

DASH-IF IOP-9 V5.0.0 (2022-01)



**DASH-IF Interoperability Points;  
Part 9: Text**



***DASH Industry Forum***

3855 SW 153rd Dr.  
Beaverton, OR 97003 - USA

---

Email : admin@dashif.org

***Important notice***

---

The present document can be downloaded from:  
<http://www.dashif.org/guidelines>

---

# 1 Contents

Intellectual Property Rights .....	4
Foreword.....	4
Modal verbs terminology .....	4
Executive Summary.....	4
Introduction .....	4
1 Scope.....	6
2 References .....	6
2.1 Normative references .....	6
2.2 Informative references .....	6
3 Definition of terms, symbols and abbreviations.....	6
3.1 Terms .....	6
3.2 Symbols .....	7
3.3 Abbreviations.....	7
4 CMAF Media Profiles.....	7
5 Adaptation Set requirements and recommendations .....	8
5.1 Content requirements .....	8
5.1.1 Text tracks.....	8
5.1.2 Video tracks .....	10
5.1.2.1 General .....	10
5.1.2.2 CTA 608/708 .....	10
5.1.3 Chunks and gaps .....	10
5.2 Client recommendations .....	10
<b>Annex A (informative): Change History .....</b>	<b>12</b>

---

# Intellectual Property Rights

## Disclaimer

This is a document made available by DASH-IF. The technology embodied in this document may involve the use of intellectual property rights, including patents and patent applications owned or controlled by any of the authors or developers of this document. No patent license, either implied or express, is granted to you by this document. DASH-IF has made no search or investigation for such rights and DASH-IF disclaims any duty to do so. The rights and obligations which apply to DASH-IF documents, as such rights and obligations are set forth and defined in the DASH-IF Bylaws and IPR Policy including, but not limited to, patent and other intellectual property license rights and obligations. A copy of the DASH-IF Bylaws and IPR Policy can be obtained at <http://dashif.org/>.

The material contained herein is provided on an "AS IS" basis and to the maximum extent permitted by applicable law, this material is provided AS IS, and the authors and developers of this material and DASH-IF hereby disclaim all other warranties and conditions, either express, implied or statutory, including, but not limited to, any (if any) implied warranties, duties or conditions of merchantability, of fitness for a particular purpose, of accuracy or completeness of responses, of workmanlike effort, and of lack of negligence.

In addition, this document may include references to documents and/or technologies controlled by third parties. Those third-party documents and technologies may be subject to third party rules and licensing terms. No intellectual property license, either implied or express, to any third-party material is granted to you by this document or DASH-IF. DASH-IF makes no warranty whatsoever for such third-party material.

Note that technologies included in this document and for which no test and conformance material is provided, are only published as candidate technologies, and may be removed if no test material is provided before releasing a new version of this guidelines document. For the availability of test material, please check <https://www.dashif.org>.

---

## Foreword

This Technical Specification (TS) has been produced by the DASH-IF Technical Working Group.

---

## Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in deliverables except when used in direct citation.

---

## Executive Summary

The present document defines the CMAF Media Profiles and the DASH signalling for text tracks, including subtitles and captions as well as open captions and subtitles in video tracks. This work was derived from IOP v4.3 [i.3], but does not contain non-CMAF profiles.

---

## Introduction

The present document is Part 9 of a multipart set of documents, collectively called "IOP V5.0.0". All the parts are:

1. Overview, architectures, and interfaces
2. Core principles and CMAF mapping
3. On-demand services
4. Live and low-latency live services
5. Ad insertion
6. Content protection
7. Video
8. Audio
9. Text

10. Events
11. Additional functionalities
12. Conformance and reference tools

---

# 1 Scope

The present document defines the CMAF Media Profiles and the DASH signalling for text tracks, including subtitles and captions as well as open captions and subtitles in video tracks. This work was derived from IOP v4.3 [i.3], but does not contain non-CMAF profiles. The use of “sidecar” files is not covered in the present document.

---

## 2 References

### 2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, DASH-IF cannot guarantee their long-term validity.

The following referenced documents are necessary for the application of the present document.

- [1] DASH-IF IOP v5.0.0, Part 2, “Core principles and core mapping”
- [2] ISO/IEC 23009-1: “Information technology -- Dynamic adaptive streaming over HTTP (DASH) – Part 1: Media presentation description and segment formats”
- [3] ISO/IEC 14496-30 “Information technology -- Coding of audio-visual objects -- Part 30: Timed text and other visual overlays in ISO base media file format”
- [4] ISO/IEC 23000-19 “Information technology -- Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media”
- [5] SCTE 214-1, “MPEG DASH for IP-Based Cable Services Part 1: MPD Constraints and Extensions”

### 2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, DASH-IF cannot guarantee their long-term validity.

The following referenced documents are not necessary for the application of the present document, but they assist the user with regard to a particular subject area.

