Question 5/21 – AI-enabled multimedia applications

(Continuation of Question 5/16)

1 Motivation

The recent success of artificial intelligence (AI) in various applications has raised study and utilization of AI technology to a new height. AI has been the apex technology of the information age. One of the most exciting aspects of the AI inflection is that "real-world" use cases abound. At the same time, deep-learning-enabled advances in computer vision and technologies such as natural language processing are dramatically improving the quality of people's work and life.

At present, the ecological pattern of AI has been established gradually. Specialized intelligent applications will be the main potential area for the future development of AI. No matter whether it is a specialized or generalized application, the AI studies will focus on analysing data at three basic levels: computing layer (base), algorithm layer (technology) and application layer. Datasets are combined with powerful technology, value is being created and competitive advantage is being gained.

Multimedia has become the pioneer, and the concept of "AI-enabled multimedia" as well as "intelligent multimedia" has already come up. Scientists, engineers all over the world are delving into some of the most exciting areas such as computer vision and speech technologies. Computers are being taught to understand and generate multimedia contents, augmenting reality to guide field technicians when operations get complex, helping computers recognize people, detect sentiment and speak with emotion, and enrich video with metadata extracted from it.

AI-enabled multimedia applications are booming, emerging technologies brings not only new opportunities, but also new challenges as well as new demands. Taking multimedia data as an example, huge volume multimedia data does not indicate high quality labelling data that AI applications could benefit. If no guidelines or standards of multimedia format, labelling are developed, multimedia data collected and labelled by company A could not be used in company B. These results in huge resource waste and prevents the data flow, which can severely hinder the development of the AI industry.

This Question focuses on artificial intelligence-enabled multimedia applications, 1) to identify challenges facing the deployment of AI-enabled multimedia applications, 2) to analyse the impact of AI technologies in standards for multimedia applications, and 3) to identify evaluation and assessment specifications of applications, algorithms and data structures for standards in AI-enabled multimedia applications, in order to boost and innovate the development of multimedia as well as AI industry.

2 Study items

Study items to be considered include, but are not limited to:

– scope and definition of AI as it relates to multimedia applications;
– identify specific use cases where AI can be applied to multimedia applications;
– identify AI techniques facilitating intelligent and automated multimedia-based tasks;
– identify use cases, framework and requirements of multimedia applications using AI generated content (AIGC), including those utilizing large foundational models, techniques enabling AIGC are to be studied, works related to content itself, such as creation, inspection, regulation, etc., are out of the scope of this Question;
– data preparation for use with AI-enabled multimedia applications;
– specific system characteristics for AI-enabled multimedia applications;
– assessment and evaluation techniques for AI-enabled multimedia services;
– identification of how AI may impact existing multimedia applications;

– accessibility of AI enabled multimedia applications for all, to help persons with disabilities.

3 Tasks

Tasks include, but are not limited to:

– determine the scope and definitions of AI as it relates to multimedia applications;
– identify and collect specific use cases where AI can be applied to multimedia applications;
– identify data preparation requirements, including but not limited to data collection, data labelling, data control and data delivery;
– identify requirements, framework and architecture of AI systems/platforms enabling multimedia applications;
– identify multimedia related AI applications in vertical industries, such as manufacturing industry, energy industry, etc.;
– identify the requirements for evaluation and assessment methodologies for quantifying the performance of AI-enabled multimedia applications;
– identify and collect use cases on accessibility of AI enabled multimedia applications;
– maintain deliverables under the responsibility of the Question, including: ITU-T F.742.1, F.746.13， F.746.15, F.746.16, F.747.12, F.748.14, F.748.15, F.748.17, F.748.18, F.748.19, F.748.20, F.748.21.

An up-to-date status of work under this Question is contained in the SG21 work programme (https://itu.int/ITU-T/workprog/wp_search.aspx?sp=18&q=5/21).

4 Relationships

Recommendations:

– F.700-series

Questions:

– All Questions of Study Group 21

Study groups:

– ITU-T SGs 12, 13, 15, 17 and 20

Other bodies:

– ISO, IEC, ISO/IEC, ETSI, IEEE
– Artificial Intelligence Industry Alliance
– China Communications Standards Association