

Recommendation **ITU-T F.748.27 (01/2024)**

SERIES F: Non-telephone telecommunication services

Multimedia services

Framework and requirements for the construction of 3D intelligent driven digital human application systems

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.748.27

Framework and requirements for the construction of 3D intelligent driven digital human application systems

Summary

Recommendation ITU-T F.748.27 outlines the framework and requirements for the construction of three dimensional (3D) intelligent driven digital human application systems. With the advancement of modelling, driving, rendering and interactive technologies, an increasing number of new services and applications involving 3D intelligent driven digital humans are emerging. It defines the concept, related terms, and fundamental functions of 3D intelligent driven digital human to specify the framework of 3D intelligent driven digital human application systems, including image generation, speech generation, animation generation, interaction processing, multimodal input and output modules with its specified functions and construction requirements. In addition, the Appendix presents some use cases of the workflow of 3D intelligent driven digital human.

History*

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T F.748.27	2024-01-13	16	11.1002/1000/15761

Keywords

3D intelligent driven digital human, animation generation module, image generation, interactive processing module, multimodal input module, multimodal output module, speech generation module.

* 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 3D intelligent driven digital human application systems	3
7 Framework of 3D intelligent driven digital human application systems.....	3
8 Requirements for the construction of 3D intelligent driven digital human application systems	5
8.1 Image generation	5
8.2 Speech generation.....	6
8.3 Animation generation	6
8.4 Interactive processing.....	7
8.5 Multimodal input	8
8.6 Multimodal output	9
Appendix I – Use cases of workflow of 3D intelligent driven digital human application systems.....	11
I.1 Use case of image generation process for 3D intelligent driven digital human application systems.....	11
I.2 Use case of 3D intelligent driven digital human application systems for news broadcasting	11
I.3 Use case of 3D intelligent driven digital human application systems for intelligent assistant	12
Bibliography.....	13

Recommendation ITU-T F.748.27

Framework and requirements for the construction of 3D intelligent driven digital human application systems

1 Scope

This Recommendation specifies the framework and requirements for the construction of three dimensional (3D) intelligent driven digital human application systems.

The scope of this Recommendation includes:

- Overview of 3D intelligent driven digital human application systems;
- Framework of 3D intelligent driven digital human application systems;
- Requirements of 3D intelligent driven digital human application systems.

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.748.15] Recommendation ITU-T F.748.15 (2022), *Framework and metrics for digital human application systems*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 application system [b-ISO/IEC TR 10032]: A collection of application processes which utilizes the services provided by the human-computer interface, communications facility, and data management system to perform the processing necessary to meet the requirements of the information system.

3.1.2 digital human [ITU-T F.748.15]: A computer application that integrates the technologies of computer graphics, computer vision, intelligent speech and natural language processing. It can be used for digital content generation and human-computer interaction to help improve content production efficiency and user experience.

3.1.3 intelligent driven digital human [ITU-T F.748.15]: The digital human is driven by a computer system to automatically complete a series of actions by technical means, including text-driven, audio-driven and video-driven.

3.1.4 text-to-speech synthesis (TTS) [b-ITU-T P.10]: A TTS process generates a speech signal from text codes. It is usually composed of two parts:

- a language-dependent text processing part (the high-level processing part), which generates from the character string (by reading rules, vocabulary and semantic analysis) a set of phonetic, prosodic, etc., parameters which are used by:
- an acoustical signal generating part, the synthesizer itself which generates the audible speech.