INTERNATIONAL STANDARD

ISO/IEC 20071-23

First edition
2018-09

Information technology — User interface component accessibility —

Part 23: Visual presentation of audio information (including captions and subtitles)

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

Partie 23: Présentation visuelle d'informations sonores

Reference number
ISO/IEC 20071-23:2018(E)

© ISO/IEC 2018



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2018

All rights reserved. Unless otherwise specified, or required in the context of its implementation, 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
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword ... **vi**
Introduction ... **vii**
1 Scope ... **1**
2 Normative references ... **1**
3 Terms and definitions ... **1**
4 Framework for the visual presentation of audio information ... **3**
4.1 Purpose ... 3
4.2 Motivation ... 4
4.3 Locations of presentations ... 4
4.4 Modes of presentations ... 4
4.5 Modes of access ... 5
4.6 Modes of display ... 5
4.7 Levels of importance ... 5
4.7.1 General ... 5
4.7.2 Essential information ... 5
4.7.3 Significant information ... 6
4.7.4 Helpful information ... 6
4.7.5 Unhelpful information ... 6
5 Applicability of requirements and recommendations ... **6**
5.1 Predictable audio contents ... 6
5.2 Unpredictable audio contents ... 6
6 Production of visual presentations of audio information ... **7**
6.1 Consideration to match the intended meaning ... 7
6.2 Ease of understanding of visual presentations of audio information ... 7
6.3 Consideration of output devices ... 7
6.4 Verification of visual presentations of audio information with the intended output devices ... 7
6.5 Connecting visual presentations of audio information data with content data ... 7
6.6 Combining multiple visual presentations of audio information ... 8
6.7 Update of visual presentations of audio information data ... 8
6.8 Evaluation ... 8
6.9 Evaluations including contribution of typical users ... 8
7 Visual design ... **8**
7.1 General ... 8
7.2 Personalization ... 8
7.3 Engagement ... 8
7.4 Synchronization of presentations ... 9
7.5 Avoidance of information obstruction ... 9
7.6 Font size ... 9
7.7 Font type ... 10
7.8 Font face ... 10
7.9 Upper, lower, and mixed case letters ... 10
7.10 Contrast and use of colour ... 10
7.11 Speed ... 11
7.12 Number of lines ... 11
7.13 Spacing between characters and lines (kerning and leading) ... 11
7.14 Correct punctuation ... 11
7.15 Spacing between words and phrases ... 12
7.16 Transitions between presentations ... 12
7.17 Sentence segmentation ... 12
7.18 Indication of sentence breaks over multiple visual presentations ... 12
7.19 Additional duration for location change ... 13

© ISO/IEC 2018 – All rights reserved

7.20 Modes of display 13

8 Visual alternative container (VAC) 13
8.1 General 13
8.2 VAC position and area 13

9 Describing speech 13
9.1 Describing verbal content 13
9.2 Grammar 14
9.3 Vulgar verbal content and slang 14
9.4 Language variation 14
9.5 Foreign accents 14
9.6 Indiscernible audio content 14
9.7 Spelling 14
9.8 Abbreviations 15
9.9 Homophones, homonyms, homographs, heteronyms, and heterographs 15
9.10 Long speech 15
9.11 Describing multiple simultaneous information 15
9.12 Confirmation by content producers when producing visual presentations of speech 15
9.13 Sources of information 15

10 Non-speech information (NSI) 16
10.1 General 16
10.2 Describing NSI 16
10.3 Correct description of NSI 16
10.4 Well-known sound descriptions 16
10.5 Onomatopoeia 16
10.6 Sound effects in speech 16
10.7 Censored language 17
10.8 Paralinguistic sound effects 17
10.9 Discrete and sustained sound effects 17
10.10 Confirmation by content producers when producing visual presentations of NSI 17
10.11 Sources of information 17

11 Music 18
11.1 General 18
11.2 Describing presence of music 18
11.3 Describing the reason or purpose for the music 18
11.4 Provide information that identifies the music 18
11.5 Clarification of music descriptions 19
11.6 Presentation of lyrics 19
11.7 Distinction of lyrics from speech 19
11.8 Confirmation by content producers when producing visual presentations of music 20
11.9 Sources of lyrics 20

