TECH 3381

CARRIAGE OF EBU-TT-D IN ISOBMFF

VERSION: 1.0

SOURCE: SP/MIM – XML SUBTITLES

Geneva
October 2014

Conformance Notation

This document contains both normative text and informative text.

All text is normative except for that in the Introduction, any section explicitly labelled as 'Informative' or individual paragraphs which start with 'Note:'.

Normative text describes indispensable or mandatory elements. It contains the conformance keywords 'shall', 'should' or 'may', defined as follows:

'Shall' and 'shall not':	Indicate requirements to be followed strictly and from which no deviation is permitted in order to conform to the document.
'Should' and 'should not':	Indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others.
	OR indicate that a certain course of action is preferred but not necessarily required.
	OR indicate that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.
'May' and 'need not':	Indicate a course of action permissible within the limits of the document.

Default identifies mandatory (in phrases containing "shall") or recommended (in phrases containing "should") presets that can, optionally, be overwritten by user action or supplemented with other options in advanced applications. Mandatory defaults must be supported. The support of recommended defaults is preferred, but not necessarily required.

Informative text is potentially helpful to the user, but it is not indispensable and it does not affect the normative text. Informative text does not contain any conformance keywords.

A conformant implementation is one which includes all mandatory provisions ('shall') and, if implemented, all recommended provisions ('should') as described. A conformant implementation need not implement optional provisions ('may') and need not implement them as described.

* Page intentionally left blank. This document is paginated for two sided printing

Contents

Status of this document (Informative)

This document is a stable document and may be used as reference material or cited from another document.

This document is part of a series of EBU-TT (EBU Timed Text) documents. The full list of published and planned EBU-TT documents is given below.

Part 1: EBU-TT Subtitling format definition (EBU Tech 3350)

Introduction to EBU-TT and definition of the XML based format.

Part 2: STL (Tech 3264) Mapping to EBU-TT (EBU Tech 3360)

How EBU-TT provides backwards compatibility with EBU STL.

EBU-TT Live - Authoring and Contribution

How to use EBU-TT for the production and contribution of live subtitles.

EBU-TT Annotation

How EBU-TT can be used in future scenarios for 'authoring of intent'.

EBU-TT User Guide

General guide ('How to use EBU-TT').

EBU-TT-D (EBU Tech 3380)

EBU-TT content profile for TTML that can be used for the distribution of subtitles over IP based networks.

Carriage of EBU-TT-D in ISOBMFF (EBU Tech 3381)

How EBU-TT-D can be stored using the storage format of the ISO Base Media File Format (ISO/IEC 14496-12).

* Page intentionally left blank. This document is paginated for two sided printing

8

Carriage of EBU-TT-D in ISOBMFF

EBU Committee	First Issued	Revised	Re-issued
TC	2014		

Keywords: subtitling, STL, XML, W3C, TTML, DFXP, captions, EBU Timed Text, ISOBMFF, DASH.

Scope (Informative)

The aim of this publication is to specify how EBU-TT-D as defined in EBU Tech 3380 *EBU TT-D Subtitling distribution format* [1], hereafter referred to as EBU-TT-D, can be stored using the storage format of the ISO base media file format defined in ISO/IEC 14496-12 *Information technology - Coding of audio-visual objects - Part 12: ISO base media file format*, hereafter referred to as ISO/IEC 14496-12 [2].[1]

This publication builds upon ISO/IEC 14496-12:2012/Amd 2:2014 *Information technology - Coding of audio-visual objects - Part 12: ISO base media file format, AMENDMENT 2: Carriage of timed text and other visual overlays* [3], hereafter referred to as ISO/IEC 14496-12:2012/Amd 2:2014, and ISO/IEC 14496-30:2014 *Information technology - Coding of audio-visual objects - Part 30: Timed text and other visual overlays in ISO base media file format* [4], hereafter referred to as ISO/IEC 14496:30:2014.

1.   Generic constraints

If not noted otherwise the constraints of ISO/IEC 14496-12 [2], ISO/IEC 14496-12:2012/Amd 2:2014 [3] and ISO/IEC 14496-30:2014 [4] shall apply.

