Recommendation

ITU-T H.862.6 (09/2023)

SERIES H: Audiovisual and multimedia systems

E-health multimedia systems, services and applications – Multimedia e-health data exchange services

Functional requirements for counselling services based on artificial emotional intelligence

ITU-T H-SERIES RECOMMENDATIONS

Audiovisual and multimedia systems

CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS	H.100-H.199
INFRASTRUCTURE OF AUDIOVISUAL SERVICES	H.200-H.499
MOBILITY AND COLLABORATION PROCEDURES	H.500-H.549
VEHICULAR GATEWAYS AND INTELLIGENT TRANSPORTATION SYSTEMS (ITS)	H.550-H.599
BROADBAND, TRIPLE-PLAY AND ADVANCED MULTIMEDIA SERVICES	H.600-H.699
IPTV MULTIMEDIA SERVICES AND APPLICATIONS FOR IPTV	H.700-H.799
E-HEALTH MULTIMEDIA SYSTEMS, SERVICES AND APPLICATIONS	H.800-H.899
Personal health systems	H.810-H.819
Interoperability compliance testing of personal health systems (HRN, PAN, LAN, TAN and WAN)	H.820-H.859
Multimedia e-health data exchange services	**H.860-H.869**
Safe listening	H.870-H.879

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

Recommendation ITU-T H.862.6

Functional requirements for counselling services based on artificial emotional intelligence

Summary

Recommendation ITU-T H.862.6 proposes service requirements and functional specifications for counselling services based on artificial emotional intelligence (AEI) technologies. This Recommendation proposes a service model for counselling services using several scenarios. At a time when artificial intelligence (AI) technologies are being widely proposed and used, the relevant standards can be an important opportunity to facilitate the development of the technology in the industry.

---

* 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/.

Table of Contents

Recommendation ITU-T H.862.6

Functional requirements for counselling services based on artificial emotional intelligence

1 Scope

This Recommendation introduces a concept of counselling services based on artificial emotional intelligence, and analyses its common characteristics. This Recommendation also identifies high-level requirements, and presents functional specifications of the services.

The scope of this Recommendation includes:

– Service requirements and future use cases regarding future service needs
– Functional specifications
– Identification of technology challenges.

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.

None.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following term defined elsewhere:

3.1.1 interoperability [b-ITU-T Y.101]: The ability of two or more systems or applications to exchange information and to mutually use the information that has been exchanged.

3.2 Terms defined in this Recommendation

This Recommendation defines the following term:

3.2.1 emotion enabled multimodal user interface: A multimodal user interface (UI) based on emotion analysis with some properties and illustrations and data.

NOTE – Definition adapted from [b-ITU-T H.862.5].

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AEI   Artificial Emotional Intelligence

CBT   Cognitive Behavioural Therapy

IoT   Internet of Things

OTT   Over-The-Top

RWD   Real-World Data

RWE   Real-World Evidence

UI   User Interface

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 need not be present to claim conformance.

–   The keywords "can optionally" and "may" indicate an optional requirement which is permissible, without implying any sense of being recommended. These terms are not intended to imply that the vendor's implementation must provide the option and the feature can be optionally enabled by the network operator/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with the specification.

6   Background

Emotion recognition technology has been carried out through facial recognition, speech recognition, gesture recognition and biometrics based on core technologies of artificial intelligence (AI) such as deep learning, neural networks and big data.

Artificial emotional intelligence (AEI) is a technology that combines emotion recognition, emotion generation and emotion augmentation technologies with AI. The challenges for these technologies can be identified as shown in Table 1.

Table 1 – Technology challenges for AEI

Emotion technologies	Challenges for AEI	Examples
Recognition technology	It can recognize human emotions through facial expressions, voices, texts, gestures, etc.	Multi modal emotion recognition
Generation technology	It can express emotion through avatar's facial expression, voice, text, gesture, etc. through AI.	Lip animation
Augmentation technology	It can strengthen the user's emotions through conversations, multimedia contents, Internet of things (IoT) devices, etc.	Haptic devices

The conceptual diagram of a counselling service using AEI is shown in Figure 1.

