

FOCUS GROUP ON EMBODIED ARTIFICIAL INTELLIGENCE FOR MULTIMEDIA TECHNOLOGIES (FG-EAI)

TERMS OF REFERENCE

Background

This Focus Group is established under the policy framework of Recommendation ITU-T A.7 <https://www.itu.int/rec/T-REC-A.7/en>

1 Rationale

Embodied artificial intelligence (EAI) represents a transformative paradigm in which AI is integrated into physical bodies. Its embodied interaction and native multimodal perception naturally enhance and extend next-generation multimedia applications. Meanwhile, EAI systems are no longer designed merely to perform pre-programmed tasks in structured settings. Driven by its core capabilities, such as cognition, learning, and collaboration, EAI systems are now being designed to coexist and collaborate with humans in open, dynamic environments, supporting a broad range of application domains, including smart manufacturing, public service, logistics, healthcare, and home automation, and creating growing demands for interoperable platforms, multimodal frameworks, and scalable system architectures.

Despite its potential, current EAI technologies remain technically immature, rapidly evolving, and fragmented, with significant gaps in reliability, accessibility, and usability. There is currently no global consensus on terminology, architecture, or evaluation approaches for evaluating performance, aligning with human values and needs, and guarantee safety, and achieving interoperability. These stand in stark contrast to the ITU's vision of trusted, safe, and inclusive development of AI¹ technologies, underscoring the urgent need for a coordinated effort across institutions and stakeholders.

2 Scope

The scope of this Focus Group reflects the ITU-T SG21 remit and is limited to embodied artificial intelligence for multimedia technologies. In this context the Focus Group will canvass EAI for human-centric multimedia applications by studying its technical capabilities, including cognition to enhance understanding of the physical world, learning to adapt tasks through trial and error, and collaboration among humans, machines, and environments. It will investigate foundational technologies including real/virtual-world datasets, closed-loop control, world-model training and advanced human-machine interaction. It will also consider application-level challenges, such as insufficient capabilities and instability in long-horizon tasks, that contribute to today's fragmented development and hinder industrial-scale deployment. Safety and human factors will also be addressed, clarifying technology readiness and human oversight mechanisms.

The Focus Group will identify stakeholders with whom ITU-T could collaborate, and for enabling the inclusion of non-members to contribute to the technical pre-standardization work. Potential collaboration can be built across academia, industry, regulatory bodies, and standards development organizations (SDOs), to ensure inclusive and interoperable technical foundations.

3 Objectives

According to its scope (clause 2), the Focus Group aims to identify key technologies, conduct pre-standardization work, and facilitate standard adoption for enabling EAI, including for human-centric multimedia applications. It has the following objectives:

¹ [HTTPS://WWW.ITU.INT/EN/ACTION/AI/PAGES/DEFAULT.ASPX](https://www.itu.int/en/action/ai/pages/default.aspx)

- To study and harmonize concepts, terminologies, working definitions, vision, and ecosystem relevant to EAI.
- To analyse the standardization gaps toward EAI-enabled services.
- To identify and study enabling technologies, trends, and standardization needs relevant to EAI capabilities.
- To identify use cases to guide future EAI development for multimedia applications across various service domains, facilitating EAI adoption.
- To identify criteria and benchmarks for assessing performance, accessibility, usability, safety, and reliability of multimedia services enabled by EAI.
- To take into account key aspects of trustworthy EAI-enabled multimedia applications, including safety and human factors and privacy.
- To identify opportunities for collaboration among relevant ITU-T Study Groups (SGs), external SDOs, industrial entities, and academic institutions, including both ITU and non-ITU members, to guarantee interoperability and foster collaboration.
- To provide a platform to share findings and for dialogue on technical policy and regulatory implications of EAI related to multimedia aspects.
- To facilitate community engagement and best practice, knowledge dissemination to accelerate consensus in the technical approaches, ensuring alignment with the evolving yet maturing landscape of EAI.
- To study and gather information to develop a pre-standardization roadmap.