2.   Track format

As defined in ISO/IEC 14496-30:2014 section 6.4 TTML streams shall be carried in subtitle tracks. Subtitle tracks are defined in ISO/IEC 14496-12:2012/Amd 2.

3.   Sample Entry format

EBU-TT-D subtitle streams shall use the XMLSubtitleSampleEntry format as defined by ISO/IEC 14496-30:2014.

---

[1] This storage format is implemented by file formats such as MPEG-4 (ISO/IEC 14496-14) and JPEG 2000 (ISO/IEC 15444-3).

4.   Sample format

An EBU-TT-D subtitle sample consists of one XML document that conforms to EBU-TT-D.

An EBU-TT-D document in a sample is stateless and contains all information needed by a presentation processor to render it (e.g. region and style definitions).

Note: An ISOBMFF fragment conforming to this specification satisfies the requirements for a Type 1 Stream Access Point in MPEG DASH [5].

5.   Layout and Styling

The layout behaviour shall be as defined in ISO/IEC 14496-30:2014 with the following constraints:

When the subtitle track is associated with a video object the subtitles shall be rendered onto a plane coincident with the plane on which any associated video is applied. The width and height of the subtitle track header should therefore not be set, and shall not be used for processing or displaying the subtitles.

If no video track is associated with the subtitle track the presentation system shall allocate a suitable rendering plane on which to present the subtitles. In this case the width and height of the subtitle track header may be set and used for processing or displaying the subtitles.

6.   Timing

The timing behaviour for processing EBU-TT-D subtitle samples shall be the same as defined in ISO/IEC 14496-30:2014 Section 5.2 and 6.3. In accordance with this definition all time expressions in the EBU-TT-D document shall be relative to the start of the subtitle track and not to the start of the sample.

The following additional constraints are made:

The begin and end of an EBU-TT-D document may not be coincident with the begin and end of the sample. All subtitle content that falls partially or wholly within the duration of the sample shall be present in the EBU-TT-D document and shall be displayed. Subtitle content that falls entirely outside the duration of the sample shall be omitted from the EBU-TT-D document that is contained within that sample.

Any timed content that falls partially outside of the sample boundaries may be used for error recovery (e.g. if a subsequent sample has not arrived in time for processing), however if a subtitle sample corresponding to a particular video frame visibility period has been received then that subtitle sample shall the only one from which subtitles shall be displayed.

Note: Care must be taken if edit lists are applied to the related video track to ensure that the media times within the subtitle track match. For example if an edit list offsets the start of video presentation by 4 frames (so frame 'zero' is displayed 4 frames into the playback timeline) then measures need to be taken to ensure that the offset of subtitle timing is adjusted to match. This can be achieved either by retiming the subtitles within the EBU-TT-D document to add the required offsets, or more straightforwardly by applying a similar edit list to the subtitle track to achieve the offset functionality within the ISOBMFF track layout. If re-use of the EBU-TT-D sample is a requirement, the latter option is recommended.

Presentation systems must honour the track timings and edit lists to present correctly timed subtitles.

In ISO/IEC 14496-30:2014 section 6.3 an 'empty' sample shall be a sample containing a TTML document that has no content. A processor SHALL accept either of the following two documents as TTML documents that have no content:

- an EBU-TT-D document with no body element
- an TTML document with the XML content:
  ```
  <tt xml:lang="" xmlns="http://www.w3.org/ns/ttml" />
  ```

NOTE: This definition is not intended to override the ISOBMFF definition of an empty sample, merely to indicate two concrete examples of documents with no content that must be accepted as empty samples.

7.   Bibliography

1	EBU-TT-D	EBU Tech 3380, EBU TT-D Subtitling distribution format
2	ISO/IEC 14496-12	Information technology - Coding of audio-visual objects - Part 12: ISO base media file format
3	ISO/IEC 14496-12:2012/Amd 2:2014	Information technology - Coding of audio-visual objects - Part 12: ISO base media file format, AMENDMENT 2: Carriage of timed text and other visual overlays
4	ISO/IEC 14496-30:2014	Information technology - Coding of audio-visual objects - Part 30: Timed text and other visual overlays in ISO base media file format
5	ISO/IEC 23009-1 2012-04-01	Information technology – Dynamic adaptive streaming over HTTP (DASH) – Part 1: Media presentation description and segment formats