SMPTE ST 2052-1:2013
Revision of
SMPTE ST 2052-1:2010

SMPTE STANDARD

Timed Text Format (SMPTE-TT)

Table of Contents

Approved
June 5, 2013

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices, and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in Part XIII of its Administrative Practices.

SMPTE ST 2052-1 was prepared by Technology Committee 24TB.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Engineering Document. However, attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. SMPTE shall not be held responsible for identifying any or all such patent rights.

Introduction

This section is entirely informative and does not form an integral part of this Engineering Document.

Timed text is an important essence element of information and entertainment programming. Various regulatory bodies and administrations around the world have, in both technical specification and quantitative requirement, made various forms of timed text (such as captions or subtitles) a necessary component of content.

This Standard provides a framework for timed text to be supported for content delivered via broadband means, taking into consideration the economic and practical necessity of the ability to use source data delivered in pre-existing and regionally-specific formats (such as CEA-708, CEA-608, DVB Subtitles, and WST (World System Teletext)) as well as text that may be authored or provided in other formats.

It is a goal of this Standard to make timed text essence more useful for broadband content, with equal status to the associated video and audio essence. Considering however that contemporary broadband content on many platforms (e.g., media players, media extenders, home servers, home networks, etc.) may be intended for display in a consumer electronic environment, and that there will be a transition period where such equipment can only utilize pre-existing formats, this Standard also defines functionality for the preservation of legacy formats. It is the intent, however, that new display devices will utilize the more capable TTML timed text structures rather than legacy data formats.

1 Scope

This Standard defines the SMPTE profile of W3C Timed Text Markup Language (TTML), designated SMPTE-TT, which may represent Captions or Subtitles. It identifies the features from TTML required for interoperability between display systems for the format. In order to preserve certain semantic features of the input formats, SMPTE-TT also defines some standard metadata terms to be used, and some extension features not found in TTML.

2 Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "Note:"

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword “reserved” indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be defined in the future.

A conformant implementation according to this document is one that 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.

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows: Normative prose shall be the authoritative definition; Tables shall be next; followed by formal languages; then figures; and then any other language forms.

3 Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

IETF RFC 2396 (August 1998), Uniform Resource Identifiers (URI): Generic Syntax

IETF RFC 4648 (October 2006), The Base16, Base32, and Base64 Data

ISO/IEC 15948:2003. Information Technology — Computer Graphics and Image Processing — Portable Network Graphics (PNG): Functional Specification

W3C Timed Text Markup Language (TTML) 1.0, W3C Recommendation 18 November 2010
http://www.w3.org/TR/ttaf1-dfxp/

W3C Extensible Markup Language (XML) 1.0 (Fifth Edition), W3C Recommendation 26 November 2008

W3C Extensible Stylesheet Language (XSL) Version 1.1, W3C Recommendation 05 December 2006

4 Definitions and Acronyms

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

4.1 Definitions

Caption: textual representation of the audio track, usually including all sounds, and usually in the same language as the audio track dialog, intended for hearing impaired audiences.

Chunk: A segment of a media timeline that forms part of a larger contiguous sequence of timeline segments where each segment has a 1 to 1 mapping with a SMPTE-TT document.

Closed Captions: Captions that may be displayed or not displayed depending on user preference.

Datatype: A metadata description defining the encoding format of data.

Div: A container section of a TTML document as defined by the W3C TTML standard.

Infoset: an abstract Data Model describing the information available from an XML document.

Namespace: mechanism for scoping the definitions of names in an XML document.

Presentation processor: Hardware or software that creates a visible rendition of a document as defined in W3C TTML.

Subtitle: textual representation of the audio track, usually just the dialog and usually in a language other than the audio track dialog, intended for foreign language audiences.

SMPTE-TT document: an encoding of some XML content conforming to the SMPTE-TT document type.

Timed Text: generically all text data sources, including Captions and Subtitles.

XInclude: A mechanism for including some XML content within another XML document.

