

# Recommendation **ITU-T F.748.23 (04/2024)**

SERIES F: Non-telephone telecommunication services

Multimedia services

---

## **Requirements and framework for intelligent crowdsensing multimedia interaction based on deep learning**

ITU-T F-SERIES RECOMMENDATIONS  
**Non-telephone telecommunication services**

|                                                                 |                    |
|-----------------------------------------------------------------|--------------------|
| TELEGRAPH SERVICE                                               | F.1-F.109          |
| Operating methods for the international public telegram service | F.1-F.19           |
| The gentex network                                              | F.20-F.29          |
| Message switching                                               | F.30-F.39          |
| The international telemesssage service                          | F.40-F.58          |
| The international telex service                                 | F.59-F.89          |
| Statistics and publications on international telegraph services | F.90-F.99          |
| Scheduled and leased communication services                     | F.100-F.104        |
| Phototelegraph service                                          | F.105-F.109        |
| MOBILE SERVICE                                                  | F.110-F.159        |
| Mobile services and multideestination satellite services        | F.110-F.159        |
| TELEMATIC SERVICES                                              | F.160-F.399        |
| Public facsimile service                                        | F.160-F.199        |
| Teletex service                                                 | F.200-F.299        |
| Videotex service                                                | F.300-F.349        |
| General provisions for telematic services                       | F.350-F.399        |
| MESSAGE HANDLING SERVICES                                       | F.400-F.499        |
| DIRECTORY SERVICES                                              | F.500-F.549        |
| DOCUMENT COMMUNICATION                                          | F.550-F.599        |
| Document communication                                          | F.550-F.579        |
| Programming communication interfaces                            | F.580-F.599        |
| DATA TRANSMISSION SERVICES                                      | F.600-F.699        |
| <b>MULTIMEDIA SERVICES</b>                                      | <b>F.700-F.799</b> |
| ISDN SERVICES                                                   | F.800-F.849        |
| UNIVERSAL PERSONAL TELECOMMUNICATION                            | F.850-F.899        |
| ACCESSIBILITY AND HUMAN FACTORS                                 | F.900-F.999        |

*For further details, please refer to the list of ITU-T Recommendations.*

## Recommendation ITU-T F.748.23

### Requirements and framework for intelligent crowdsensing multimedia interaction based on deep learning

#### Summary

Artificial intelligence (AI) techniques can greatly improve the efficiency and effectiveness of crowdsensing tasks execution and enable intelligent multimedia interaction in crowdsensing. Recommendation ITU-T F.748.23 outlines specific scenarios for crowdsensing multimedia interaction that leverage AI techniques, and subsequently defines the corresponding requirements and framework in detail.

#### History \*

| Edition | Recommendation | Approval   | Study Group | Unique ID          |
|---------|----------------|------------|-------------|--------------------|
| 1.0     | ITU-T F.748.23 | 2024-04-26 | 16          | 11.1002/1000/15608 |

#### Keywords

AI, crowdsensing, framework, multimedia data, requirements.

---

\* To access the Recommendation, type the URL <https://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID.

## FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

## NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

## INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2024

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

## Table of Contents

|                                                                                        | Page |
|----------------------------------------------------------------------------------------|------|
| 1 Scope.....                                                                           | 1    |
| 2 References.....                                                                      | 1    |
| 3 Definitions .....                                                                    | 1    |
| 3.1 Terms defined elsewhere .....                                                      | 1    |
| 3.2 Terms defined in this Recommendation.....                                          | 2    |
| 4 Abbreviations and acronyms .....                                                     | 2    |
| 5 Conventions .....                                                                    | 2    |
| 6 Overview of intelligent crowdsensing multimedia interaction based on AI .....        | 2    |
| 7 Scenarios and requirements for AI-based intelligent crowdsensing data interaction... | 4    |
| 7.1 Scenarios and requirements for crowdsensing task assignment.....                   | 5    |
| 7.2 Scenarios and requirements for efficient aggregation of multimedia data.....       | 9    |
| 7.3 Scenarios and requirements for crowdsensing data quality optimization .....        | 11   |
| 7.4 Requirements for multimedia crowdsensing data privacy enhancement .....            | 14   |
| 8 Framework of crowdsensing multimedia interaction based on AI.....                    | 15   |
| 8.1 Centralized crowdsensing platform layer .....                                      | 15   |
| 8.2 Crowdsensing resource layer.....                                                   | 17   |
| 8.3 AI capabilities.....                                                               | 17   |
| Bibliography.....                                                                      | 19   |



# Recommendation ITU-T F.748.23

## Requirements and framework for intelligent crowdsensing multimedia interaction based on deep learning

### 1 Scope

This Recommendation specifies the requirements and framework for intelligent crowdsensing multimedia interaction based on artificial intelligence (AI).

The scope of this Recommendation is proposed to include the following aspects:

- General overview of intelligent crowdsensing multimedia interaction base on AI.
- Scenarios and requirements for crowdsensing multimedia interaction based on AI.
- Framework of crowdsensing multimedia interaction based on AI.

### 2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T F.748.21] Recommendation ITU-T F.748.21 (2022), *Requirements and framework for feature-based distributed intelligent systems*.

