ITUPublications
Recommendations

International Telecommunication Union
Standardization Sector

Recommendation

ITU-T F.746.16 (12/2022)

SERIES F: Non-telephone telecommunication services

Multimedia services

Technical requirements and evaluation methods of intelligence levels of intelligent customer service systems

ITU-T F-SERIES RECOMMENDATIONS

NON-TELEPHONE TELECOMMUNICATION SERVICES

TELEGRAPH SERVICE	
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 telemessage 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	
Mobile services and multidestination satellite services	F.110–F.159
TELEMATIC SERVICES	
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	
Document communication	F.550–F.579
Programming communication interfaces	F.580–F.599
DATA TRANSMISSION SERVICES	F.600–F.699
MULTIMEDIA SERVICES	**F.700–F.799**
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.746.16

Technical requirements and evaluation methods of intelligence levels of intelligent customer service systems

Summary

The intelligent customer service system (ICSS) can provide more convenient, efficient, and stable services for users through the application of artificial intelligence (AI) technologies such as speech recognition, text to speech and natural language processing. There is value therefore in improving and evaluating the intelligence levels of the intelligent customer service system. Recommendation ITU-T F.746.16 specifies the requirements and methods for evaluating the system intelligence of intelligent customer service systems through four aspects which are; the basic functions, the core technologies of AI, the system maturation and the service experience.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T F.746.16	2022-12-14	16	11.1002/1000/15190

Keywords

Intelligent customer service, intelligence levels.

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

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 F.746.16

Technical requirements and evaluation methods of intelligence levels of intelligent customer service systems

1 Scope

This Recommendation aims to establish an indicator system to evaluate the intelligence levels of intelligent customer service systems (ICSS) for human-machine interactions. The intelligent customer service system mentioned in this Recommendation includes a variety of product forms such as outbound service robots, inbound service robots, and text service robots.

This Recommendation considers four main aspects:

– The basic functions of ICSS;

– The artificial intelligence core technical abilities of ICSS;

– The system maturation of ICSS;

– The service experience of enterprises.

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

None.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 intelligent customer service system: A customer service system based on natural language processing, speech recognition, text to speech and other artificial intelligence technologies. After converging the user's questions through single or multiple rounds of conversation, the intelligent customer service system relies on a maintained background knowledge base and interface components to realize intelligent dialogue.

3.2.2 single-round dialogue: A conversation between a human and an intelligent customer service system.

3.2.3 multi-round dialogue: More than one conversation between a human and an intelligent customer service system in the process for solving complex problems.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

APP	Application
BLEU	Bilingual Evaluation Understudy
CRM	Customer Relationship Management
ICSS	Intelligent Customer Service System
KM	Knowledge Management
MOS	Mean Opinion Score
NLG	Natural Language Generation
NLP	Natural Language Processing
NLU	Natural Language Understanding
OCR	Optical Character Recognition
PII	Personally Identifiable Information
SER	Sentence Error Rate
TTS	Text to Speech
WER	Word Error Rate

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.

– 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 Architecture overview

Figure 1 shows the functional structure of the intelligent customer service system.

