ITUPublications
Recommendations

International Telecommunication Union
Standardization Sector

Recommendation

ITU-T F.748.48 (03/2025)

SERIES F: Non-telephone telecommunication services

Multimedia services

Technical specification for artificial intelligence cloud platforms – Artificial intelligence model deployment

Documents
Document Sharing Hub

Full document details: ITU-T F.748.48 2025

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 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	F.110-F.159
Mobile services and multidestination 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.

AVA Documents
Document Sharing Hub

Full document details: ITU-T F.748.48 2025

Recommendation ITU-T F.748.48

Technical specification for artificial intelligence cloud platforms – Artificial intelligence model deployment

Summary

Recommendation ITU-T F.748.48 specifies a framework for the cloud-based deployment of artificial intelligence (AI) models. It covers the functional requirements of AI cloud platforms to enable the deployment of AI models. It establishes the technical specifications of the platform's supporting functional modules, core functional modules and auxiliary functional modules.

History*

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T F.748.48	2025-03-16	21	11.1002/1000/16238

Keywords

AI, compression, model deployment, monitoring, optimization, requirement.

* 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.

Documents Document Sharing Hub

Full document details: ITU-T F.748.48 2025

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.

Documents
Document Sharing Hub

Full document details: ITU-T F.748.48 2025

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 AI model deployment	2
7	Requirements of AI model processing	3
7.1	Model conversion	3
7.2	Model optimization	4
7.3	Model evaluation	4
8	Requirements of deployment management	5
8.1	Cloud deployment	5
8.2	End-edge deployment	6
Appendix I – An example of AI model deployment on cloud		8
I.1	Model preparation	8
I.2	Model optimization	8
I.3	Model evaluation	8
I.4	Cloud deployment	8
I.5	Monitoring and maintenance	9
Appendix II – An example of AI model deployment on end device		10
II.1	End device selection	10
II.2	Model preparation	10
II.3	Model conversion	10
II.4	End device deployment	10
Bibliography		11

Documents
Document Sharing Hub
Full document details: ITU-T F.748.48 2025

Documents
Document Sharing Hub

Full document details: ITU-T F.748.48 2025