[ITU-T Y.4217] Recommendation ITU-T Y.4217 (2022), *Service requirements and capability framework for Internet of things-related crowdsourced systems*.

### 3 Definitions

#### 3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

**3.1.1 crowdsensing** [b-ITU-T Y.4205]: The practice of employing crowdsourcing in collecting sensory data.

**3.1.2 crowdsourcing participant** [b-ITU-T Y.4205]: The person contributing to a crowdsourced system by providing and potentially benefitting from ideas, content, or access to services, hardware or other system resources. In the general case, the participant is a member of a large, open and potentially undefined group of people.

**3.1.3 crowdsourcing resource** [b-ITU-T Y.4205]: The device, hardware module, service or other system resource that a crowdsourced system has been given access to by crowdsourcing participants.

**3.1.4 crowdsourcer** [b-ITU-T Y.4205]: The initiator of a crowdsourcing task. It can refer to a single person, a group of people, a corporation, an organization, etc.

**3.1.5 crowdsourced systems** [b-ITU-T Y.4205]: Systems that employ crowdsourcing in order to augment their constituent infrastructure and the set of provided services or collected information.

**3.1.6 artificial intelligence (AI)** [b-ETSI GR ENI 004]: Computerized system that uses cognition to understand information and solve problems.

**3.1.7 machine learning** [b-ITU-T Y.3172]: Processes that enable computational systems to understand data and gain knowledge from it without necessarily being explicitly programmed.

### **3.2 Terms defined in this Recommendation**

None.

## **4 Abbreviations and acronyms**

This Recommendation uses the following abbreviations and acronyms:

|     |                               |
|-----|-------------------------------|
| AI  | Artificial Intelligence       |
| DL  | Deep Learning                 |
| GPS | Global Positioning System     |
| IoT | Internet of Things            |
| ML  | Machine Learning              |
| NLP | Natural Language Processing   |
| OCR | Optical Character Recognition |
| RNN | Recurrent Neural Network      |
| UAV | Unmanned Aerial Vehicle       |
| UE  | User Equipment                |

## **5 Conventions**

In this Recommendation:

- The keywords "is required to" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this document is to be claimed.
- The keywords "is recommended" indicate a requirement which is recommended but which is not absolutely required. Thus, this requirement needs not be present to claim conformance.

## **6 Overview of intelligent crowdsensing multimedia interaction based on AI**

With the development and popularization of the intelligent mobile devices, wearable devices and Internet of things (IoT) and its built-in sensors (such as accelerometer, gyroscope, camera, compass, global positioning system (GPS), microphone, connected car communication module, etc.), intelligent crowdsensing has gradually developed into a new perception mode. More and more industries require the use of intelligent crowdsensing technology to perceive spatial-temporal information, extract knowledge, and understand multimodal data. During the execution of crowdsensing tasks, crowdsensing technology can leverage the advantages of large scale participants' widespread distribution, flexible mobility, and real-time connectivity. However, the following issues still need to be optimized by artificial intelligence (AI) technologies such as machine learning (ML) or deep learning (DL) when implementing intelligent crowdsensing:

Firstly, in crowdsensing systems, there are usually multiple parallel tasks that need to be assigned. Due to the heterogeneity, complexity, and uncertainty in both the crowdsensing nodes and crowdsourcing participants, it is necessary to introduce AI schemes and jointly optimize the assignment of multi tasks based on the characteristics of varied sensing nodes and participants. This can enhance collaboration among multiple sensing nodes and better integrate comprehensive perception of multimedia data.

Secondly, when executing the crowdsensing tasks, the multimedia data interaction will consume the network resources and incur costs for crowdsourcing participants. In order to save network resources and reduce data transmission costs, it is necessary to carry out efficient aggregation of multimedia data by extracting multimedia key features, filter redundant information through AI techniques on crowdsourcing resources, and transmit key features and parameters instead of all the media content to the crowdsensing system for further processing.

Thirdly, the perceived multimedia data may have the characteristics of multimodality, low quality, information redundancy, and large data volume. If manual screening is used to extract multimedia data information, the crowdsensing tasks executing efficiency will be very low. Therefore, in order to improve the efficiency of crowdsensing tasks, it is necessary to adopt AI techniques and design adaptive frameworks for evaluating the quality of multimedia data, selecting high-quality data, extracting multimedia data features, and efficiently analysing multimedia data.

Lastly, privacy and security is a crucial concern in crowdsensing, as the perception of multimedia data may involve information such as location, image or voice. AI techniques such as data encryption, anonymization and secure transmission can be used to ensure the security and privacy of crowdsensing tasks. Moreover, intelligent mechanisms can be developed to detect abnormal or malicious behaviours, protecting the reliability and effectiveness of data perception, processing, and transmission.

In conclusion, AI schemes can significantly enhance the efficiency and effectiveness of crowdsensing tasks executing, with benefits in task assignment, multimedia data processing, data transmission efficiency, and privacy and security protection. By establishing requirements and a framework for crowdsensing multimedia interactions based on AI, it can offer low-cost, intelligent capabilities for various crowdsensing scenarios and achieve effective crowdsensing results. An overview of intelligent crowdsensing multimedia interaction based on AI is shown in Figure 6-1.