TECHNICAL SPECIFICATION

ISO/IEC TS 20071-25

First edition
2017-02

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

Technologies de l'information — Accessibilité du composant interface utilisateur —

Partie 25: Lignes directrices sur la présentation audio de texte dans des vidéos, y compris les légendes, sous-titres et autres textes à l'écran

Reference number
ISO/IEC TS 20071-25:2017(E)

© ISO/IEC 2017



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2017, Published in Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword ... **iv**

Introduction ... **v**

1 Scope ... **1**

2 Normative references ... **1**

3 Terms and definitions ... **1**
3.1 General terms ... 2
3.2 Audio-specific terms ... 2
3.3 Language of presentation terms ... 2
3.4 Production terms ... 3

4 Framework and process considerations ... **3**
4.1 General ... 3
4.1.1 Alternate names for the audio presentation of text in videos ... 3
4.1.2 Motivation for the audio presentation of text in videos ... 4
4.2 Types of text in videos ... 4
4.2.1 Captions/subtitles ... 4
4.2.2 Other on-screen text ... 5
4.3 Types of audio presentations of text in videos ... 5
4.3.1 Live and recorded spoken captions/subtitles and spoken on-screen text ... 5
4.3.2 Pre-mixed and receiver-mixed productions ... 6
4.4 Creating and delivering audio presentations of text in videos ... 6
4.4.1 Narrator preparation ... 6
4.4.2 Volume ... 6
4.4.3 Audio quality ... 7
4.4.4 Inclusion of users in creating the audio presentation of text in videos ... 7
4.4.5 Inclusion of users in evaluating the audio presentation of text in videos ... 7
4.5 Synchronization ... 7
4.6 Text-files ... 7

5 Guidance on the audio presentation of text in videos ... **8**
5.1 General considerations ... 8
5.1.1 Different strategies ... 8
5.1.2 User considerations ... 8
5.1.3 Availability across technologies ... 8
5.1.4 Consistency within a video and programme series ... 9
5.2 Developing the audio presentation of text in videos ... 9
5.2.1 Clarity in the audio presentation of text in videos ... 9
5.2.2 Reading/delivering the audio presentation of text in videos ... 9
5.3 Levels of importance ... 9
5.4 Guidance on identifying the audio presentation of text in videos ... 9
5.5 Guidance on identifying characters ... 10
5.6 Guidance on text and speech adjustments ... 10

Bibliography ... **12**

© ISO/IEC 2017 – All rights reserved

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by ISO/JTC1, *Information technology*, Subcommittee SC 35, *User interfaces*.

A list of all parts in the ISO/IEC/TS 20071 series can be found in the ISO website.

© ISO/IEC 2017 – All rights reserved

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 ISO/IEC/TS 20071-21), 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 document 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.

© ISO/IEC 2017 – All rights reserved

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

1 Scope

This document provides recommendations on the audio presentation of captions/subtitles and other on-screen text for use in all type of videos regardless of the language and technology being used to transmit and present the recorded or live video.

This document applies to making captions/subtitles and other on-screen text accessible to users with various needs, including but not limited to people with learning and reading disabilities, people with cognitive disabilities, people who are blind or have low vision, older people, and non-native language speakers. It does not apply to captions/subtitles or other on-screen text whose content is already provided in the soundtrack in a language and a way users can access.

This document provides guidance on spoken captions/subtitles as a stand-alone access service but it also provides guidance on how to integrate spoken captions/subtitles, other spoken on-screen text and audio description, if needed, in different types of videos.

NOTE 1 Extensive guidance on audio description is provided in ISO/IEC/TS 20071-21.

This document does not consider the devices or transmission mechanisms used to deliver and play the content or the audio presentation of text in videos. These devices include, but are not limited to televisions, computers, wireless devices, projection equipment, DVD and home cinema equipment, cinema equipment and other forms of user interfaces technology. Therefore, this document does not consider transcoding files for the various video and audio outputs.

NOTE 2 Technical matters of transmission and distribution are covered by other documents (e.g. MPEG standards and other technical standards such as IEC 62731:2013).

This document acknowledges the various needs and preferences of users, as well as the different approaches to the audio presentation of text in videos.

It applies to audio presentations intended to be heard simultaneously along with the original video.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at http://www.iso.org/obp

© ISO/IEC 2017 – All rights reserved

3.1 General terms

3.1.1
video
combination of audio and visual content presented together in a synchronized manner via ICT

Note 1 to entry: While the visual content is often presented using a screen, it might also be presented via other technologies, e.g. a projected hologram.

3.1.2
programme
complete unit of a recorded or live video

3.1.3
programme category
classification of programmes

Note 1 to entry: Programme categories are not necessarily mutually exclusive.

EXAMPLE Programme categories include documentary, news and information, and drama.

3.2 Audio-specific terms

3.2.1
spoken captions/subtitles
audio captions/subtitles
captions/subtitles that are voiced over the audiovisual content

Note 1 to entry: In this document, the term “spoken captions/subtitles” will be used.

3.2.2
spoken on-screen text
text, other than captions/subtitles, that is voiced over the audiovisual content

3.2.3
audio description
descriptive audio
audiovisual content described in an audio modality

Note 1 to entry: Audio description can also be used to describe locations, directions, and objects.

Note 2 to entry: Audio description can be used to describe sound not easily identified or coming from an unknown source or location.

3.2.4
narrator
person(s) and/or technology which voices the alternative audio information

Note 1 to entry: Alternative audio information includes audio description and/or spoken captions/subtitles and/or spoken on-screen text.

Note 2 to entry: In audio description, narrators are also referred to as describers or audio describers.

3.3 Language of presentation terms

3.3.1
original language
language in which audiovisual content is produced

© ISO/IEC 2017 – All rights reserved

3.3.2
dubbing
secondary audio version of a video produced in a language other than the original language of the video and timed to match the voicing of the original actors

Note 1 to entry: The secondary audio version is lip-synchronized and replaces the original dialogue, which cannot be heard.

3.3.3
voice-over
secondary audio version of a video produced in a language other than the original language of the video which overlaps with the voicing of the original actors

Note 1 to entry: The secondary audio version is not lip-synchronized and does not replace the original dialogue, which can still be heard.

3.3.4
captions/subtitles
transcription or translation of audio content visually presented together with the content

Note 1 to entry: Transcriptions or translations include speech and/or non-speech information.

Note 2 to entry: Transcriptions or translations are often suitable for use as an alternative or a complement to the audio content.

3.4 Production terms

3.4.1
pre-mixed production
process that involves delivering alternative audio information together with the audio stream of the video as one single audio track

Note 1 to entry: Although pre-mixed is used in the definition to refer to the audio mix, it can also refer to a production process in which the original audio (including both the soundtrack and the dialogues), the alternative audio information in the video and the original visual content are delivered together.

Note 2 to entry: Pre-mixed productions are also referred to as broadcast-mixed productions.

3.4.2
receiver-mixed production
process that involves delivering alternative audio information separately from the audio stream of the video and having them mixed in a device controlled by the user

Note 1 to entry: Alternative audio information can be delivered to the user or be downloaded or streamed from the Internet as separate services or as services mixed in different combinations.

4 Framework and process considerations

4.1 General

4.1.1 Alternate names for the audio presentation of text in videos

Terms used in this document vary according to country, language, region, video content and type of text in videos.

They include audio subtitles, spoken subtitles, spoken captions, audio captions, to refer to the audio presentation of captions/subtitles. For the purpose of this document, these terms are synonymous, as defined in Clause 3.

© ISO/IEC 2017 – All rights reserved