4.2 Acronyms

CEA: Consumer Electronics Association

DVB: Digital Video Broadcasting

MIME : Multipurpose Internet Mail Extensions

PNG: Portable Network Graphics

SMPTE-TT: SMPTE Profile of TTML

TTML (W3C): Timed Text Markup Language

URI: Universal Resource Identifier

W3C: World Wide Web Consortium

XML: eXtensible Markup Language

XSL: Extensible Stylesheet Language

5 SMPTE-TT Format

5.1 Translation Modes

5.1.1 Requirements

A SMPTE-TT document describes closed caption or subtitle data for association with a given media asset. Such documents can either be created directly or translated from an existing data source. In this Standard, the term SMPTE-TT document is used to describe either a directly created document or one translated from an existing data source, which will be referred to in this Standard as a legacy format.

In order for users to correctly interpret a translation of a legacy format, it may be required that all display features of the original are preserved in case they are being used to convey meaning. Similarly, the timing of data extraction and preservation of latency, display modes, and any other features that affect a display "instant" may also be crucial. Translation accuracy may be required to ensure that no meaning is missed by the viewer after the translation.

SMPTE-TT, however, allows for features, especially in the area of internationalization and styling that cannot be specified in some legacy formats such as CEA-608. In some cases, it may be preferable to replace legacy features with new presentation modes where this does not affect the meaning of the content.

In order that additional features of SMPTE-TT may be used when translating, this Standard defines two modes of translation: Preserved and Enhanced to meet these two cases.

5.1.2 Modes

The mode used for translation shall be recorded in the smpte:information element as defined in Section 5.7.4. Each translated document shall use one of these modes exclusively. Specifics of what constitutes each of these modes for a given legacy format is outside the scope of this Standard and is addressed in the RPs defining the mappings from other formats into SMPTE-TT.

5.1.2.1 Preserved mode

In Preserved mode, all legacy data shall be correctly converted (e.g., CEA-608 pop-on, roll-up, etc.) such that the resulting TTML rendering is as accurate a representation of the original rendering as possible, given the constraints of the formats.

In Preserved mode, if some presentation feature of the original file cannot be captured by any SMPTE-TT feature, this shall be indicated in the header of the file using the smpte:information element as defined in Section 5.7.4. The translation shall continue to preserve as much as possible of the remainder of the document.

5.1.2.2 Enhanced mode

In Enhanced mode, SMPTE-TT shall be used to convey the semantic information of the legacy data, but may differ in its exact presentation. Legacy data may be more liberally interpreted so that exact presentation

fidelity may not be preserved (e.g., matching of visual effects, typography, and coloring may not be required). All content presented in the same style in the legacy data however, shall be translated to a common destination style.

In Enhanced mode, exact output positioning need not be preserved, but may be mapped to approximately equivalent grouped areas (e.g., upper third, lower third).

In Enhanced mode, exact timing need not be preserved (e.g., in the case of CEA-608, frame accurate replication of onset of each pair of letters is not required, and all characters in a line may be aggregated into a single onset time where the onset time of the aggregated group is equal to the original onset time of first visible character in the group, and the outgoing time of the aggregated group, the outgoing time of the last character in the group).

5.2 Default Styles

TTML leaves some initial style values to be defined in context. When computing style inheritance, the set of initial style values shall be set according to Table 1.

Table 1 – Initial style values

tts:color	white

When a SMPTE-TT document with no region defined (using the TTML default region) is presented, the style initial values of Table 2 shall be used.

Table 2 – Initial style values – no region defined

tts:displayAlign	after
tts:textAlign	center

5.3 SMPTE Namespaces

TTML allows for metadata to be added in private namespaces. SMPTE-TT documents need not use the prefixes in Table 3 although such prefixes are used in examples throughout this specification. It is assumed that if a full namespace declaration is not given in an example XML fragment then the definitions in Table 3 are set in an outer context.