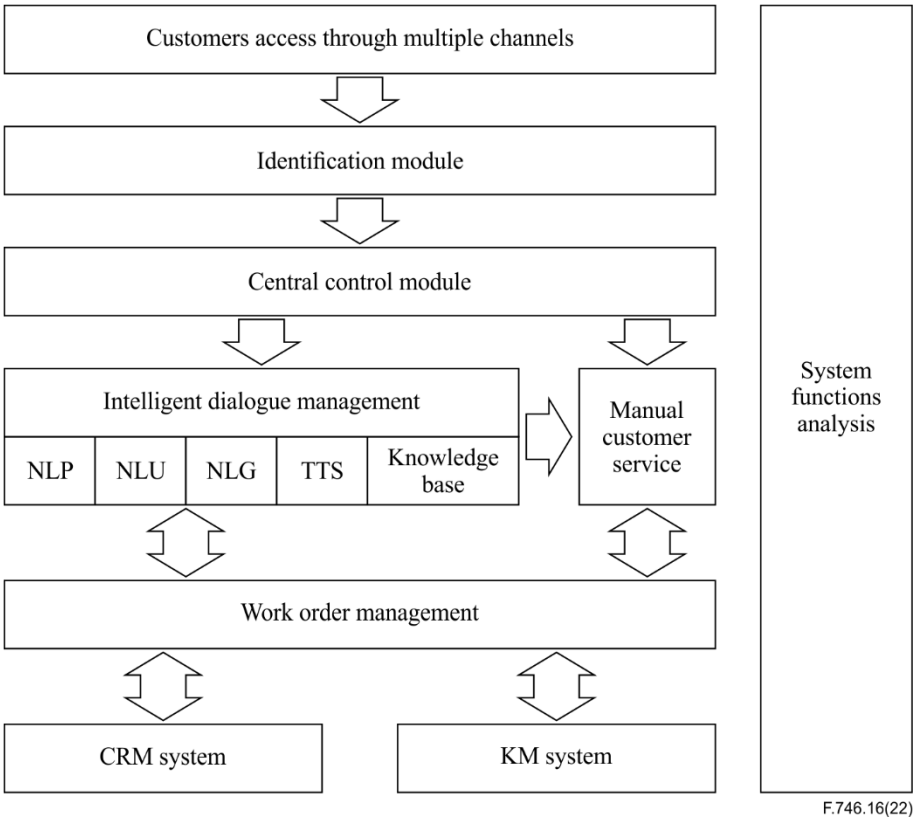

Figure 1 – Functions structure of ICSS

Customers access through multiple channels: This module provides users with web pages, apps, phones and other channels to access the ICSS and obtain relevant services.

Identification module: This module of the ICSS mainly recognizes the user's speech and transforms it into text content for subsequent analysis.

Central control module: This module further process the user's message and converts it into input content that can be recognized and understood by AI, and determines whether the message is answered by the intelligent customer service system or by a manual customer service.

Intelligent dialogue management: This module is to understand the user's intention, match the relevant answers in the corpus, and feed back to the user through text or speech. It mainly involves natural language processing (NLP), natural language understanding (NLU), natural language generation (NLG) and other AI technologies.

Manual customer service: Manual customer service is an important part of the ICSS, which is responsible for dealing with problems that are difficult to solve automatically by the system.

Work order management: When the system encounters problems that are difficult to solve, the content of human-machine conversation will be saved in the form of work orders, and finally the centralized manual processing will be carried out.

CRM system and KM systems: Customer relationship management (CRM) and knowledge management (KM) are auxiliary systems of the ICSS, which can help the ICSS better manage customer relations and business knowledge. CRM system refers to an information system for enterprises to collect, manage, analyse and use customer information by using software, hardware and network technology. KM system is an information system that collects, processes and shares all the knowledge of an organization.

7 Intelligence levels of ICSS

According to the application maturity of involved intelligent technologies, the intelligence ability of ICSS is divided into five levels. The definition and description of each level are shown in Table 1.

Table 1 – Intelligence levels for ICSS

Intelligence levels	Definition	Description
Level-1	No intelligence	The system only has simple basic functions and system maturity, the user experience is not satisfactory.
Level-2	Partial intelligence	Based on some basic functions and system maturity, the system has simple AI core technical capabilities, but its performance and service experience are mediocre, and users can barely accept the system.
Level-3	Conditional intelligence	The system has many basic functions, and the system supports a variety of AI core technical capabilities, and its performance is good. The service capability and system maturity of the enterprise are basically acceptable to users.
Level-4	High intelligence	The system has rich basic functions, and the system supports a variety of AI core technical capabilities, and its performance is excellent. The service ability and system maturity of the enterprise are strong, and the user acceptance is strong.
Level-5	Full intelligence	In addition to the basic functions, the system also has other extended functions, the system supports a variety of AI core technical capabilities, and its performance is excellent. The service capability and system maturity are completely acceptable to users.

