

**AIRLINE PASSENGER EXPERIENCE  
ASSOCIATION**

**APEX SPECIFICATION 0814  
Captions and Subtitles for  
Inflight Entertainment Systems**

**Version 1**  
**2016-05-24**

**© 2016 Airline Passenger Experience Association. All Rights Reserved.**

The Airline Passenger Experience Association (APEX) is the author and creator of this specification for the purpose of copyright and other laws in all countries throughout the world. The APEX copyright notice must be included in all reproductions, whether in whole or in part, and may not be deleted or attributed to others. APEX hereby grants to its members and their suppliers a limited license to reproduce this specification for their own use, provided it is not sold. Others should obtain permission to reproduce this specification from the Airline Passenger Experience Association, Attn: Executive Director, 355 Lexington Avenue, 15th Floor, New York, NY 10017 USA; Phone: (212) 297-2177; Fax: (212) 370-9047; email: info@apex.aero.

#### IMPORTANT NOTICES

This document is a specification adopted by the Airline Passenger Experience Association (APEX). This document may be revised by APEX. It is intended solely as a guide for companies interested in developing products that can be compatible with other products developed using this document. APEX makes no representation or warranty regarding this document, and any company using this document shall do so at its sole risk, including specifically the risks that a product developed will not be compatible with any other product or that any particular performance will not be achieved. APEX shall not be liable for any exemplary, incidental, proximate or consequential damages or expenses arising from the use of this document. This document defines only one approach to compatibility, and other approaches may be available to the in-flight industry.

This document is an authorized and approved publication of APEX. Only APEX has the right and authority to revise or change the material contained in this document, and any revisions by any party other than APEX are unauthorized and prohibited.

Compliance with this document may require use of one or more features covered by proprietary rights (such as features which are the subject of a patent, patent application, copyright, mask work right or trade secret right). By publication of this document, no position is taken by APEX with respect to the validity or infringement of any patent or other proprietary right. APEX hereby expressly disclaims any liability for infringement of intellectual property rights of others by virtue of the use of this document. APEX has not and does not investigate any notices or allegations of infringement prompted by publication of any APEX document, nor does APEX undertake a duty to advise users or potential users of APEX documents of such notices or allegations. APEX hereby expressly advises all users or potential users of this document to investigate and analyze any potential infringement situation, seek the advice of intellectual property counsel, and, if indicated, obtain a license under any applicable intellectual property right or take the necessary steps to avoid infringement of any intellectual property right. APEX expressly disclaims any intent to promote infringement of any intellectual property right by virtue of the evolution, adoption, or publication of this document.

## Table of Contents

Foreword.....	4
1. Introduction.....	4
2. Scope.....	4
3. System Reference Model.....	5
3.1 Introduction.....	5
3.2 Workflow.....	5
3.3 Formats.....	5
3.4 Conversions.....	6
3.5 Editing.....	6
3.6 QC.....	6
3.7 Delivery.....	6
3.8 Timeline.....	6
4. Normative References.....	67
5. Informative References.....	7
6. Terms, Definitions and acronyms.....	7
7. Normative Clauses.....	119
7.1 Caption Format.....	119
7.2 CEA (608) Conversions.....	119
8. Informative Annexes.....	119
8.1 Acquisition workflows.....	119
8.2 Caption Quality.....	119
8.3 IMF OPL for IFE workflows.....	119
8.4 Output conversion to Legacy Format of APEX 0403.....	119
8.5 Output conversion to Alternate Formats (WebVTT).....	1210
8.5.1 Introduction and background on WebVTT format:.....	1210
8.5.2 Overview of data type and genesis:.....	1210
8.5.3 Use case, why WebVTT is required by some systems:.....	1210
8.5.4 Constraints required of IMSC source, if any.....	1210
8.5.5 Feature constraints specified for compliance: WebVTT Device Limitations.....	1211
8.5.6 Player or render feature capability requirement or expectation.....	1311
8.5.7 W3C conversion recommendations, with any application notes.....	1412
8.6 List of Participants.....	1514

## Foreword

As inflight entertainment (IFE) has evolved, so have the media requirements. The earliest digital systems specified unique media delivery formats, however systems have been evolving toward adopting commonly available formats. This evolution has generally helped to streamline and simplify the supply chain. The purpose of this standard is to specify a broadly available format for captions and subtitles, and indicate how to use that format for IFE. This specification is a voluntary specification; however achieving compliance will bring economies of effort and timesaving's to IFE workflows.