This Recommendation proposes requirements for counselling services using artificial emotional intelligence technology.
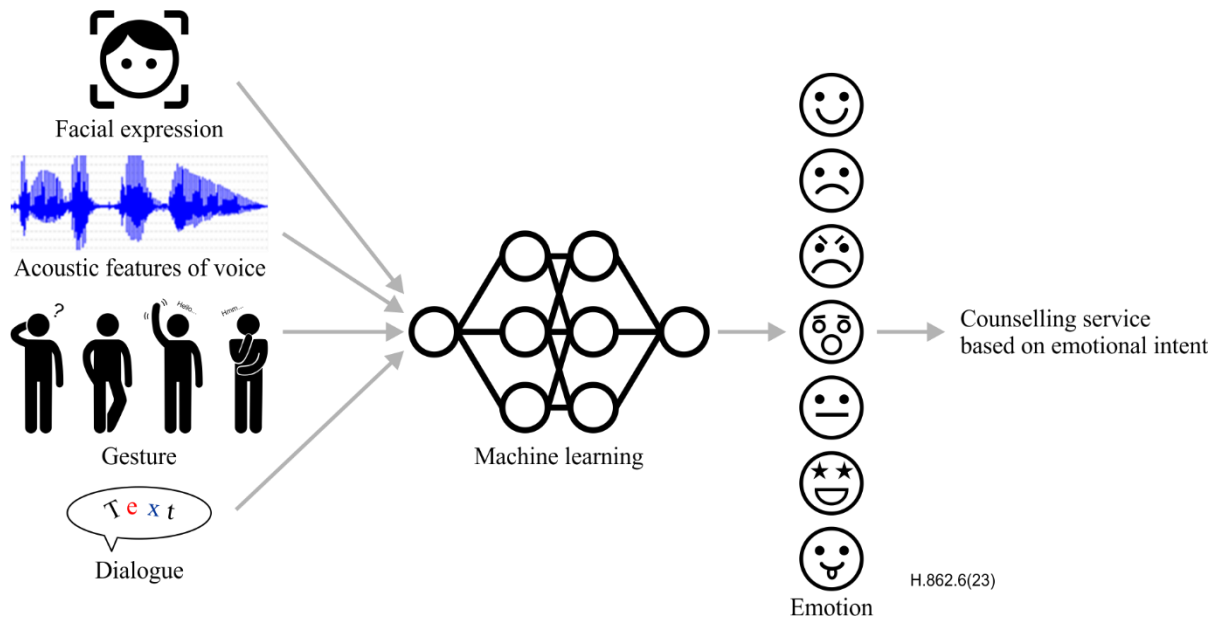
Figure 1 – Conceptual diagram of a counselling service using AEI

7 Services using artificial emotional intelligence

7.1 Counselling services

Artificial emotional intelligence is used in counselling services to recognize various human emotions and sensitivity through biometric information, facial expressions, actions, and voices. By learning the information through an AI algorithm, a personalized emotional ability improvement solution can be presented. It provides advice on appropriate behaviour to users in situations such as interviews through comprehensive analysis of human faces, voices, and gestures. By analysing subtle patterns such as speed and size of voice when talking on the phone, it predicts and informs extreme changes in bipolar mood. It contributes to promoting mutual communication, improving relationships, and restoring relationships by identifying the emotional relationship characteristics of interactions (conversation patterns, language use, etc.) with people around, such as family and friends. When difficult and exhausting emotions are identified from various personal data such as voice, language expression or facial expression, emotional activities such as movies, music or games are recommended to alleviate them. By applying AI technology, it helps to improve emotional and psychological stability and emotional ability by grasping the emotional state through biometric information, facial expressions, images or voice.