8 Basic functions of ICSS

Basic functions include multi-channel management, multimedia message interaction, intelligent recommendation, identity recognition, knowledge base management, work order management, human-machine collaboration, and session abnormal management. The intelligence level description of this dimension can be seen in Table 2.

Table 2 – Intelligence levels for basic function of ICSS

Intelligence levels	Description
Level-1	1) Supports at least one way to access channel. 2) Supports at least one multimedia message type.
Level-2	New relatives to the previous level: 1) Supports at least two ways to access channel. 2) Supports at least two types of multimedia messages. 3) Supports at least one way to identify the identity.
Level-3	New relatives to the previous level: 1) Supports at least three ways to access channel. 2) Supports at least three types of multimedia messages. 3) Supports the intelligent dialogue function, and predict the user's intention according to the user's behaviour track and historical question information.

Table 2 – Intelligence levels for basic function of ICSS

Intelligence levels	Description
	4) Supports at least two ways to identify the identity. 5) Supports knowledge base management.
Level-4	New relatives to the previous level: 1) Supports at least four ways to access channel. 2) Supports at least four types of multimedia messages. 3) Supports intelligent dialogue, and can automatically complete, prompt or correct the entered content in the process of user entering questions. 4) Supports at least three ways to identify the identity. 5) Supports common operations such as adding, deleting, modifying, and querying the knowledge base. 6) Supports work order management functions, including rapid creation of work orders, flexible configuration of work orders, one click to work orders, work order prioritization, and work order status tracking. 7) Supports intelligent customer service such as responding quickly and transferring to manual customer service in time.
Level-5	New relatives to the previous level: 1) Supports work order management functions other than the above functions. 2) Supports intelligent customer service with human-computer cooperation function, and provide process records, problem push and other assistance for manual customer service during the interaction process. 3) Supports abnormal session management.

8.1 Multi-channel management

ICSS is recommended to support websites, APPs, e-mails and other channels.

Evaluation methods:

a) Determine whether the customer service product supports the above channel access.

b) Determine whether the customer service product supports access through channels other than the above channels.

8.2 Multi-media message interaction

Determine whether the interactive methods supported by intelligent customer service are abundant. Customers can use speech, text, pictures, videos, expressions, custom messages and other real-time interactions during communication with the customer service to increase the friendliness and fun.

Evaluation methods:

a) Judge the types of messages supported by the customer service system through material review and technical means, and give corresponding scores.

8.3 Intelligent dialogue

The ICSS recommends related questions based on the user's behaviour tracking, historical consultation and other information. During the interaction process, the system automatically recommends questions related to user consultation, which usually appear in the feedback answers or interaction interface.

Evaluation methods:

a) When the user enters the conversation interface, judge whether the customer service can predict user questions based on historical question records or popular questions, and provide relevant pushes.

b) When the users ask questions, the system will automatically analyse the content that has been entered to provide completion or relevant prompts.

8.4 Identity recognition

This indicator is to evaluate whether the intelligent customer service supports the function of identification. Considering the personally identifiable information (PII) protection, the ICSS can optionally have the function of identifying the user's identity after obtaining the user's permission. For example, the system can use voiceprint recognition, face recognition, short message service (SMS) verification code, voice call and other technologies to determine the user's identity information.

Evaluation methods:

a) Check the smart customer service product information provided by the manufacturer to determine whether the smart customer service supports the identification function.

b) According to the product description, the test system can correctly identify the user's identity.

8.5 Knowledge base management

This indicator is to evaluate the richness and self-learning function of the knowledge base. The ICSS is recommended to include corpus, question database and other related databases.

Evaluation methods:

a) Check the smart customer service product information provided by the manufacturer to determine whether the smart customer service supports the knowledge base management function.

b) Check the richness of customer service knowledge base.

c) Test whether the knowledge base can support self-learning and establish custom questions and answers.

8.6 Work order management

This indicator is to evaluate whether the ICSS has the work order management function. When encountered problems cannot be solved, customer service staff is required to create work orders to support functions such as flexible work order configuration, prioritization, and work order status tracking.