Table 3 – Namespaces

Name	Prefix	Value
SMPTE	smpte:	http://www.smpte-ra.org/schemas/2052-1/2013/smpte-tt

5.4 Tunneling Transport

5.4.1 Carriage of Binary Data

SMPTE-TT is an XML (as defined in W3C Extensible Markup Language) based format; in order to carry binary data, that data shall be converted into a text format suitable for inclusion into XML. In this section, mechanisms and protocols are described for converting the input formats into standard binary formats and then encoding these into SMPTE-TT.

This carriage of the original binary data is known as “tunneling” and allows content producers to preserve the pre-existing Timed Text format in cases where the ultimate destination may not be capable of utilizing the TTML translation or where archiving of the exact presentation is required.

Binary data shall be carried in SMPTE-TT using a TTML Metadata element and the smpte:data element as shown in Table 4.

Table 4 – Example of binary data

```
<metadata>
   <smpte:data encoding=”Base64” datatype=”type”>
     encoded data here.
   </smpte:data >
 </metadata>
```

The value of the datatype attribute is specific to each binary format and is not defined in this document, but may be found in the RPs defining the mappings from other formats into SMPTE-TT.

5.5 Pre-rendered Backgrounds

5.5.1 Introduction (Informative)

Some legacy formats (e.g., DVB Subtitles (see ETSI EN 300 7430)) are encoded as image data and no original text version is available. In addition, they may contain logos or other non-text presentations which cannot be readily converted into text. This complicates pure translation into SMPTE-TT since it requires some kind of image processing techniques, such as optical character recognition, to recover the text information. This process might not be practical or might be lossy – especially if the font information is variable or unknown.

While tunneling of the original data is possible as a complete stream, in some instances a hybrid approach is desirable where the image data is used in the TTML presentation and the recovered text is made available computationally but not rendered.

TTML defines that the background rectangle of a rendered area may be a single color. SMPTE-TT provides an extension: the smpte:backgroundImage attribute which is defined for the element. This allows a pre-rendered image be used as the fill for a specified div in addition to its background color.

5.5.2 smpte:backgroundImage

The smpte:backgroundImage attribute shall be used to specify a style property that defines the background image of a rendered area.

This attribute may be specified by any element type that permits use of attributes in the TT Style Namespace; however, this attribute shall apply as a style property only to those element types indicated in Table 5.

Table 5 – Element types that can be used with smpte:backgroundImage

Values:	<uri-specification> | none
Initial:	None
Applies to:	Div
Inherited:	No
Percentages:	N/A
Animatable:	None

The tts:backgroundImage style is illustrated by the example shown in Table 6.