## 1. Introduction

Captions and subtitles are included extensively in inflight entertainment, both for accessibility and to increase the languages offered. This standard will set out the format of caption and subtitle data, and packaging with program media. This document will also include typical workflows for acquisition and conversions. Information will be provided describing conversions and interface with legacy standards, such as APEX 0403. The standard is based on W3C and SMPTE standards. It provides a path to discontinue open captions and promote user selectable captions.

The authors recognize that this standard notwithstanding, the contribution format provided may be as mutually agreed by content provider and media service provider.

## 2. Scope

The specification defines a caption and subtitle data format for IFE use. This format may be provided early in the content delivery process, such as when an IMF file source includes captions in this format as part of that package. Conversions into this format can be performed at subsequent stages from other available formats as part of the preparation process. When a compatible media format is utilized, this format is suitable for inclusion in that package for delivery to a passenger screen. The scope includes deliveries where playback originates on the aircraft, or for distribution paths intended solely for IFE viewers.

This format shall be the basis for workflows and data interchange between entities in the delivery supply chain. Compliant systems will utilize this format to the passenger screen.

Systems designed following the adoption of this standard are within its scope. IFE systems typically have a long life cycle, as long as 12 years from design through deployment and retirement. There is no requirement to update existing systems from legacy formats that are in use currently.

Mindful of potential accessibility requirements by the US Department of Transportation, meeting the overall goal of providing captions and subtitles is our main objective.

### 3. System Reference Model

#### 3.1 Introduction

The process of acquiring captions and subtitles in IFE deliveries comprise legacy workflows and integrated workflows. Legacy workflows typically have multiple variations, and formats. An integrated workflow uses the data formats native to this specification.

#### 3.2 Workflow

A highly integrated workflow would originate with an IMF package. Both caption and subtitle data may be included in that package, using the IMSC format specified in this document. That data would accompany the media through any editing and formatting processes, and be included in an IFE output profile product, as anticipated in APEX 0415.

Current IFE deliveries typically have three defined stages in the process. Content owner to media services (lab), media services to media integrator (CSP or HW manufacturer), and delivery to the onboard system and ultimately the passenger. While our intent is to recommend that this format ~~will~~ be utilized in the delivery from content owner to media services provider, the available format will be most efficient for this purpose. The remaining stages of the process are required to utilize this format for compliance.

#### 3.3 Formats

Several caption and subtitle legacy formats exist. The formats differ for the two types. A lowest common denominator would be a text file containing words and timing information. For text information the character coding can vary. Unicode, as UTF-8 or UTF-16 is a worldwide standard. Language specific coding's were used before Unicode and may be part of legacy formats. Current caption formats include .scc and .cap files. Subtitle formats include .pac files, and SRT files.

Source formats would include TTML (IMSC), SCC files (TV), new captions, Digital Cinema captions, IMF containing IMSC.

### 3.4 Conversions

Conversions between text-based formats require reformatting the data and possibly converting the timing information. A conversion from text based to image based requires rendering the characters into a graphics format, usually TIFF or BMP. Some common tools are, GIC DDP, Easytitle, EZ-Sync, Scenarist ST2, BluRay image and timing formats.

### 3.5 Editing

Captions may need to be edited to conform to a specific version that has been created for IFE. This editing, or reformat of CC and SUB files, is usually done after the media editing. An editing workflow that simultaneously conforms the files is desirable. It has not until recently been possible as an automated process, and is becoming available in new tools.

