|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ITU logo | INTERNATIONAL TELECOMMUNICATION UNION  **TELECOMMUNICATION STANDARDIZATION SECTOR**  STUDY PERIOD 2017-2020 | | | TSAG-TD430 |
| **TSAG** |
| **Original: English** |
| **Question(s):** | | N/A | | Geneva, 10-14 December 2018 |
| **TD** | | | | |
| **Source:** | | ITU-T SG16 Chairman | | |
| **Title:** | | New Question 5/16 (Artificial intelligence-enabled multimedia applications): Outcome of informal consultations | | |
| **Purpose:** | | Action | | |
| **Contact:** | | | Noah Luo Huawei Technologies China | Tel: +44 (11) 8920 8954 Fax: +44 (11) 8920 8900 E-mail: [noah@huawei.com](mailto:noah@huawei.com) |

Following the discussions at the RG-WP session on Tue 11 December concerning the text of new Question 5/16 (Artificial intelligence-enabled multimedia applications) proposed by SG16 as found in [TSAG-TD357](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T17-TSAG-181210-TD-GEN-0357) (Annex A), the SG16 chairman held informal consultations and the proposed amended text found in [Annex A](#AnxA) hereinafter was agreed for input to TSAG RG WP.

Annex A:  
(Amended) draft new Question 5/16 (Q.MM-AI) - Artificial intelligence-enabled multimedia applications

### 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 such technologies 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. In future years, 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 analyzing data at three basic levels: computing layer (base), algorithm layer (technology) and application layer. AI is not just “tech for tech”. Where large data sets are combined with powerful enough 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 video, 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, but focused studies are far behind. Emerging technologies brings not only new opportunities, but also new challenges as well as new demands. Taking multimedia data as an example, image, video and sound data are the fuel of AI applications such as recognition, sentiment classification, etc. However, huge volume multimedia data does not indicate high quality labelling data that AI applications could benefit. 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, which results in huge resource waste and prevent the data flow, which severely hindered the development of AI industry.

This Question focusses on artificial intelligence-enabled multimedia applications, 1) to identify challenges facing the deployment of AI-enabled multimedia applications, 2) to analyze 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 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, such as video surveillance, content screening, image recognition etc.;
* Data preparation for use with AI-enabled multimedia applications;
* Specific system characteristics for AI-enabled multimedia applications;
* Assessment and evaluation techniques for AI-enabled service platforms, such as intelligent speech, natural language processing, machine translation, deep-learning based face recognition and verification, etc.;
* 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 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.

### 4 Relationships

**Recommendations**

* F.700-series

Questions

* All questions of Study Group 16

Study Groups

* SG12, SG13, SG15, SG17, SG20

Other bodies

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

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_