Table 6 – Example of embedding a logo (XML)

```
<?xml version="1.0" encoding="UTF-8"?>
<tt xmlns:smpte="http://www.smpte-ra.org/schemas/2052-1/2013/smpte-tt"
    xmlns="http://www.w3.org/ns/ttml"
    xmlns:ttm="http://www.w3.org/ns/ttml#metadata"
    xmlns:tts="http://www.w3.org/ns/ttml#styling"
    xml:lang="en">
  <head>
    <ttm:title>Example using pre encoded images in metadata</ttm:title>
    <ttm:desc>This example embeds SMPTE Logo in timed text; </ttm:desc>
    <metadata>
      <smpte:image xml:id="SMPTE_logo16" imageType ="PNG" encoding="Base64">
        iVBORw0KGgoAAAANSUhEUgAAAGQAAABHCAMAAADGBBL+AAAAGXRFWHRTb2Z0
        d2FyZQBBZG9iZSBJbWFnZVJlYWR5ccllPAAAADBQTFRFKSO+/f39mJ/ax8fH
        kZGSub3lfcPmTU1vcnjPm+HwSUvA7vDysLCw5OTn1tbc////oFSeJQAAABB0
        Uk5T//////////////////AOAjXRkAAAXhSURBVHja7FnRltsqDERgbIJA
        +f+/7YywkzgbZ+Pdnvtyy/bYJLEZRhoJRMP1P2jhH8hvQIRtvV7HdXzf7da6
        iqj/PbWHIXbffAG5Pgx/e0zatExoDX/oLVOXr+1ppm9AvjYRS6klMOlbs5ZS
        l7/oE0mtcXhlW6+O00z+EojYlBwCo06NbZpSHzhp+pTNexCdwAJDJhiMgydT
        gCXDHax6a/p7ENgeEHAJYNzV5oIyfk6ESa3/EkTobQ4n0gkGG+Eb7xi/wm9d
        W/oViLTWu2DGmPoENmwtGXuW2iCkJlP6BYhjwAmcM50/DOVGowpaWsnIB1wO
        QYDhnqCsbqGXbj3qGqBE+Z5LOJJuokhVEXV2D+90zyTW8QgeE5NvA/MApLut
        QMVglHuWsDsInA/dKbnod0p+DSITMTBX3QampMxi8vyoGyC4mgeqnAehsYhh
        HKvlGEuptYYa0HCrtZQYW4YZ1chX2nuDhdeBzijPGB6Dh8MGsBgjw1LbeZCW
        3o++hwqlIHi+BZGes/m/cWsWzrUKHfublrEg5c4F4jo6NxAtu1eWVE6CBEvL
        2uvXHgoAVK4a6iOIWahWQgQDAE5niYRQ5A6ilSAXgJTyCJIzJlNDzKHmsrR4
        GqRqc5Ss0OI8Q49Vr/vlVx9HrUuqq0cpoJuSAoVbivddxzuUbA5iEkO9XLz7
        5PgdSJi63yIEHErPA8NyDVlygWFLj3BwtFijN3dg1MFE8M48M5T0WV2PIMvk
        45aI0ThwHTYv1BAYSOQveAO3jA8l5mGviSiza6jOl4te34IMbWVcI9wkY6K8
        FQLmzi94IU+fQL07Zb44yOUVSH4EMR+WV9i9jN8iZ1h5yVrXXnWqjyJGBpgj
        nRVnj5OrHDCpi/mbFt31JVPPNfYbEw0rSNiB5LSAAIAHF/UY10Mm4824jhM5
        cySOjYnx2drHj3oXYOwbyMyP5CAXPfLJNJjUztFhcEqv+Kh0/MDemChSXLyD
        lMhEhoYHxi73gMkyrWFShCixIlBrGU4QpKcSHphIhprvIGv0M07m9+q6JZUo
        nRKKitkNOek6651PygqiW2IZII8rzKFPVt0igIVL1lDXLf43n9wdDyYLXVjK
        rHANmRybq95B6A7M01IZUy+3bP1KXX3EyTxfRJlf9HKZ+zETdy0vcEdZ41FX
        x79i4pfkySvrZQPBtmDWXRb+GoxxRH0c4THU9ZVJuGl4RHxF4pq5AyhFx8p1
        YK6Wt4hn9geAMWU5k2cQYT72DFll5PoRjAyYL1k4b6Rdw7Ym4czsRCrUzsjC
        dZOEv5GxbOc88qlO4QaCBFlDfyfhsKyaoRm2ZaOO9WSVK+JmZEV+Xx/8zgQ5
        XqPM7BgE8jq/MgZNG8j4TC7xKU6ym7b4bfnRGs/lpMJOcXAsc6l5X/327r6M
        3Z27SD0LkhjvFcIKio0EOjVcpJf4CJLH4uHm5LKVT++7XFtDVIgSgtjTluhq
        cW2Z/1BHnaSSZWp4UzUjPDTOF529k99sU7FZz77J/oADNJebNDu/q8cemlUO
        4iCW+hrLt/Y5N69Z+w823FeUNagdWOiwpAKalw/kxivMiirFKwevglCd6g9A
        hJUoUVCpqa21lWLMtQBaKylUDaOeSzJOYU7WjCzY4RqgeQV8KxS3HjngB2Ks
        xuKymeVU9UsajoKartmr6pfYxqdGySguzNeHOuGIJoqnUWNj0hPL1HthCj9M
        BPajFtWpr1mjlhzPgVxtUj9Yof1b2k4kvBmPQrywhK0mW6kDpOcPQHZwPKdx
        MvS43s9WxJF5hOPOXyNEoOdDc8keRHbFvImTMdbCN8fTSoAAjS4Px0RKlI8c
        v7ecNCfDMwOepigdz05KXr2LTulhe4VF/UDD4ckhT5+cjOMYz+zWkzvSg9nS
        9OEJYZC3z4m2xlM6P7/rPK3zjodqa58eQn4D4tmSB0Y8dTSeqJq7vredpX4N
        wuI4TXTDlk94HmknTmw/AnGfgs96AN2S6hmIIeHP2qDBBPnvfx3+gfzfQP4I
        MABfygf6VwIk7QAAAABJRU5ErkJggg==
      </smpte:image>
    </metadata>
    <layout>
      <region tts:extent="100px 71px" tts:origin="40% 5%” xml:id="logoArea">
      </region>
    </layout>
  </head>
  <body region="logoArea">
    <div smpte:backgroundImage="#SMPTE_logo16">
      <p tts:display=”none”>SMPTE Logo</p>
    </div>
  </body>
</tt>
```