Evaluation methods:

a) Check the smart customer service product information provided by the manufacturer to determine whether the smart customer service supports the work order management function.

b) Check the operating instructions of the work order management system, whether it supports flexible configuration of work order rules, one-click transfer of work orders, provision of hierarchical services, priority sorting, and work order status tracking functions according to the customer's customer service needs.

c) Other functions supported by the work order management system.

For example: The enterprise can set the work orders of users such as VIP and special strategy risks as priority work orders. When customer service personnel of a certain business type receive other business, they can carry out internal transfer orders. Both the internal system and users can understand the progress of the corresponding work order processing.

8.7 Human-machine collaboration

This indicator is to evaluate the ability of the intelligent customer service and the manual customer service to solve customer problems collaboratively. When encountered problems cannot be solved by the customer service system of the machine, it supports transfer to manual service. Then the human service and the customer service system will cooperate to solve the customers' problems.

Evaluation methods:

a) Determine whether there is support for the transfer to manual services.

b) Determine whether the smart customer service supports video calls, voice calls, text chats, etc.

For example: When the machine customer service is not smooth, such as unrecognizable, multiple fuzzy recognition, user dissatisfaction, etc., observe whether the machine customer service responds as always, and judge whether it is a manual combination of machine response.

8.8 Abnormal session management

Determine whether the intelligent customer service can detect abnormal conditions such as abnormal session duration and abnormal connection, and implement corresponding countermeasures.

When abnormalities such as disconnection occur in the session, the ICSS is required to save the previous records, determine the cause of the abnormality, and try to reconnect. When the system contacts the user again, it is recommended to be able to continue the topic before the interruption.

Evaluation methods:

a) Check the session duration to determine whether it is abnormal.

b) When the session is abnormal, check whether there is an attempt to reconnect.

c) After the session is reconnected, check whether the system can continue to record the breakpoint.

9 Artificial intelligence core technical ability of ICSS

This indicator contains three types of algorithms: speech processing, natural language processing (NLP), and image processing. The intelligence level description of this dimension can be seen in Table 3.

Table 3 – Intelligence levels for artificial intelligence core technical ability

Intelligence levels	Description
Level-1	1) Supports speech recognition, but the performance is poor, the user's experience is poor. 2) Supports speech synthesis, and users reluctantly accept the naturalness of synthetic speech. 3) Supports single-round dialogue, but the performance is poor, the user's experience is poor.

Table 3 – Intelligence levels for artificial intelligence core technical ability

Intelligence levels	Description
Level-2	New relatives to the previous level: 1) Supports speech recognition, the system performance is mediocre, the user's experience is flat. 2) Supports speech synthesis, and users basically accept the naturalness of synthetic speech. 3) Supports single-round dialogue and the system performance is mediocre, the user's experience is flat. 4) Supports multiple-round dialogue, but the performance is poor, the user's experience is poor.
Level-3	New relatives to the previous level: 1) Supports speech recognition and the system performance is good. 2) Supports speech synthesis, and users can accept the naturalness of synthetic speech. 3) Supports single-round dialogue, the system performance is good. 4) Supports multiple-round dialogue and the system performance is mediocre, the user's experience is flat. 5) Supports multilingual interaction.
Level-4	New relatives to the previous level: 1) Supports speech recognition and the performance is perfect. 2) Supports speech synthesis, and users can fully accept the naturalness of synthetic speech. 3) Supports single-round dialogue and the performance is perfect. 4) Supports multiple-round dialogue and the performance is good. 5) Supports optical character recognition (OCR) and the performance is good.
Level-5	New relatives to the previous level: 1) Supports multiple-round dialogue and the performance is perfect. 2) Supports OCR and the performance is perfect. 3) Supports picture content recommendation, and the recommended pictures are highly relevant. 4) Supports the ability to construct, use, and maintain a knowledge graph.

