

Recommendation **ITU-T F.748.46 (03/2025)**

SERIES F: Non-telephone telecommunication services

Multimedia services

Requirements and evaluation methods of artificial intelligence agents based on large scale pre-trained models

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
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.748.46

Requirements and evaluation methods of artificial intelligence agents based on large scale pre-trained models

Summary

Recommendation ITU-T F.748.46 provides a framework of requirements and evaluation methods for AI agents based on pre-trained large models. It attempts to help enterprises and users understand AI agent technology and help them build and implement products and applications faster and more accurately. Furthermore, this Recommendation is intended to promote the healthy development of the industry.

History *

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T F.748.46	2025-03-16	21	11.1002/1000/16236

Keywords

AI agent, evaluating methods, technical 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, and 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 <https://www.itu.int/ITU-T/ipr/>.

© ITU 2025

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.....	1
4 Abbreviations and acronyms	1
5 Conventions	2
6 Framework of requirement and evaluation methods for artificial intelligence agents based on large scale pre-trained model	2
7 Requirements and evaluation methods of artificial intelligence agents	3
7.1 Perception and cognition	3
7.2 Planning capability	8
7.3 Memory capability.....	10
7.4 Execution capability	11
Appendix I – Formulae for common indicators.....	14
Bibliography.....	16

Recommendation ITU-T F.748.46

Requirements and evaluation methods of artificial intelligence agents based on large scale pre-trained models

1 Scope

This Recommendation mainly focuses on AI agents based on large scale pre-trained models. It attempts to establish technical requirements and evaluation methods for intelligent products and applications to measure their performance. The capabilities involved in this Recommendation include four aspects:

- **Perception and cognition:** This comprises of recognition, comprehension, dialogue, generation, and reasoning.
- **Planning:** This includes task planning, task scheduling and task optimization.
- **Memory:** This includes short-term memory and long-term memory.
- **Action:** This includes virtual environment execution capability and real-world environment execution capability.

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 terms defined elsewhere:

3.1.1 AI agent [b-ISO/IEC 22989]: Automated entity that senses and responds to its environment and takes actions to achieve its goals.

3.1.2 artificial intelligence [b-ISO/IEC 2382]: An interdisciplinary field, usually regarded as a branch of computer science, dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AI	Artificial Intelligence
CV	Computer Vision
SCR	Sentence Correct Rate

SER Sentence Error Rate
WCR Word Correct Rate
WER Word Error Rate

5 Conventions

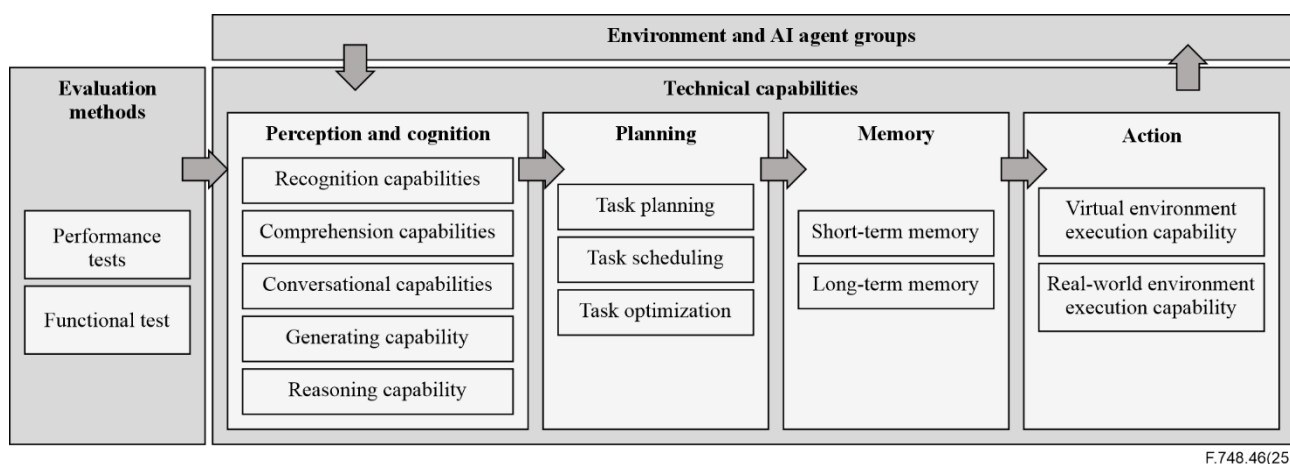
In this Recommendation:

- The keywords "**is required to**" indicate a requirement that must be strictly followed and from which no deviation is permitted if conformance to this document is to be claimed.
- The keywords "**is recommend to**" indicate a requirement that is recommended but which is not absolutely required. Thus, this requirement need not be present to claim conformance.
- The keywords "**optional**" indicate an optional requirement that 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 Framework of requirement and evaluation methods for artificial intelligence agents based on large scale pre-trained models

The AI agents based on the pre-trained large models are required to have the following four capabilities: perception and cognition, planning, memory and execution capability. In practical applications, the intelligent system is recommended to consider the interaction of the external environment and the agent groups. Based on the above, this Recommendation divides the overall technical structure into two modules, one focusing on the technical capabilities of the single AI agent and the other focusing on the interaction between the AI agent and the external elements.

This Recommendation evaluates the capability levels of AI agents in two aspects: performance superiority and functional completeness.



F.748.46(25)

Figure 1 – Framework of requirement and evaluation methods for artificial intelligence agents based on large scale pre-trained models