The above XML results in the image in Figure 1.

Figure 1 – Example of embedding a logo (resulting image)

The semantics of the style property represented by this attribute shall follow those defined in W3C XSL 1.1, Section 7.8.3.

5.5.3 smpte:backgroundImageHorizontal

The smpte:backgroundImageHorizontal attribute shall be used to specify a style property that defines the position of the background image of a rendered area.

This attribute may be specified by any element type that permits use of attributes in the TT Style Namespace; however, this attribute shall apply as a style property only to those element types indicated in Table 7.

Table 7 – Element types that can be used with smpte:backgroundImageHorizontal

Values:	<percentage> | <length> | left | center | right | inherit
Initial:	Center
Applies to:	Div
Inherited:	No
Percentages:	See XSL 1.1, Section 7.8.5
Animatable:	None

The semantics of this attribute are defined by W3C XSL 1.1, Section 7.8.5.

5.5.4 smpte:backgroundImageVertical

The smpte:backgroundImageVertical attribute shall be used to specify a style property that defines the position of the background image of a rendered area.

This attribute may be specified by any element type that permits use of attributes in the TT Style Namespace; however, this attribute shall apply as a style property only to those element types indicated in Table 8

Table 8 – Element types that can be used with smpte:backgroundImageVertical

Values:	<percentage> | <length> | left | center | right | inherit
Initial:	Center
Applies to:	Div
Inherited:	No
Percentages:	See W3C XSL 1.1, Section 7.8.6
Animatable:	None

The semantics of this attribute are defined by W3C XSL 1.1, Section 7.8.6.

5.5.5 Supported image types

If the URI reference has a #fragment extension, then this shall refer to an element in the same document with the id equal to the value of the #fragment. A #fragment is the trailing part of a URI often used to point to a subsection of a resource pointed to be the leading part. This element shall be a smpte:image element and the data in that element shall be treated as an encoded image in the format defined by the element. If the image referencing or decoding fails, the tts:backgroundColor attribute shall be used.

If the URI reference does not have a #fragment extension, then it shall refer to an external image resource. This reference shall resolve to an image of the pre-rendered content of that area. If the URI reference is external to the document, then the filename extension in the URI shall match the encoding type of the image when using one of the MIME types in Table 9. The URI mapping is not specified by this Standard, but is left to the definition of the container or delivery system in which the SMPTE-TT is used. The image formats in Table 9 shall be supported by a SMPTE-TT presentation processor.

Table 9 – Supported image formats

Format	Code	MIME type	File extension	Reference
PNG	PNG	image/png	.png	Portable Network Graphic ISO/IEC 15948:2003

5.5.6 Rendering

The referenced image shall be rendered in accordance with the XSL background-image trait (see W3C XSL 1.1, Section 7.8.3). The min-height and min-width of the area associated with the image shall be set to the intrinsic height and width of the image source. The foreground color for additional marks should be set to transparent. Presentation processors may render foreground marks over a background-image.