7.2 Education services

Emotional synthesis technology has been developed for education services based on deep learning that can express various emotions that can control not only emotions such as sadness and happiness, but also the intensity of emotions, such as "a little sad" or "very sad". By applying cross-language learning technology through foreign language AI learning, mutual learning in the other person's mother tongue is possible. Based on the know-how of college entrance exam experts and AI-based emotional recognition technology, it analyses videos to provide feedback on eye contact, facial expressions and likability, and comprehensively provides cover letters and interview consulting to students. Personalized education for learners can be provided through machine learning-based user emotion recognition technology.

7.3 Customer support services

For customer support services an over-the-top (OTT) service is a content recommendation service based on emotional analysis that recommends video content by analysing emotions with facial

expressions. It recommends content that suits the user's mood such as joy, sadness or anger, after reading the user's facial expression on the "Scanner search for reading my emotions" option.

7.4 Mobility services

In mobility services the core of emotional driving is to provide drivers with optimized interior environment elements such as music, temperature, lighting, vibration, and scent to suit the driver's condition after the vehicle has learnt biometric and current emotional state information based on AI.

7.5 Robot services

The intelligent emotional robot industry, which better understands human emotions through emotional analysis, is expected to have growth potential throughout the entertainment, sales, criminal investigation and social activities sectors.

8 Functional requirements

8.1 Basic requirements

In an emotion-based counselling service, it is important to understand the user's emotion and intention and to sympathize with the conversation.

8.1.1 Emotion recognition

–   An emotion-based counselling service is required to include a method for recognizing one or more emotions in order to understand the user's emotional state. For emotion recognition, emotion expression is sensed based on a single or multiple sensors and recognized through a computer program.

   NOTE – Human emotions are expressed in various forms such as conversation content, voice, facial expression, gesture, and bio-signals. In addition, these emotions have characteristics of each human being, and it is possible to predict the intention of the conversation partner according to the emotions by gathering these emotional expressions.

–   Emotion recognition may be used in counselling services using single or multiple sensors.

Table 2 shows the elements in which emotions are expressed and how to recognize them in general.

Table 2 – Elements for emotional expressions

Data	Sensor	Method
Expression	Camera	Recognition of emotional changes by detecting key points on the face
Conversation	Speaker	Conversation text is classified according to emotional change with a document classifier
Voice	Speaker	Recognition of emotional changes based on phonetic characteristics
Gesture	Camera	Recognition of movement changes expressed according to emotions through the camera
Biosignal	Bio-sensor	Emotion recognition through the electrical signals of the living body that change according to emotions, such as electromyography and brain waves

8.1.2 Distinguish false emotions

The important thing in a counselling service is to be able to identify exact emotions. In the emotional expression criteria, the recognition of emotional expression depends on the accuracy of the recognizer. False emotion is a case in which conversation content and facial expression are

shown with opposite emotions. In this case, it may be necessary to distinguish between the actual user's emotion and the expressed emotion.

8.1.3 Empathy for users

– Counselling services are required to be able to empathize with the emotions of others through emotional communication with users.

8.1.4 Understanding intentions based on emotions

– Intentions appearing in conversations or actions are recommended to be interpreted differently depending on the state of emotion. In order to distinguish between them, it may be necessary to understand the intention according to the emotion.

8.1.5 Monitoring data

Monitoring data refers to data generated during the interaction between the service bot and the user in the counselling service.

– Preservation of previous state with users

Preservation of data from previous interactions with users may provide a way to remember and respond to users. In addition, it may be a basic data for analysing users for effective counselling.

– Analysis of user data

Analysis of data with users may be the basis for effective counselling. Through the analysed data, it is possible to have a conversation tailored to the user's situation.

8.1.6 User avatar

In a counselling service, there may be cases where the user does not want his or her appearance to be exposed other than that the user's actual appearance is shown to the other party or himself. In this case, an avatar in which the user's emotion, mouth shape, and gesture are synchronized may be used.

8.2 Service requirements

In this clause the requirements for a counselling service using emotion recognition technology are specified.