9.1 Speech recognition

This indicator is to test whether the intelligent customer service has speech recognition capability. This is the ability to test the accuracy of the intelligent customer service's speech recognition mainly through the technical test method. Through recognition and understanding, the voice signal is converted into corresponding text or commands.

Evaluation methods:

a) Construct a speech data set, obtain the text of speech recognition, and separately count the word accuracy and sentence accuracy of speech recognition.

b) Word error rate (WER): Compare the recognized sequence with the label sequence of characters, insert, replace, and delete the total number of characters, divided by the percentage of the total number of words in the label sequence. Each test data will get a word error rate, and the average value is the word error rate of the test set.

$$WER = \frac{NW_{misrecognized}}{NW_{total}}$$

$$NW_{misrecognized}: the\ number\ of\ misrecognized\ words.$$

$$NW_{total}: Total\ words.$$

c) Sentence error rate (SER): If there is a misrecognition of a word or character in a sentence, the sentence is considered to be misrecognized. The number of sentence recognition errors divided by the total number of sentences is the sentence error rate.

$$SER = \frac{NS_{misrecognized}}{NS_{total}}$$

$$NS_{misrecognized}: the\ number\ of\ misrecognized\ sentences.$$

$$NS_{total}: Total\ sentences.$$

9.2 Text to speech

This indicator is to test whether the intelligent customer service has text to speech capability. This is the ability to convert any text into speech with a high degree of naturalness.

Evaluation methods:

a) After the multi-round call task initiated by the outbound call system under test is completed, the text and audio are extracted.

b) Many testers use the mean opinion score (MOS) indicators (five-point scale) to rate the naturalness of the robot's speech synthesis in audio. The MOS scoring instructions can be seen in Table 4.

Table 4 – MOS scoring instructions

Scores	Levels	Description
5	Excellent	Fluent and natural, can be completely relaxed, no need for attention.
4	Good	The distortion is not obvious and requires attention, but does not need to be clearly concentrated.
3	General	It is clearly distorted and requires a moderate level of attention;
2	Bad	Severe distortion and need for concentration.
1	Very bad	The quality of the speech is so poor that it is difficult to understand even if you try to listen.

c) In the case of artificial recording and speech synthesis audio splicing, the overall fluency is required to be determined and the jumping and non-smooth conditions are required to be recorded.

d) In order to reflect the customizability and diversity of speech synthesis dialogue in the external call process, the situation regarding the richness of timbre and variety is to be recorded.

9.3 Single-round dialogue

Evaluate whether the intelligent customer service product has the capability of single-round conversation. A single round of conversation refers to a simple question and answer. The question can be described in one sentence and does not depend on the context.

Evaluation methods:

a) A single round of dialogue test data set is used to observe the dialogue and question answering effects of intelligent customer service.

b) The recognition situation is divided into: accurate recognition, fuzzy recommendation is correct, fuzzy recommendation is wrong, not recognized.

c) Accurate recognition and fuzzy recommendation are regarded as successful intention recognition. The ratio of successful recognition to total tests is the intention recognition rate. The ratio of accurate recognition to total tests is the accurate recognition rate.

$$Intent\ recognition\ rate = \frac{NT_{true\&vague}}{NT_{total}}$$

$$NT_{true\&vague}: Accurately\ identify\ or\ indistinctly\ recommend\ the\ correct\ quantity$$

$$NT_{total}: Total\ task$$

$$Accurate\ recognition\ rate = \frac{NT_{true}}{NT_{total}}$$

$$NT_{true}: The\ number\ of\ accurate\ identifications.$$

9.4 Multi-round dialogue

This indicator is to evaluate the ability of intelligent customer service products to support multiple rounds of dialogue. In the customer service scenario, many problems cannot be solved by simple question and answer in a single round of conversation. When the user's needs are more complex and may need to be presented in multiple rounds of conversation, the machine can help the user find satisfactory results by asking, clarifying, or confirming. Such a round-trip interactive session is called the multi-round dialogue.