### 3.6 QC

Quality control processes should insure completeness, timing accuracy, and placement on screen relative to the media content. These checks may be manual observation by a human QC operator, or performed by an automated program.

### 3.7 Delivery

Packaging caption and subtitle data is delivery format specific. For the legacy APEX0403 format, the image files are multiplexed into an MPEG Transport Stream. The APEX 0415 format will directly utilize the IMSC format of this specification. Tools include Manzanita MP2-TSME.

### 3.8 Timeline

It should be noted that some content owners advise that the captioning process may add 5-7 days to delivery timeline. This might be due to processing time or approval requirements.

## 4. Normative References

TTML Text and Image Profiles for Internet Media Subtitles and Captions 1.0 (IMSC-1) W3C Recommendation 21 April 2016  
SMPTE RP2052-10:2012 Conversion from CEA-608 Data to SMPTE TT

## 5. Informative References

W3C - TTML1 Timed Text Markup Language 1 (TTML1) (Second Edition). 24 September 2013  
SMPTE - ST 2052-1, Timed Text Format (SMPTE-TT)  
FCC 14-12 Report and Order, Declaratory Ruling, and Further Notice of Proposed Rulemaking  
European Broadcasting Union (EBU) - Tech 3380, EBU-TT-D Subtitling Distribution Format Version 1.0 (EBU-TT-D)  
EMA - Best Practices For Closed Captioning Of Internet Protocol- Delivered Video Programming

## 6. Terms, Definitions and acronyms

**Analog caption decoder** *A device mandated to be included in US television sets that can convert the data sent on line 21 of the analog video signal into an on-screen caption overlay visible to the viewer*

**Bit map** *An image represented by a generally rectangular grid of pixels. A rendered picture of the information such as a caption in raster format*

**Burned in subtitle/captions** *A process where caption or subtitle information has been rendered into a bit map picture and made a part of the video replacing the underlying image*

**Captions** *Text version of spoken words in a program, along with audio cues intended for hearing impaired viewers*

**CTA** *Consumer Technology Association, formerly CEA*

**CEA-608-E R-2014** *Technical standard for line 21 data services*

**Closed Captions** *Enabled by viewer selection. They are conveyed along with the visual data, but separate from it. When a viewer chooses to enable captions, they are combined with the picture at the display*

**Content Owner**

Formatted: Normal

Formatted: Font: Bold

Formatted Table

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

<u>CSP</u>	<u>Content Service Provider</u>	Formatted: Font: 12 pt
<u>CVAA</u>	<u>21st Century Communications and Video Accessibility Act , a 2010 US Law.</u>	Formatted: Font: Bold
<u>Dynamic subtitles</u>	<u>User selectable, as opposed to burnt-in</u>	Formatted: Font: Bold
<u>EBU</u>	<u>European Broadcast Union</u>	Formatted: Font: Bold
<u>Embedded IFE</u>	<u>An IFE system attached to an aircraft, usually seen by a passenger as a seatback video display unit</u>	Formatted: Font: Bold
<u>IFE</u>	<u>Inflight Entertainment</u>	Formatted: Font: Bold
<u>IFE HW manufacturer</u>		Formatted: Font: Bold
<u>IMF</u>	<u>Interoperable Master format</u>	Formatted: Font: 12 pt
<u>IMF OPL</u>	<u>IMF Output Profile List</u>	Formatted: Font: Bold
<u>IMSC</u>	<u>Internet Media for Subtitles and Captions</u>	Formatted: Font: Bold
<u>Open Captions</u>	<u>Are inherently visible and cannot be turned off. They are burned in at the source as the specific media format for each system is created</u>	Formatted: Font: Bold
<u>Render</u>	<u>The process of converting information from text or data into a raster image or picture</u>	Formatted: Font: Bold
<u>Scenarist ST-@</u>	<u>Subtitle format used by Scenarist DVD authoring</u>	Formatted: Font: Bold
<u>SMPTE</u>	<u>Society of Motion Picture and Television Engineers</u>	Formatted: Font: Bold
<u>SMPTE-TT</u>	<u>SMTE Timed Text</u>	Formatted: Font: Bold
<u>Subtitle</u>	<u>Render the spoken dialog and narrative into a different language by translation. Subtitles are intended for hearing passengers who have a preference for a different language than the original</u>	Formatted: Font: Bold
<u>Timed Text</u>	<u>Text that is displayed in sync with a media program, such as captions. The standard that is used to represent timed text</u>	Formatted: Font: Bold
<u>TTML</u>	<u>Timed Text Markup Language</u>	Formatted: Font: Bold