8.2.1 Tracking continuous emotional changes

Tracking not only emotional changes at a single point in time but also long-term emotional changes may provide important predictive data for the user's accurate emotional changes and behavioural changes. Consequently it may be necessary to track continuous changes.

– Emotional change at a specific point in time: A single point in time change is recommended to include a recognition error. It is difficult to track the user's daily emotions and changes.

– Changes over short periods: Integrating emotional changes in a short period rather than emotional recognition at a specific point in time may estimate changes more accurately.

– Long-term emotional change: Through continuous tracking, additional information such as user propensity may be estimated.

8.2.2 Sentiment management

In order to provide customized services to users, it is recommended to record and manage users' emotions.

8.2.3 User interaction

– In order to increase the level of satisfaction with the service, the service is required to be provided according to the user's emotion.

Feedback from the service can trigger new user behaviour. This series of actions is called interaction with the user.

9 Functional model for counselling services

The functional model for the counselling service is composed of functions for receiving user input through the user interface as shown in Figure 2, and for conversation and interaction with the user. In addition, a part for collecting and managing actual use data for utilizing user data can be added.
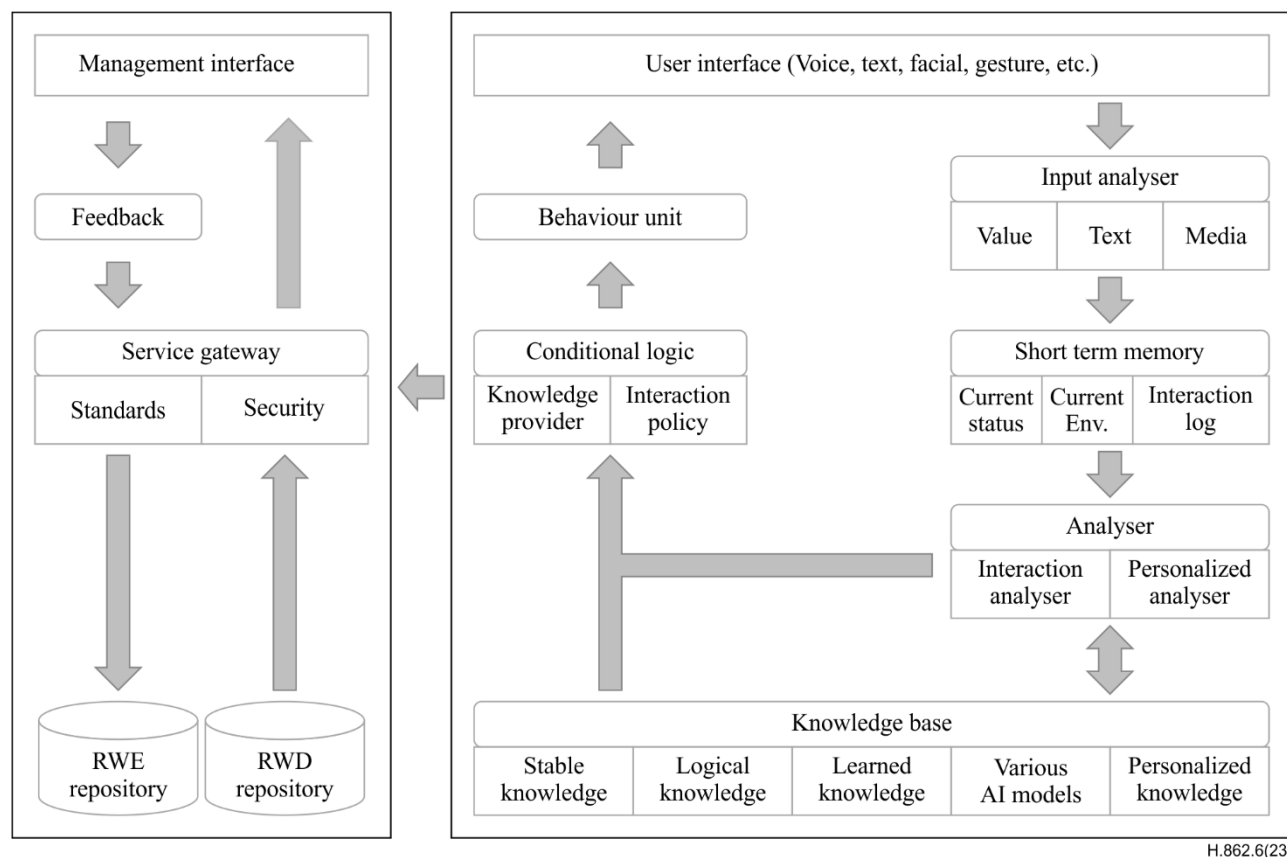