The background-repeat property shall be considered to have the value no-repeat.

The presented image shall not be scaled and the XSL background-color trait shall be visible for any background areas of the outside the image, therefore, authors should ensure that the div will be sized to match the given pre-rendering.

5.6 Font resolution

TTML specifies fonts by named strings. SMPTE-TT does not define specific fonts or font embedding semantics; however, any system delivering SMPTE-TT shall define a mechanism for mapping from <fontFamily> and <genericFontFamily> strings in a SMPTE-TT document to a set of known or delivered font resources.

5.7 SMPTE Metadata XML Vocabulary

5.7.1 Conventions

The grammar definitions in this Standard follow the conventions of TTML. An informative XSD schema module of the SMPTE additions to TTML is published as ST 2052-1a, and online here:
http://www.smpte-ra.org/schemas/2052-1/2013/smpte-tt.xsd

5.7.2 smpte:data

The data element shall be used to record binary data of the input format used to generate the SMPTE-TT document. The data element shall have as a content model a text string that is the encoded binary data in the encoding format indicated by the encoding attribute.

5.7.2.1 XML Representation – Element Information Item: data

<data
encoding = (Base64)

datatype = string

xml:id = ID

>

Content: PCDATA

</data>

Data which is not one of the data types defined by this Standard may also be tunneled using this mechanism. In such a case, the datatype attribute value shall have a prefix x- to indicate a private data tunneling, as shown in the example in Table 10:

Table 10 – Example of proprietary data

```
<metadata>
  <smpte:data encoding="Base64" datatype="x-privateTextType">
    encoded data here.
  </smpte:data >
</metadata>
```

All other datatype values are SMPTE Reserved.

Each smpte:data element present in a conforming SMPTE-TT document shall be a child element of a TTML <metadata> element. The smpte:data element may be present in any location supported by <metadata>.

The presentation semantics of this element shall be the reconstitution of the legacy format, e.g., for use in a consumer device that cannot display TTML timed text. Transformation engines shall preserve this data if, and only if, the transformation they perform preserves the presentation semantics of the document; otherwise the transformation shall remove this element.

5.7.3 smpte:image

The image element shall be used to record a pre-rendered image (e.g., for DVB subtitle) where the underlying text or font is not available. This may be referenced by the smpte:backgroundImage style attribute.

5.7.3.1 XML Representation – Element Information Item: image

<image

encoding = (Base64)

imageType PNG

xml:id = ID

>

Content: PCDATA

</image>

Each smpte:image element present in a conforming SMPTE-TT document shall be a child element of a TTML <metadata> element.

The encoding attribute indicates the form of encoding, the defined value is Base64 indicating the encoding format of IETF RFC 4648.

The presentation semantics of this element shall be as defined in Section 5.5.2.

5.7.4 smpte:information

The information element shall record details about the conversion process, including the type of data from which the document was translated.

5.7.4.1 XML Representation – Element Information Item: information

<information
origin = (<URI>| **NONE**)

threshold = <float>

mode = (Enhanced | Preserved)

xml:id = ID

{one or more attributes in smpte metadata namespaces}

>

Content: EMPTY

</ information>

Each smpte:information element present in a conforming SMPTE-TT document shall be a child element of a TTML <metadata> element. Only a single smpte:information element shall be present in a conforming SMPTE-TT document and this shall be a child element of a TTML <head> element, within <metadata>. No presentation semantics for this element are defined

Transformation engines shall preserve this data if, and only if, the transformation that they perform preserves the presentation semantics of the document; otherwise the transformation shall remove this element.

The origin attribute specifies the URI (see IETF RFC 2396 (August 1998),) for the source format for the translation; the default value shall be NONE. The NONE value shall only be used for this attribute if the file was not translated from any prior data (e.g., if it is generated from an authoring tool), otherwise a specific value shall be used. A proprietary value is any fully qualified URI indicating transformation of a format not defined by any SMPTE specification.

