, multi-step operations and multi-tenant vulnerabilities — directly relevant to how agents invoke external tools/APIs.

XSTR.aam **— AI accountability manifest**

Scope: Normative mechanisms for cryptographic accountability of AI agents through verifiable manifest declarations binding agent identity, declared intent and operational scope; covers manifest format, intent-verification procedures, accountability-chain integrity for multi-agent delegation, and interoperability provisions.

Summary: Provides cryptographic accountability binding agent identity to operator-declared intent and operational scope, with intent verification and accountability-chain integrity for multi-agent delegation, complementing the agentic trust-framework work.

5 Actions requested of the IETF and the IRTF

SG17 invites the IETF and the IRTF, for action by 31 August 2026:

1. To the IAB — to share the IETF and IAB updates and future vision on the security, identity and trust of Agentic AI (including relevant RFCs, Internet-Drafts, BoFs, Working Groups and Research Groups), and to explore, together with the SG17 management team, new approaches to joint collaboration; the points below give first examples of such collaboration.
2. To the IESG — to identify overlaps and complementarities with the SG17 work items above. SG17 would greatly benefit if the IESG could consider developing its own mapping table, as it holds the authoritative knowledge of its structure and work items, to bring clarity and avoid duplication.
3. To the IAB, the IESG and the relevant existing or future Working Groups related to Agentic AI, in particular in the SEC and ART areas (ADs) — to engage with SG17 on the security, identity and trust of Agentic AI, including participation in the 9 July 2026 SG17 workshop and in the new Focus Group FG-TIDA.
4. As a first illustration of such collaboration, to take X.dpidm-aAI as a first candidate to agree a common identity stack for Agentic AI and humans.
5. As a further illustration, to propose concrete ways to align the leadership of the two organizations.
6. Several IETF and IRTF groups may wish to invite SG17 to present its results at the next IETF and IRTF meetings during IETF 126 in Vienna — for example, X.dmst for HRPC and/or DINRG; this overall liaison statement during the IAB open meeting; a presentation at any BoF aiming to establish new Agentic-AI-related Working Groups; and leadership discussions between the two organizations, or between specific IETF Area Directors and SG17.

SG17 looks forward to a close and continuing collaboration with the IETF and the IRTF on the trust and identity of humans and agentic AI. The SG17 Chair will be pleased to consider any request from IETF or IRTF participants to attend any of the restricted SG17 meetings as invited experts, free of charge.

Attachments: 6

For those recipients (e.g. IETF participants) who are also ITU-T Sector Members, ITU-T SG17 Associate members, or from ITU Academia, the ITU-T SG17 Temporary Documents (TDs)

referenced in this liaison are provided below, with their links, by courtesy. Access to these documents is restricted to ITU-T TIES users.

- SG17-TD295/P — Terms of Reference of the Focus Group on Trust and Identity for Humans and Agentic AI (FG-TIDA)
- SG17-TD237/P — Executive Summary – SG17 Workshop on "Trustable and Interoperable Digital Identities for Human and Agentic AI" (30–31 March 2026)
- SG17-TD261/P — Draft report of ITU Workshop on Trustable and Interoperable Digital Identities for Human and Agentic AI (Geneva, 30–31 March 2026)
- SG17-TD260/P — Draft report of Fifth ITU-T X.509 Day Event (virtual, 12 May 2026)
- SG17-TD256/P — Debriefing session of ITU-T SG17 workshop on 2 June 2026
- SG17-TD241/P — Draft programme — ITU Workshop "Designing Trust Management for Agentic AI" (9 July 2026, Geneva)

_____