**WebVTT** Web Video Text Tracks. A standard for adding captions or subtitles to the HTML5 video standard as supported by Internet browsers

Formatted: Font: Bold

**W3c** The World Wide Web Consortium (W3C) is an international community that develops open standards to ensure the long-term growth of the Web.

Formatted: Font: Bold

Formatted: Normal

## 7. Normative Clauses

**Analog caption decoder**—A device mandated to be included in US television sets that can convert the data sent on line 21 of the analog video signal into an on-screen caption overlay visible to the viewer

**Bit map**—An image represented by a generally rectangular grid of pixels. A rendered picture of the information such as a caption in raster format

**Burned in subtitle/captions**—A process where caption or subtitle information has been rendered into a bit map picture and made a part of the video replacing the underlying image

**Captions**—text version of spoken words in a program, along with audio cues intended for hearing impaired viewers

**CTA**—Consumer Technology Association, formerly CEA

**CEA-608-E-R-2014**—Technical standard for line 21 data services

**Closed Captions**—Enabled by viewer selection. They are conveyed along with the visual data, but separate from it. When a viewer chooses to enable captions, they are combined with the picture at the display

**Content Owner**—

**CSP**—Content Service Provider

**CVAA**—21st Century Communications and Video Accessibility Act, a 2010 US Law.

**Dynamic subtitles**—User selectable, as opposed to burnt in

**EBU**—European Broadcast Union

**Embedded IFE**—An IFE system attached to an aircraft, usually seen by a passenger as a seatback video display unit

**IFE**—Inflight Entertainment

**IFE HW manufacturer**—

**IMF**—Interoperable Master format

**IMF OPL**—IMF Output Profile List

**IMSC**—Internet Media for Subtitles and Captions

**Open Captions**—are inherently visible and cannot be turned off. They are burned in at the source as the specific media format for each system is created

Formatted: Font: Italic

~~Render~~ ———— ~~The process of converting information from text or data into a raster image or picture~~

### 7.1 Caption Format

Caption and subtitle data shall be formatted per the latest version of W3C Recommendation IMSC, TTML Text and Image Profiles for Internet Media Subtitles and Captions.

### 7.2 CEA (608) Conversions

Conversions from CEA formats should be per SMPTE RP2052-10:2012 and SMPTE RP2052-11:2013

## 8. Informative Annexes

### 8.1 Acquisition workflows

IMSC-1 is a super set of SMPTE RP2052, TTML. RP2052 is designated as the safe harbor data format by the FCC in the CVAA. This places IMSC as the central common format in the current world of captions and subtitles.

### 8.2 Caption Quality

The US Federal Communications Commission has published a guideline for Caption Quality that is highly informative to IFE needs. Document FCC-14-12A1 contains a set of Minimum Captioning Quality Standards for Accuracy, Synchronicity, Completeness, and Placement. ~~APEX~~ will address these requirements with respect to the limitations for IFE systems.

### 8.3 IMF OPL for IFE workflows

The caption and subtitle data type required by this specification is also incorporated in SMPTE ST 2067, the Interoperable Master Format. As such, this is a native source for captions and subtitles. The IMF workflow is designed to support distribution file creation by an Output Profile List (OPL) that may be implemented by multiple vendors. Creating an IFE OPL will be developed as a part of Apex Specification 0415.

### 8.4 Output conversion to Legacy Format of APEX 0403

The currently available tools and process requires that captions or subtitles first be converted to DVB standard image format. While this is a known process, a desirable direct conversion will require new tools that use IMSC as an input format and convert to DVB data in Transport Stream delivery.

## 8.5 Output conversion to Alternate Formats (WebVTT)

The intent of the annex is to provide information and guidance on setting up a conversion workflow from the APEX format, IMSC, to WebVTT. Source constraints for best mapping are detailed, as this is not an exact conversion due to lack of feature parity. Please refer to W3C recommendations in all cases.

### 8.5.1 Introduction and background on WebVTT format:

[From Wikipedia]

WebVTT (Web Video Text Tracks) is a [W3C](#) standard for displaying [timed text](#). WebVTT files provide captions or subtitles for video content, text video descriptions, and more generally any form of metadata that is time-aligned with audio or video content. WebVTT files are delivered as timed-text files separate from the video and associated video-audio and are to be displayed, or not displayed, based on passenger input. The W3C WebVTT specification is still in draft stage but the basic features are already supported by all major browsers.

### 8.5.2 Overview of data type and genesis:

The .srt format is a precursor to WebVTT. For over ten years, plug in filters have been available that allow Windows Media Player to show the contents of .SRT files.

### 8.5.3 Use case, why WebVTT is required by some systems:

The most important use case for WebVTT is for “BYOD” IFEC; IFEC systems where passengers are allowed to Bring Your Own Device. For systems where the captioning and subtitles are delivered to COTS devices, like in a WiFi IFE System that streams to the passenger devices, there is a need to stream in the formats that are supported by these devices. At the time of writing, WebVTT has the broadest support among web browsers and media players on PCs and mobile devices.

### 8.5.4 Constraints required of IMSC source, if any.

- Pop-on, pop-off text only
- No scrolling text
- No flashing text
- UTF-8 Unicode scripts/text only

### 8.5.5 Feature constraints specified for compliance: WebVTT Device Limitations

When serving today’s WebVTT clients, it’s acceptable to employ multiple regions as long as a single event timeline is maintained. This graphic design constraint is not guaranteed, particularly when separately authored forced subtitle and caption tracks are combined in a downstream process.

As an example, at the beginning of a scene, imagine the top of the frame includes a warning sign that is located above a door.

In this example, that warning sign is in a language unfamiliar to the characters in the story. It is important that the audience understand what the sign says even though the characters in the scene do not. It is common for a forced subtitle track to translate the sign into the preferred language of the viewer. When our characters say, “I can’t read the sign above the door; let’s go inside and see what’s there.” the closed caption track, which might be enabled by the audience, includes an event that straddles the timeline of the forced subtitle track.

In this scenario, to avoid the creation of multiple timelines, the authorizing tool would have to clear the screen, pop-on the forced subtitle near the top of the screen and leave it up. Next, when our characters begin to speak, that event must end and a new event including a region for the forced subtitle at the top and also a region for the caption at the bottom of the screen must begin.

As of this writing, popular captioning tools do not issue a warning when tracks with overlapping events are combined in this manner. However the file will fail upon export to a constrained WebVTT format.

It is strongly recommended that workflows which combine tracks in this manner employ the latest version of all tools. Service providers and laboratories are encouraged to read ANNEX ??? [Telestream’s technical bulletin] on this subject and observe the timeline segmentation guidance. Setting aside the implementation specifics, timeline segmentation generally needs to be turned on (or practiced manually). This is not merely a tool issue, it’s an ecosystem consideration that will endure as long as the current crop of “BYOD” and COTS devices remain in circulation.

#### 8.5.6 Player or render feature capability requirement or expectation

##### Minimum WebVTT Cue support

Cue	Support*	Description
line	Yes	Specifies where text appears vertically. If vertical is set, line specifies where text appears horizontally.
position	Yes	Specifies where the text will appear horizontally. If vertical is set, position specifies where the text will appear vertically.
size	Yes	Specifies the width of the text area. If vertical is set, size specifies the height of the text area.
align	Yes	Specifies the alignment of the text. Text is aligned within the space given by the size cue setting if it is set.
vertical	No	Indicates that the text will be displayed vertically rather than horizontally, such as in some Asian languages.

region	No	Provides configuration options regarding the dimensions, positioning and anchoring of the display region.
--------	----	---

#### Minimum WebVTT Text Tag support

Text Tag	Support	Description
(Timestamp)	No	Allows words or phrases to be added to the display, as is used in Karaoke
c	No	Class Object. Style the contained text using a CSS defined class.
i	Yes	Italics. Italicize the contained text.
b	Yes	Bold. Bold the contained text
u	Yes	Underline. Underline the contained text
ruby	No	Ruby. Used with ruby text tags to display <a href="#">ruby characters</a> (i.e. small annotative characters above other characters).
rt	No	Ruby Text. Used with ruby tags to display <a href="#">ruby characters</a> (i.e. small annotative characters above other characters).
v	No	Voice. Similar to class tag, also used to style the contained text using CSS.
lang	No	Language. Used to annotate parts of the cue where the <a href="#">applicable language</a> might be different than the surrounding text's language.

#### 8.5.7 W3C conversion recommendations, with any application notes

W3C mapping document (draft)

<http://w3c.github.io/ttml-webvtt-mapping/>

There is a good summary of the history of the format at:

<http://www.balisage.net/Proceedings/vol10/html/Tai01/BalisageVol10-Tai01.html>

Conversions

<http://docs.brightcove.com/en/perform/brightcove-player/guides/webvtt-converter.html>

## 8.6 List of Participants

This specification could not have been produced without the dedicated involvement of many individuals and companies. The following persons participated in the creation of this document by attendance at one or more meetings of the APEX CCWG in this work. Their company affiliation at the time of their participation is also given.

Kfir Adam	LinguisTech
Richard Ashton	Sony Pictures Releasing
Andy Beer	Spafax
Kevin Birchmore	DMD
Scott Braun	United Airlines
Deborah Chariton	Sony Pictures Releasing
Michael Childers	Lufthansa Systems (Co-Chair, CCWG)
Caroline Foley_Comer	LinguisTech
Geoff Freed	NCAM
Giovanni Galvez	TRelestream
Larry Goodman	Telestream
Mike Gromek	WestJet
Henry Gu	GIC
Victor Hernandez	Ackila
Diane Johnson	Descriptive Works
Greg Kadonada	Global Eagle Entertainment
Robert Kisor	Entertainment Technology Consultant
Sam Larkin	Bitlogic
Pierre-Anthony Lemieux	MovieLabs
Mike Metz	Gogo
Orit Michiel	MPAA
Kevin Minshull	WestJet
Mike Moffatt	Gogo
Jon Norris	Lumexis Corporation
John O'Connor	Cine Magnetics
Alex Olugbile	NBCUniversal
Dirk Ottens	Lufthansa Systems
Chana Ozawa	LinguisTech
Toshiro Ozawa	Thales Avionics
Crystal Pham	Sony Pictures Releasing
Christine Ringger	Swiss International Air Lines
Mary Rogozinski	GoGo
Andy Rosen	Bitlogic
Bryan Rusenko	APEX Technical Director (Co-Chair, CCWG)
Pierre Schuberth	Thales Avionics
David Schultens	Open Video Systems
Randy Schwarz	Panasonic Avionics Corporation
Carsten Schwarzbach	Lufthansa German Airlines
Dave Sherwin	WestJet
Justin Smith	Lumexis Corporation
Albert Stopniewicz	Panasonic Avionics Corporation
Jeff Tsaur	Thales Avionics
Phil Watson	Panasonic Avionics Corporation