Evaluation methods:

a) Use multiple rounds of dialogue to test the data sets. Before the intelligent customer service product gives the user accurate results, the system will ask diverging questions, add details and repeatedly ask questions based on the results to see whether it can be accurately identified.

b) Multiple rounds of test data sets covering 2-4 rounds.

c) All rounds of multiple dialogues were successfully identified as task completion, and the ratio of test data volume to task completion rate is given by:

$$Task\ completion\ rate = \frac{T_{completed}}{T_{total}}$$

$$T_{completed}: the\ number\ of\ completed\ tasks.$$

$$T_{total}: the\ total\ number\ of\ tasks.$$

d) The ratio of the number of successful rounds identified in the multi-round test to the total number of test rounds is the recognition rate:

$$Recognition\ rate = \frac{T_{co-i}}{T_{total}}$$

$$T_{co-i}: The\ number\ of\ correctly\ identified\ multi-round\ dialogues.$$

9.5 Knowledge graph

The purpose of this indicator is to evaluate whether the ICSS has the ability to construct, use, and maintain a knowledge graph. The ICSS can optionally have the ability to construct a knowledge graph, and can improve the functions lacking in the traditional knowledge base.

Evaluation methods:

a) Determine whether the ICSS has the ability to build a knowledge graph through material reviews.

b) Test the ability of ICSS to build, generate and maintain knowledge graph through technical means.

c) Comprehensive evaluation of the construction, generation and maintenance of knowledge graphs and other indicators to determine whether customer service products have integrated knowledge graph technology.

9.6 Multilingual interaction

This indicator is to evaluate whether the ICSS has the ability to interact in multiple languages. The intelligent customer service can optionally have the ability of multilingual interaction, that is, the function of bilingual conversion can be automatically performed by machine translation.

Evaluation methods:

a) Judge whether the ICSS is capable of multi-language interaction through material reviews.

b) Test the intelligibility of multilingual interaction of the ICSS and bilingual evaluation understudy (BLEU) score through test corpus.

9.7 Optical character recognition

The purpose of this indicator is to evaluate whether the ICSS has the capability of text image recognition. The ICSS can optionally have intelligent visual capabilities such as OCR to assist the ICSS in receiving and processing multi-modal user input.

Evaluation methods:

a) Judge whether the ICSS has the capability of image and text recognition through material review.

b) Through the building of text image data to test whether the ICSS has text image recognition capability.

$$Accuracy\ rate = \frac{Ima_{correctly}}{T_{total}}$$

$$Ima_{correctly}: The\ number\ of\ correctly\ identified\ images.$$

9.8 Picture content recommendation

This indicator is to evaluate whether the ICSS has the ability to recommend related content of pictures. The ICSS can optionally identify pictures and recommend the same or similar items in specific fields such as e-commerce.

Evaluation methods:

a) Judge whether the ICSS has the ability to recommend related pictures through material review.

b) Test whether the ICSS has the ability to recommend related contents of pictures by building image data, and score the relevancy of pictures on a scale of 0-5.

10 ICSS system maturity

Indicators of maturation of the ICSS system include data visibility, usability, reliability, security, maintainability, and auditability. The intelligence level description of this dimension can be seen in Table 5.

Table 5 – Intelligence levels for ICSS system maturation

Intelligence levels	Description
Level-1	1) The usability performance is poor, the user's experience is poor. 2) The security performance is poor, the user's experience is poor.
Level-2	New relatives to the previous level: 1) The usability performance is general, and the user experience is basically acceptable. 2) The reliability performance is general. 3) The security performance is good, and the user experience is basically acceptable. 4) Supports auditability.
Level-3	New relatives to the previous level: 1) The usability performance is good. 2) The reliability performance is good. 3) The security performance is perfect. 4) Supports data visibility function, which can generate visual reports for business data in real time.
Level-4	New relatives to the previous level: 1) The usability performance is excellent. 2) The reliability performance is excellent, such as weather and news inquiries. 3) Supports auditability. 4) Supports automatic operation and maintenance. 5) Supports data visibility function, and the system can automatically generate analysis results and documents and export them.
Level-5	New relatives to the previous level: 1) Supports operation monitoring, and the system can quickly locate system operation problems.