H.862.6(23)

Figure 2 – Functional model

9.1 User interface

The user interface plays a role in receiving various information inputs such as voice, gesture, video, and sensor information from the user and passing it to the input analyser module. In the user interface, data is collected multimodally according to available input methods.

9.2 Input analyser module

A module for analysing user input into data that can be applied for analysis. This module can include the functions shown in Table 3. By analysing the input, it is possible to detect and track changes in the user's current emotional state and behaviour. Functions of the input analyser module are shown in Table 3.

Table 3 – Functions of the input analyser module

Type	Function
Voice	Converting speech to text or extract speech features
Image	Extracting facial feature points or analysing images
Gesture	Motion tracking or action estimation

9.3 Short term memory module

A module for remembering and processing previous interactions with users in order to provide personalized counselling services. In this module, previous conversations and states with the user are memorized, and consultation is performed by reflecting the current input of the user. For example, if you took some medicine yesterday, the counselling chatbot can ask if there is any change in your condition after taking the medicine yesterday.

9.4 Interaction analyser module

A module for analysing the user's intention and predicting the action by utilizing the signal recognition transmitted from the user interface through the user's emotion recognition, gesture recognition, text and sensors.

9.5 Knowledge base

The knowledge base includes knowledge for performing the functions of a counselling chatbot and artificial intelligence models (LLM, VAQ, etc.). It also manages personal information for personalized services.

9.6 Conditional logic module

This counselling chatbot module is configured to respond to the counselee with a module for providing knowledge and an interaction policy, and includes dialog creation and action creation. The interaction policy is intended to improve the user's emotional state according to the user's emotional state. It conducts a conversation according to cognitive behavioural therapy (CBT) methodology. Alternatively, the type of role providing content can be delivered to the behaviour module.

9.7 Behaviour module

The behaviour module determines the movement of the avatar of the counselling chatbot and the operation of the mutually operating sensors according to the counselling status. Avatars on the system can provide a rich user experience through words, facial expressions and gestures.

9.8 Management interface

The management interface refers to the user interface that manages consultation contents. This includes data visualization that is reported to counsellors or psychiatrists. In the management interface, the counselling performance of the user can be summarized by reflecting the feedback from the counsellor or psychiatrist, who is the user. It is also possible to reflect back to the service.

9.9 Real-world data (RWD) repository

The real-world data (RWD) repository refers to a place where logs of mutual interactions between users and counselling chatbots are accumulated. This repository stores the contents of the conversation between the counselee and the counselling chatbot, changes in the user's emotions, and the counselling effect. It can be used to improve the system.

9.10 Real-world evidence (RWE) repository

Data generated in the real-world evidence (RWE) repository is created by feedback from counsellors or psychiatrists. This data is structured and stored so that the counselling results can be utilized. In addition, this data can be used for purposes such as counselling performance management and research.

Bibliography

[b-ITU-T H.862.5]   Recommendation ITU-T H.862.5 (2021), *Emotion enabled multimodal user interface based on artificial neural networks*.

[b-ITU-T Y.101]   Recommendation ITU-T Y.101 (2000), *Global Information Infrastructure terminology: Terms and definitions*.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	**Audiovisual and multimedia systems**
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems