

I n t e r n a t i o n a l T e l e c o m m u n i c a t i o n U n i o n

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

T.701.25

(03/2022)

SERIES T: TERMINALS FOR TELEMATIC SERVICES
User interfaces - Accessibility and human factors

Guidance on the audio presentation of text in videos, including captions, subtitles and other on-screen text

Recommendation ITU-T T.701.25

ITU-T

ITU-T T-SERIES RECOMMENDATIONS
TERMINALS FOR TELEMATIC SERVICES

Facsimile – Framework	T.0–T.19
Still-image compression – Test charts	T.20–T.29
Facsimile – Group 3 protocols	T.30–T.39
Colour representation	T.40–T.49
Character coding	T.50–T.59
Facsimile – Group 4 protocols	T.60–T.69
Telematic services – Framework	T.70–T.79
Still-image compression – JPEG-1, Bi-level and JBIG	T.80–T.89
Telematic services – ISDN Terminals and protocols	T.90–T.99
Videotext – Framework	T.100–T.109
Data protocols for multimedia conferencing	T.120–T.149
Telewriting	T.150–T.159
Multimedia and hypermedia framework	T.170–T.189
Cooperative document handling	T.190–T.199
Telematic services – Interworking	T.300–T.399
Open document architecture	T.400–T.429
Document transfer and manipulation	T.430–T.449
Document application profile	T.500–T.509
Communication application profile	T.510–T.559
Telematic services – Equipment characteristics	T.560–T.619
General multimedia application frameworks	T.620–T.649
User interfaces - Accessibility and human factors	T.700–T.799
Still-image compression – JPEG 2000	T.800–T.829
Still-image compression JPEG XR	T.830–T.849
Still-image compression – JPEG-1 extensions	T.850–T.899

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

Recommendation ITU-T T.701.25

Guidance on the audio presentation of text in videos, including captions, subtitles and other on-screen text

Summary

Audiovisual content (such as video) often contain text, which cannot be easily accessed by a wide section of the audience. While captions/subtitles provide text alternatives to audio elements in audiovisual content, other on-screen text may have various functions. It can be part of the story (as a message written on a piece of paper by one of the characters) or it can provide additional information (such as graphs, emergency alerts or superimposed titles). Complementarily, audio description provides a description of audiovisual content auditorily, including captions/subtitles and other on-screen text if present and are of particular benefit to persons who, for different reasons, cannot access on-screen text. However, some users may only require captions/subtitles and other on-screen text to be made accessible as audio because they already have access to other visual content such as the images.

Recommendation ITU-T T.701.25 provides guidance for audiovisual content producers, distributors and exhibitors on the audio presentation of captions/subtitles and other on-screen text. It acknowledges the relationship with existing access services such as audio description. While considering current implementations, as well as future possibilities suggested by research, and bearing in mind possible trade-offs between quantity and quality, this document positions itself for situations in which various access services coexist and users are given the choice to select those best suited to their needs.

Recommendation ITU-T T.701.25 is twin with the published ISO/IEC TS 20071-25:2017 "Information Technology – User interface component accessibility – Part 25: Guidance on the audio presentation of text in videos, including captions, subtitles and other on-screen text" developed by ISO/IEC JTC1 SC35.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T T.701.25	2022-03-29	16	11.1002/1000/14973

Keywords

Accessibility, audio presentation of text in videos, captions, on-screen text, subtitles.

* 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/>.

© ITU 2022

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	3
5 Conventions	3
6 Framework and process considerations	3
6.1 General	3
6.2 Types of text in videos	4
6.3 Types of audio presentations of text in videos	5
6.4 Creating and delivering audio presentations of text in videos	6
6.5 Synchronization.....	6
6.6 Text-files.....	7
7 Guidance on the audio presentation of text in videos.....	7
7.1 General considerations	7
7.2 Developing the audio presentation of text in videos	8
7.3 Levels of importance	8
7.4 Guidance on identifying the audio presentation of text in videos	9
7.5 Guidance on identifying characters	9
7.6 Guidance on text and speech adjustments	9
Bibliography.....	11

Introduction

Videos are omnipresent in our digital society and are used to inform, educate and entertain audiences. Videos often contain text, which cannot be easily accessed by a wide section of the audience. This text includes captions/subtitles and other on-screen text that is part of the visual content.

Captions/subtitles provide text alternatives to audio elements. Other on-screen text may have various functions. It can be part of the story, as a message written on a piece of paper by one of the characters. It can also provide additional information, such as graphs, emergency alerts or superimposed titles.

Persons who, for different reasons, cannot access on-screen text will benefit from an audio presentation. This oral rendering is often part of audio description (see [b-ITU-T T.701.21], an access service providing a description of audiovisual content auditorily, including captions/subtitles and other on-screen text if present. However, some users may only require captions/subtitles and other on-screen text to be made accessible because they already have access to other visual content such as the images.

This Recommendation provides guidance for video producers, distributors and exhibitors on the audio presentation of captions/subtitles and other on-screen text. It acknowledges the relationship with existing access services such as audio description. While considering current implementations, as well as future possibilities suggested by research, and bearing in mind possible trade-offs between quantity and quality, this document positions itself for situations in which various access services coexist and users are given the choice to select those best suited to their needs.