12 Emotions 20
12.1 General 20
12.2 Describing intended emotional nuance 20
12.3 Describing the reason or purpose for the emotional nuance 20
12.4 Confirmation by content producers when producing visual presentations of audio information of emotions 21

13 Silence 21
13.1 General 21
13.2 Describing intentional silence 21
13.3 Describing prolonged silence 21
13.4 Describing the reason or purpose for the silence 21

14 Identifying speakers 21
14.1 General 21
14.2 Means of identifying speakers 22
14.3 Identifying speakers by word 22

 © ISO/IEC 2018 – All rights reserved

14.4 Identifying speakers by pictogram 22
14.5 Identifying speakers by colour 22
14.6 Identifying speakers by position 22
14.7 Identifying change of speakers by changing position 22
14.8 Multiple visual presentations of audio information 23
15 Evaluating quality of visual presentations of audio information 23
15.1 General 23
15.2 Quality review process 23
Annex A (informative) **Evaluation index and references of guidelines and reports for accessibility 24**
Bibliography 26

© ISO/IEC 2018 – 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 document 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 and IEC 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 Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 35, *User interfaces*.

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

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

© ISO/IEC 2018 – All rights reserved

Introduction

Captions and/or subtitles for audio content provide visual alternatives for audio information in audiovisual content. This document provides requirements and recommendations on the production and design of the visual presentation of audio information (including captions and subtitles) that supports users who cannot make use of the audio content.

The use of this document helps to support universal and inclusive media content production practices. It provides guidance for producers, exhibitors, or distributors of audio content (including the medium of distribution and the medium of delivery) to support the accessibility and usability of visual alternatives of audio content.

Different jurisdictions have different expectations of what belongs in a caption or a subtitle. From the point of view of the user, what is important is receiving the information in an accessible design, not the technological means of its delivery. This information can include text conveying speech, sound information, verbatim transcription of the spoken word content, translations of the spoken word content, etc. This document uses "visual presentations of audio information" to include all audio information needed to be made accessible for some users.

Standardized guidance for producing visual presentations of audio information is important to meet a variety of needs. For example, it is important to recognize acceptable values for specifying typography variables such as the letter size and/or number of characters in visual alternatives that rely on text.

Providing visual presentations of audio information (including captions and subtitles) can be beneficial to all, and in particular to diverse users who cannot hear or understand the audio content in diverse contexts, including: persons with hearing loss, persons who are deaf or hard of hearing, persons with learning difficulties or cognitive disabilities, persons watching a movie in a non-native language, persons who need the content to be in another language, persons who cannot hear the audio content due to environmental conditions, or circumstances where the sound is not accessible (e.g. noisy surroundings), the sound is not available (e.g. muted, no working speakers), or the sound is not appropriate (e.g. a quiet library). Although this guidance acknowledges the need of visual presentations of audio information to provide non-visual presentations for diverse users, it does not include guidance for producing non-visual presentations, such as spoken captions/subtitles (see ISO/IEC TS 20071-25 for further reference) and tactile displays (e.g. Braille). The production, delivery, and exhibition of visual presentations of audio information based on this standard are not intended to interfere with or change the meaning of the audio content.

The production, delivery and exhibition of visual presentations of audio information vary according to the time and methodology of production, the technology used for its production, the system of delivery, and the display (including the brand and model of the display).

© ISO/IEC 2018 – All rights reserved

Information technology — User interface component accessibility —

Part 23: Visual presentation of audio information (including captions and subtitles)

1 Scope

This document provides guidance for producers, exhibitors, and distributors on the visual presentation of alternatives to audio information in audiovisual content, such as captions/subtitles.

This document provides requirements and recommendations that are intended to support users who are not able to use the audio information, prefer to use a visual representation of audio information, or prefer both audio and visual presentations.

NOTE Many users do not have a choice, for instance, when in a noisy environment (e.g. bar, restaurant, etc.). In these situations, the user does not select a visual presentation of audio information but is offered the content with captions/subtitles.