- [i.1] CTA 608-E, “Line 21 Data Services”
- [i.2] CTA 708-E, “Digital Television (DTV) Closed Captioning”
- [i.3] DASH-IF IOP v4.3, “Guidelines for Implementation: DASH-IF Interoperability Points”
- [i.4] DASH-IF IOP v5.0.0, Part 7, “Video”
- [i.5] SCTE 128-1:2020, “AVC Video Constraints for Cable Television Part 1- Coding”
- [i.6] W3C, “TTML Profiles for Internet Media Subtitles and Captions 1.0.1 (IMSC1)”
- [i.7] W3C, “WebVTT: The Web Video Text Tracks Format”

---

## 3 Definition of terms, symbols and abbreviations

### 3.1 Terms

For the purposes of the present document, the following terms apply:

**captions** - text that transcribes the audio dialogue, often includes non-verbal sounds, and primarily for accessibility

**closed captions** – captions encoded in a separate text track or encoded in the video track SEI

**closed subtitles** – subtitles encoded in a separate text track or encoded in the video track SEI

**open captions** - captions encoded in the video track pixels

**open subtitles** - subtitles encoded in the video track pixels

**subtitles** – text that transcribes the audio dialogue, often in a language different from the audio, and primarily for translation

## 3.2 Symbols

For the purposes of the present document, the following symbols apply:

## 3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

CMAF	Common Media Application Format
CTA	Consumer Technology Association
DASH	Dynamic Adaptive Streaming over HTTP
IEC	International Electrotechnical Commission
IMSC	Internet Media Subtitles and Captions
IOP	InterOperability Points
ISO	International Standards Organization
Kbps	Kilobits per second
MPD	Media Presentation Description
SEI	Supplemental Enhancement Information
VTT	Video Text Tracks
XML	eXtensible Markup Language

---

## 4 CMAF Media Profiles

Text CMAF Media Profiles shall be as defined in ISO/IEC 23000-19 [4] clauses 11, A.4 and A.5, which includes IMSC, WebVTT and CTA 608/708. @codecs for WebVTT and IMSC1 shall be as defined in ISO/IEC 14496-30.

CTA 608/708 may be carried in a CMAF Video Track in SEI messages and does not have a separate @mimeType or @codecs value. The file brand is used with the video file brand to signal their presence. Only CTA-608 is defined for use here, but it is carried in a CTA-708 DTVC wrapper, so is referred to “CTA-608/708”.

CTA 608/708 does not have a timing model but it is bound to video frames. To the extent the video is properly groomed for the Period, so will the CTA 608/708 data.

Open captions and open subtitles are addressed in DASH-IF IOP v5.0.0, Part 7, Video [i.4].

Table 1 is a comprehensive list of the text media profiles that support CMAF packaging along with some important parameters. These codecs, along with any additional codecs added since this publication can be found in the online DASH-IF specification repository found here: <https://dashif.org/codecs/text>.

**Table 1 CMAF Media Profile parameters**

Media Profile	File Brand	@mimeType	@codecs	Reference
IMSC1 Text	im1t	application/mp4	stpp.ttml.im1t	ISO/IEC 23000-19 [4]
IMSC1 Image	im1i	application/mp4	stpp.ttml.im1i	ISO/IEC 23000-19 [4]
WebVTT	cwvt	application/mp4	wvtt	ISO/IEC 23000-19 [4]
CTA 608/708	ccea	n/a	n/a	ISO/IEC 23000-19 [4]

Additional non-CMAF text (“subtitle”) profiles can be found in DASH-IF IOP v4.3 [i.3].

## 5 Adaptation Set requirements and recommendations

### 5.1 Content requirements

#### 5.1.1 Text tracks

In addition to the general provisions defined in IOP V5 part 2 [1], additionally Adaptation Sets shall comply with the provisions of Table 2.



Table 2 Text track Adaptation Set attributes and elements

DASH Attribute or Element	Use for media type	Detailed Usage in DASH-IF IOPs
@mimeType	M	See ISO/IEC 23009-1 [1], clause 5.3.7.2 Table 13. This shall be set to one of the @mimeType values defined in Table 1.
@codecs	M	See ISO/IEC 23009-1 [1], clause 5.3.7.2 Table 13. If @mimeType is set to "application/mp4", then this element shall be present and set to one of the @codecs values defined in Table 1. For CTA 608/708, this attribute is not applicable since it is signalled in a video Adaptation Set and an appropriate video value is used.
@lang	M	See ISO/IEC 23009-1 [1], clause 5.3.3.2 Table 5. The @lang attribute shall be present, and shall be set according to DASH , clause 8.12.4.3, AdaptationSet Constraints, using the same settings when @contentType is audio. Language is used as the primary selection based on a user preference setting. .
<b>Accessibility</b>	0 ... N	See ISO/IEC 23009-1 [1], clause 5.3.3.2 Table 5. If the text track is closed captions, then this element shall be present with the DASH role scheme as defined in ISO/IEC 23009-1 [1], 5.8.5.5, with @schemIdUri set to: "urn:mpeg:dash:role:2011"; and the @value shall be set to caption.
<b>Role</b>	0 ... N	See ISO/IEC 23009-1 [1], clause 5.3.3.2 Table 5. Role elements may be present depending on the nature of the text. In all cases Role shall be used as defined in ISO/IEC 23009-1, 5.8.5.5, with @schemIdUri set to: "urn:mpeg:dash:role:2011" . The cases are as follows:  <ol style="list-style-type: none"> <li>1. If the text track contains subtitles then an instance of this element should be present with @value set to "subtitle". Else if the text track contains closed captions, then an instance of Role shall be present with @value set to "caption" . If neither of these values is present, then then default is "subtitle" .</li> <li>2. If the text track contains text tailored to the needs of beginning readers, then an instance of Role shall be present with @value set to "easyreader" .</li> </ol>