The mode attribute shall indicate whether the translated document is considered an exact and faithful representation of the incoming format (Preserved) or whether the translated document has been modified to take advantage of additional features of SMPTE-TT while preserving essential semantics (Enhanced). The default value is Preserved.

The threshold attribute value shall be a real number indicating duration in fractions of a second that records the threshold time used during the translation to suppress the inclusion of temporary Timed Text states. The default threshold attribute value shall be $1/20^{th}$ of a second.

This element may carry one or more attributes in the smpte namespace.

5.8 SMPTE-TT Features from TTML

The SMPTE-TT Profile defined in Table 11 indicates the features of TTML for both transformation and presentation processing in order to accommodate all of the translation rules defined by this specification. This profile shall be referenced in a conforming SMPTE-TT document by the designator: http://www.smpte-ra.org/schemas/2052-1/2013/profiles/smpte-tt-full. This designator may be used in the same manner as designators in TTML Section 5.3.

Table 11 – SMPTE-TT profile

```
<?xml version="1.0" encoding="utf-8"?>
<!-- this file defines the "SMPTE-TT" profile of ttml -->
    <ttp:profile use="http://www.w3.org/ns/ttml/profile/dfxp-full">
      <extensions xml:base=" http://www.smpte.org/23b/smpte-tt/extension/">
        <!-- required (mandatory) extension support -->
        <extension value="required">#data</extension>
        <extension value="required">#image</extension>
        <extension value="required">#information</extension>
      </extensions>
    </ttp:profile>
```

5.8.1 Time

The timeBase should be “media” and thus the timeBase’s “smpte” and “clock” should not be used.

The Time Expression can be ether offset-time or clock-time. If offset-time is used, then the metric shall be “t” (tick).

If not specified, the frameRate is 30 fps.

5.9 SMPTE-TT Specific Features

5.9.1 #data

A SMPTE-TT transformation processor shall support, in the sense defined in W3C TTML, the #data feature by recognizing and being capable of transforming the following vocabulary defined in Section 5.7.2.

smpte:data

A SMPTE-TT presentation processor shall support, in the sense defined in W3C TTML, the #data feature by implementing presentation semantic support for the same vocabulary enumerated above.

5.9.2 #image

A SMPTE-TT transformation processor shall support, in the sense defined in W3C TTML, the #image feature by recognizing and being capable of transforming the following vocabulary defined in Section 5.7.3.

smpte:image
smpte:backgroundImage
smpte:backgroundImage
Horizontal
smpte:backgroundImage
Vertical

A SMPTE-TT presentation processor shall support, in the sense defined in W3C TTML, the #image feature by implementing presentation semantic support for the same vocabulary enumerated above.

5.9.3 #backgroundImage

A SMPTE-TT transformation processor shall support, in the sense defined in W3C TTML, the #backgroundImage feature by recognizing and being capable of transforming the following vocabulary defined in Section 5.7.3.

smpte:backgroundImage
smpte:backgroundImage
Horizontal
smpte:backgroundImage
Vertical

A SMPTE-TT presentation processor shall support, in the sense defined in W3C TTML, the #image feature by implementing presentation semantic support for the same vocabulary enumerated above.

5.9.4 #information

A SMPTE-TT transformation processor shall support, in the sense defined in W3C TTML, the #information feature by recognizing and being capable of transforming the following vocabulary defined by Section 5.7.4.

smpte:information

A SMPTE-TT presentation processor shall support, in the sense defined in W3C TTML, the #information feature by implementing presentation semantic support for the same vocabulary enumerated above.

Annex A Bibliography (Informative)

Note: All references in this document to other SMPTE documents use the current numbering style (e.g. SMPTE ST 334-2:2007) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 334-2-2007). Documents with the same root number (e.g. 334-2) and publication year (e.g. 2007) are functionally identical.