This document acknowledges the various needs and preferences of viewers (end users) as well as the different approaches to visual presentation of audio information. It applies to all presentations of visual alternatives to audio information intended to be presented as captions/subtitles.

This document does not apply to the presentation devices or transmission mechanisms used to deliver the content or visual presentations of audio information. These devices could include, but are not limited to: televisions, computers, wireless devices, projection equipment, DVD and home cinema equipment, video game consoles, and other forms of user interfaces technology. This document does not apply to transcoding files and formats for the various video outputs.

This document gives guidance on visual presentations which are delivered in the same language as in the audio (i.e., intra-lingual captions/subtitles) and visual presentations which are translated into a different language (i.e., inter-lingual captions/subtitles). This document does not apply to the specific process of language translation.

This document helps to improve accessibility. This document does not establish requirements on specific industries (e.g. television broadcasting, motion pictures) nor is it intended to supersede specific international standards within their domain.

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 https://www.iso.org/obp

© ISO/IEC 2018 – All rights reserved

3.1
information
knowledge concerning objects, such as facts, events, things, processes, or ideas, including concepts, that within a certain context has a particular meaning

Note 1 to entry: Although information will necessarily have a representation form to make it communicable, it is the interpretation of this representation (the meaning) that is relevant in the first place.

[SOURCE: ISO/IEC 2382:2015, 2121271]

3.2
content
interactive or non-interactive object containing information represented by text, image, video, sound, or other media

[SOURCE: ISO/IEC/IEEE 23026:2015, 4.6]

3.3
caption/subtitle
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
open caption
open subtitle
caption/subtitle visually presented regardless of user preference

Note 1 to entry: Open captions/subtitles do not include visual elements that are a part of the original video contents.

3.5
closed caption
closed subtitle
caption/subtitle visually presented only in response to user preference

Note 1 to entry: Closed captions/subtitles are usually presented by a specialised device or decoder.

3.6
non-speech information
NSI
part of the audio content, other than spoken words

Note 1 to entry: NSI can convey information about: plot, humour, mood, or meaning of a spoken passage.

EXAMPLE Speaker identification information (e.g. off-screen speakers and multiple on-screen speakers), sound effects, music (e.g. singing, background music, instrumentation), manner of speaking (e.g. whispering, emotion, word emphasis), audience reaction (e.g. laughing, groaning, booing).

3.7
visual alternative container
VAC
opaque or translucent area visually presenting alternative content

Note 1 to entry: While VACs are largely used to provide alternatives to audio content, they can also be used to provide alternatives to other content.

Note 2 to entry: There can be multiple VACs presented at the same time.

© ISO/IEC 2018 – All rights reserved

Note 3 to entry: VACs can be displayed to indicate where the visual presentation of content will appear in the future or has appeared in the past.

EXAMPLE Caption/subtitle-boxes, -stripes or –lines are common examples of VACs.

3.8
audiovisual content
content that includes audio and visual components

Note 1 to entry: Only the audio or the visual components might be active at some times within the presentation of audiovisual content.

3.9
video
combination of audio and visual content presented together in a synchronized manner via Information and Communication Technology

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.

[SOURCE: ISO/IEC TS 20071-25:2015, 2.1.2, modified – Note 1 to entry has been added.]

3.10
content category
<audiovisual> classification of audiovisual content

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

Note 2 to entry: When content category is considered from an artistic perspective, it is often referred to as genre.

EXAMPLE Content categories include: dramas, museum and art gallery exhibits, heritage tours, comedies, documentaries, video users' guides and manuals, university lectures, meetings, sporting events, etc.

3.11
importance
level of need for users to know information in the content

3.12
essential (information)
<importance> information that is necessary for users to understand the content and/or its function

3.13
significant (information)
<importance> information that provides a more detailed understanding of the content for most users most of the time

3.14
helpful (information)
<importance> information that provides a thorough understanding of the content for some users

3.15
unhelpful (information)
<importance> information that does not help users understand the content and/or might interfere with that understanding

4 Framework for the visual presentation of audio information

4.1 Purpose

Visual presentations of audio information should aim at providing viewers with alternative or complementary visual information that meets users' needs and contexts of use (e.g. noisy environments).

© ISO/IEC 2018 – All rights reserved