10.1 Data visibility

Enterprises need to judge the effect of intelligent customer service through visual analysis. The ICSS is required to have a report analysis function to help enterprises improve their operating efficiency.

Evaluation methods:

a) Generate visual reports.

b) Analyse according to the report.

c) Export reports and data.

10.2 Usability

Evaluate whether the ICSS is easy to use. The ICSS is required to obey the principles of user friendliness and ease of use.

Evaluation methods:

a) The reply message is required to be completed within a certain time, and there is a prompt such as "reply in customer service".

b) User-centred design. Used smoothly in a set of business processes in a guided manner.

c) Interface design can optionally fully consider the user experience, friendly and clear, simple operation, obvious function points. The main functions are required to be in a prominent position, and the commonly used functions are required to be in a position easy to operate. Remove the redundant steps and get to the result.

10.3 Reliability

Evaluate whether the ICSS is mature and reliable. A mature and stable system is very important for the continuous operation of the ICSS.

Evaluation methods:

a) Stable and reliable mature technology is required to be adopted to ensure the long-term safe operation of the system.

b) The hardware, software and information resources in the system can optionally meet the requirements of reliability design.

c) The system has fault-tolerant ability, which makes management and maintenance more convenient.

d) Unified planning and analysis of network design, selection, installation, debugging and other links to ensure reliable operation of the system.

e) The system provides multi-level security control means, establishes a perfect security management system, guarantees data security and PII protection and integrity, and simultaneously has reliable anti-virus measures.

f) System failure recovery capability.

g) Log management function.

h) System disaster recovery backup.

10.4 Security

Evaluate the security of the ICSS. While providing intelligent services, intelligent customer service products must also ensure data security and legality of dialogues.

Evaluation methods:

a) User program security and PII protection, including user permissions, password management, and so on.

b) System network security, including attack protection, vulnerability detection, Trojan killing, and so on.

c) Database security, including backup, data encryption, and so on.

d) Transmission security, including data encryption transmission, and so on.

e) Content security, including disabled word filtering, image filtering, and so on.

f) Deployment of security, including the separation of the two networks, export protection, wall protection, and so on.

10.5 Maintainability

Evaluate the maintainability of the intelligent customer service system. The ICSS is required to have complete maintenance tools.

Evaluation methods:

a) Whether to have automatic operation and maintenance technology.

b) Whether there is operation monitoring.

10.6 Auditability

Evaluate whether the ICSS is auditable. The operation data of the ICSS is required to be able to be checked, reviewed and traced.

Evaluation methods:

a) Supports personnel authority management.

b) Supports operation audit.

11 Service experience of enterprises

Service experience of enterprises evaluation includes data access ability, human intervention ability, emotional pacification ability, assistant service ability, customer satisfaction. The intelligence levels description of this dimension can be seen in Table 6.

Table 6 – Intelligent levels for service experience of enterprises

Intelligence levels	Description
Level-1	1) Supports access to user information, and provides targeted services for users. 2) Supports manual intervention to deal with emergencies. 3) The user is not satisfied, and the total user experience score is [0, 60).
Level-2	New relatives to the previous level: 1) Supports the manual intervention of the single-round dialogue, and the manual intervention rate is high. 2) The user is basically satisfied, and the total user experience score is [60, 70).
Level-3	New relatives to the previous level: 1) Supports data access ability, and the data access rate is high. 2) Supports the manual intervention of the single-round dialogue, and the manual intervention rate is acceptable. 3) Supports the manual intervention of the multi-round dialogue, with a high rate of manual intervention. 4) The user is relatively satisfied, and the total user experience score is [70, 80).
Level-4	New relatives to the previous level: 1) Supports the manual intervention of multiple rounds of dialogue, and the manual intervention rate is acceptable. 2) Supports assistant service ability. 3) The user is very satisfied, and the total user experience score is [80, 90).
Level-5	New relatives to the previous level: 1) Supports emotional pacification ability, and switches to manual customer service for pacification when necessary. 2) The user is completely satisfied, and the total user experience score is [90, 100].