CEA-608-E (ANSI) (2008), Line 21 Data Services

CEA-708-D (2008), Digital Television (DTV) Closed Captioning

ETSI EN 300 743 V1.3.1 (2006-11), Digital Video Broadcasting (DVB); Subtitling Systems

SMPTE ST 334-2:2007, Caption Distribution Packet (CDP) Definition

SMPTE ST 2052-1a:2013, SMPTE-TT XML Schema

W3C XML Information Set (Second Edition), W3C Recommendation, 04 February 2004.

W3C XML Inclusions (XInclude) Version 1.0 (Second Edition), W3C Recommendation, 15 November 2006

W3C Namespaces in XML 1.0 (Third Edition), W3C Recommendation, 8 December 2009

Annex B Chunking (Informative)

This SMPTE-TT Standard does not specifically define the embedding of SMPTE-TT into a larger system; however this Annex describes some general principles of such an embedding based on a concept of chunking.

TTML defines a format that describes the Timed Text over a defined period of time. It separates concerns of styling from the text itself, with the style and region information appearing in a preamble section and the Timed Text appearing afterwards. For longer media, particularly where streaming is required, it is important that both the style and text be available when joining a media presentation part way through. To support this, a delivery mechanism may partition a media timeline into a number of smaller segments and associate with each segment a separate SMPTE-TT document. This document would then be delivered in a “chunk” and the chunks would be delivered as needed with the relevant subsection of the media timeline. The precise definition of what constitutes a chunk is outside the scope of this Standard.

To support such scenarios, a SMPTE-TT delivery mechanism may use a chunking mechanism based on Appendix M of W3C TTML. Each chunk contains a complete TTML file that describes the behavior over a sub-part of the media timeline from a specific offset. The duration of these chunks would be defined by the author or the delivery system. The chunks would be contiguous.

Example: encapsulating a single Caption that runs for the entire duration of a chunk is shown in Table B.1.

Table B.1 – Example of minimal TTML

```
<?xml version="1.0" encoding="UTF-8"?>
<tt xmlns="http://www.w3.org/ns/ttml"
xml:lang="en">
   <body>
    <div> <p>Caption Text</p>
    </div>
   </body>
</tt>
```

Chunk durations would need to be optimized based on the capabilities of the playback engine and the capabilities of the delivery system, where this is known. Such optimization is outside of the scope of this Standard. For example, for situations where the playback engine is unconstrained in resources and the TTML file is delivered out of band with the media, it may suffice to deliver a single chunk for the entire media; whereas in highly constrained situations or where the TTML is interleaved into the media stream, it may be necessary to deliver as many as one chunk per caption. It is expected that most situations will be somewhere between these two extremes.

The mapping of chunks to a file format will depend on the container format used to deliver SMPTE-TT; for example, a mapping to an MPEG-4 packet structure is possible. The TTML parser would reset its in-memory representation (e.g., using XML Infoset (see WRC XML Information Set)) to empty at each chunk boundary. No information from prior chunks should need to be retained in order to process a chunk.

Due to the way in which time extents are defined in TTML, a caption which is present in a pair of adjacent chunks, active at the out-time of the first chunk, and also active at the in-time of the second chunk will be presented without any visible interruption.

A set of chunks creates a logical sequence for the media. The combined time intervals of a logical sequence of related chunks for a particular media need to cover the entire playback time span of the media without gaps or overlap.

A chunk can be defined with no presentation semantics to act as a “filler” chunk. The chunk duration can vary and can range from 0 to the total playback time of the media.

SMPTE-TT data in a chunk could, for delivery convenience, include one or more references to other chunks or to external resources (using, for example, XML techniques such as XInclude (see W3C XML Inclusions) syntax and system specific URI references).

Such references might result in a reduction in overall data size of files (by avoiding replication of common data) but increase the dependence of the stream on its history. For this reason, they should not be required: it should be permitted for each chunk to redefine the static metadata such that no references are necessary.

In the case of multiple chunk sequences for the same media (e.g., for different languages), the embedding system would need to contain a mechanism for identifying and selecting chunks that are in the same logical sequence.