Text Adaptation Sets containing alternative content shall differ by at least by one of the following annotation labels:

- @codecs: specifies the codecs present within the Representation.
- @lang: specifies the language of the subtitle with a non-null language code
- **Accessibility** descriptor with DASH role scheme as defined in ISO/IEC 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011, with value set to "caption".
- One or more **Role** descriptors with DASH role scheme as defined in ISO/IEC 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011 .

## 5.1.2 Video tracks

### 5.1.2.1 General

In addition to text tracks, CTA 608/708 captions can be carried in video tracks. This can be done either by open subtitles or open captions where the text is coded into the video pixels, or by carrying them in the video SEI according to SCTE 128-1 [i.5]. This section addresses closed captions in the video SEI.

**Table 3 Video track Adaptation Set attributes and elements**

DASH Attribute or Element	Use for media type	Detailed Usage in DASH-IF IOPs
<b>Accessibility</b>	0 ... N	<p>See ISO/IEC 23009-1 [1], clause 5.3.7.2 Table 13.</p> <p>If the video Adaptation Set contains CTA 608/708 closed captions, then this element shall be present and used as defined in ISO/IEC 23009-1 [2] 5.8.5.5 and see section 5.1.2.2.</p> <p>For other uses of <b>Accessibility</b> with a video track, including open captions and open subtitles, see IOP v5.0.0, Part 7, Video [i.4].</p>

Since this is a video Adaptation Set, the @mimeType and @codecs are set per the video codec.

### 5.1.2.2 CTA 608/708

When the video track contains CTA 608/708 then the video Adaptation Set **Accessibility**@value shall be present and shall be as described in SCTE 214-1 [5], section 7.2.3, “Signalling CEA-608 caption service metadata”. The forms where @value is present and contains both the caption channel and language should be used, e.g.:

```
<Accessibility schemeIdUri="urn:scte:dash:cc:cea-608:2015" value="CC1=eng;CC3=spa"/>
```

When there is only a single caption channel present, then just the language form may be used, e.g.:

```
<Accessibility schemeIdUri="urn:scte:dash:cc:cea-608:2015" value="eng"/>
```

When more than one caption channel is present, the language-only form should not be used.

## 5.1.3 Chunks and gaps

Text is low bitrate, but has different characteristics compared to audio. In particular, for IMSC1 there is typically only one sample (document) in each segment. Text should not be chunked but delivered as separate segments (documents). Even a short segment duration, encoders still shall conform to the IMSC1 Hypothetical Render Model (HRM) in IMSC1[i.6], and applied inter-segment.

When a text track contains a period of no content, continuous segments shall still be present and shall contain the “empty” document as defined in 14496-30 [3].

## 5.2 Client recommendations

It is expected that DASH clients conforming to this IOP recognize the descriptors, elements, and attributes and their values as documented in Table 2 and Table 3.

If caption or subtitle text rendering is enabled, based on the text Adaptation Sets the client selects one from the signalling in clause 5.1 as follows:

1. Any text Adaptation Set for which an Essential Descriptor is present for which the scheme or value is not understood by the DASH client, is excluded from the selection.
2. Any text Adaptation Set for which the client does not have a decoder (e.g. CTA 608/708, IMSC1, WebVTT), is excluded from the selection.
3. If text language preference settings are provided to the client by the system:

- a. any Adaptation Set from the selection where @lang is absent, null or set to “und” is excluded from the selection; and
  - b. any Adaptation Set from the selection that is not in the preferred text language, is excluded from the selection.
4. Any text Adaptation Set with one or more **Role**@schemeIdUri="urn:mpeg:dash:role:2011" where none of the @value are described in the present document, is excluded from the selection.
  5. If multiple text Adaptation Sets remain, then the ones with the highest value of @selectionPriority is chosen.
  6. If multiple text Adaptation Sets remain, then the DASH client makes a choice for itself, possibly on a random basis.

NOTE: Text in video tracks does not affect video track selection.

---

## Annex A (informative): Change History

Date	Version	Information about changes
2022-01-04	5.0.0	Version published as part 9 v5.0.0.