4 Specific tasks and deliverables

According to its scope (clause 2), the Focus Group aims to develop deliverables considering the following aspects:

- develop deliverables covering concepts, working definitions and terminologies of EAI systems;
- identify standardization gaps and technical requirements to ensure interoperability between heterogeneous EAI systems;
- identify needs for improvement of service architectures in computing and network/cloud resource provisioning, scheduling, and optimization to enable EAI for multimedia applications;
- to study and gather information to develop a pre-standardization roadmap on EAI.

To develop deliverables on technologies related to core capabilities of EAI in the context of multimedia applications, including but not limited to:

- study common EAI foundational technologies;
- study cognition-related technical foundations for EAI;
- study learning-related foundations for EAI;
- study collaboration-related foundations for EAI.

To develop deliverables on use cases and technical enablers related to EAI in human-centric multimedia applications, including:

- safety and human factors in cross-sector collaboration (e.g. healthcare, manufacturing, education, retail, logistics).
- technology readiness of EAI systems assessment to facilitate the adoption of EAI applications.
- evaluation methods and assessment frameworks, procedures and metrics for accessibility, reliability, and usability of human-centric EAI applications.

To support liaison with relevant ITU-T SGs, industrial entities and external SDOs, especially:

- collaborate with related ITU-T study groups and other standards bodies working on Embodied AI and related areas;

- collaborate with academia and industry to promote interoperability and accelerate the adoption of standard approaches;
- organize workshops, panels, hackathons, and global challenges in coordination with major ITU events (e.g., AI for Good Global Summit) to collect knowledge and experience, foster innovation, and promote youth education.

Upon completion of its lifetime, the FG should provide the final report, including the complete set of deliverables to its parent group SG21.

5 Relationships

The Focus Group will work through regular meetings with ITU-T SG21 when possible. Its pre-standardization work serves to support, complement, and inform the standardization activities of SG21, without duplicating ongoing efforts. Furthermore, it will liaise with other relevant ITU-T SGs as necessary to ensure coherence, promote synergy and avoid duplication.

The Focus Group will collaborate with other relevant entities and groups in accordance with Recommendation ITU-T A.7. These include governments, non-governmental organizations (NGOs), policy makers, SDOs, industry forums and consortia, companies, academic institutions, research institutions and other relevant organizations, dealing with aspects of EAI from their own perspectives.

The Focus Group will organize thematic workshops and forums to promote the FG activities, encouraging both ITU members and non-ITU members to jointly contribute to this Focus Group and its objectives.

6 Structure

The Focus Group may establish a sub-group structure if needed.

7 Parent group

The parent group of the Focus Group is ITU-T SG21.

8 Leadership

See clause 2.3 of Recommendation ITU-T A.7.

9 Participation

See clause 3.1 of Recommendation ITU-T A.7.

A list of participants will be maintained for reference purposes and reported to the parent group.

10 Administrative support

See clause 5 of Recommendation ITU-T A.7.

11 FG Financing

See clauses 4 of Recommendation ITU-T A.7.

12 Meetings

The Focus Group will conduct regular meetings. The frequency and locations of meetings will be determined by the Focus Group management. The Focus Group will use remote collaboration tools to the maximum extent.

The meeting dates will be announced by electronic means (e.g., e-mail and website) at least four weeks in advance.

13 Deliverables

See clause 8 of Recommendation ITU-T A.7.

14 Working language

The working language is English.

15 Approval of deliverables

Approval shall be obtained by consensus according to clause 8.2 of Recommendation ITU-T A.7.

16 Working guidelines

Working procedures shall follow the procedures of Rapporteur meetings (ITU-T A.1). No additional working guidelines are defined.

17 Progress reports

See clause 3.6 of Recommendation ITU-T A.7.

18 Announcement of Focus Group formation

The formation of the Focus Group will be announced via TSB Circular to all ITU membership, via the ITU-T news log, press releases and other means, including communication with the other involved organizations.

19 Milestones and duration of the Focus Group

The Focus Group lifetime is set for one year from the first meeting with possibility of extension.

Any extension beyond the above-mentioned date is subject to the decision by ITU-T SG21, in line with ITU-T A.7.

20 Intellectual Property Rights

See clause 7 of Recommendation ITU-T A.7.