11.1 Data access ability

The purpose of this indicator is to evaluate whether smart customer service products can access user information and access capabilities. The intelligent customer service has data access capabilities, which can provide users with a more convenient and safe service experience.

Evaluation methods:

a) Businesses that need access to data build a test suite to observe data access.

b) The data access rate is the number of successful implementations divided by the number of tests.

$$Data\ access\ rate = \frac{T_{successful}}{T_{total}}$$

$$T_{successful}: successful\ data\ access.$$

11.2 Human intervention ability

This indicator is to evaluate the manual intervention of the ICSS and in order to judge the autonomous business capabilities of the intelligent customer service. In the process of providing services to users, the ICSS needs to rely on the intervention of manual customer service for specific items such as puzzles, serious questions, and user's mood swings.

Evaluation methods:

a) Count the number of manual interventions during the single-round dialogue and multi-round dialogue testing.

$$Manual\ intervention\ rate = \frac{T_{manual}}{T_{total}}$$

$$T_{manual}: the\ number\ of\ human\ interventions.$$

11.3 Emotional pacification ability

This indicator is to evaluate whether the ICSS can identify the user's mood swings and provide timely and effective pacification. Emotional pacification is a high-level demand for an ICSS, which requires the system to correctly determine the user's emotions through a series of elements such as speech, tone, wording, punctuation and tone words during the interaction with the user. When recognizing the user's emotions such as expressions of anger, it can respond and comfort in time, and if necessary, it can switch to manual customer service for pacification.

Evaluation methods:

a) A corpus of anger and requests for artificial service was designed to evaluate the capability of providing emotional consolation.

11.4 Assistant service ability

This indicator is to assess whether the ICSS provides assistant services such as weather inquiries and news inquiries. The intelligent customer service needs to provide related services such as weather inquiries and news inquiries while providing business-related services.

Evaluation methods:

a) Determine whether the ICSS has the related business ability to provide assistant services through material review.

b) Build the command corpus related to weather inquiry, news inquiry, etc., to test whether the ICSS has relevant abilities. The service ability of assistants is rated according to 0-5 score.

11.5 Customer satisfaction

This indicator is to test the effect of the user experience of the ICSS. The quality of the user experience directly determines whether the intelligent customer service can play the role of respondent on a certain product or on certain types of products. User experience is a purely subjective feeling established by users during the use of products. For any intelligent customer service product, the user experience comprehensively reflects the subjective experience of the system in terms of channel management, technical effects, and interface design. In the context of user-centred and people-oriented product design concepts, user experience has become a key indicator for evaluating the merits of human-computer interaction systems.

Evaluation methods:

a) Through the investigation and trial of the ICSS, the user experience effect is assessed. The scoring indicators are as follows:

b) Useful: The ICSS is required to be useful and actually solve the user's problems and confusion.

c) Findable: Good navigation and positioning of elements is required to be provided so that users can quickly find the interactive entry point of intelligent customer service.

d) Accessible: The basic and common functions of the intelligent customer service are required to be accessible to all users, including members and non-members.

e) Desirable: Intelligent customer service is required to satisfy users' various emotional experiences, including emotional comfort, etc.

f) Credible: Intelligent customer service replies can let the user trust the system, no false forged information.

g) Each of the five specific indicators above gets 20 points, and the sum of the five items S is the total score of user experience.

h) The total user experience score is 4 points, $S \in [0, 60)$ no score, $S \in [60, 70)$ 1 score, $S \in [70, 80)$ 2 score, $S \in [80, 90)$ 3 score, $S \in [90, 100]$ 4 score.

i) Researchers questioned: The rating group shall be composed of a variety of social roles and professionals and shall exceed 100 persons; Each person will score the ICSS independently, and the average score will be the final score of each system.

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

Published in Switzerland
Geneva, 2023