30MR-WD-ST-xxx-LMT-register-2020-04-17(r7-draft).docx Draft SMPTE Standard

Warning

This document is an unpublished, confidential work under development and shall not be referred to as a SMPTE Standard, Recommended Practice, or Engineering Guideline. It is distributed for review and comment; distribution does not constitute publication. Recipients of this document are strongly encouraged to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Information for document editors

- This template is designed to focus on content and not on style. This means that editors should:
 - Refrain from adding styles inline unless absolutely necessary
 - o Refrain from modifying styles unless absolutely necessary
 - o Use the sample styles & layouts for all headings, figures, tables and math
 - o Use the supplied auto numbering styles for headings and annexes

In the event that the content cannot be adequately expressed using this template, contact SMPTE HO.

- The blue text (like this sentence) is editing information or boilerplate. This is information used by SMPTE HQ for publication and should not be changed by the author.
- The document editor must fill in the fields in black before submitting it to the Drafting Group.
- The red text (like this sentence) is helpful information that is not part of the final document. This information shall be deleted prior to FCD Ballot.
- The last page of this document has editing hints and a list of useful reference documents.

Project Group: <The name of the SMPTE Group for this work, this may be a TC>

Project Technology Committee: <The number and name for the parent Technology Committee>

Document type: <ST, ST (Revision), RP or RP (Revision)>

Document state: <WD, CD, FCD, or DP>

Project chair(s): <The SMPTE Group chair>

Document editor(s): Bruce Devlin

Document number: ST XXX

Document title: Language Metadata Table Controlled Vocabulary Register

Title Page

This page will be provided by SMPTE HQ Staff.

See AG-16 clause 3.1 (Title Page), and ISO Directive Part 2 clause 11 (Title).

Proposed SMPTE Standard

This document is subject to change.

Copyright © 2019 SMPTE All rights reserved

Table of Contents	Page	
Foreword	5	
Introduction (Optional/Conditional)	6	
1 Scope	6	
2 Normative References	6	
3 Terms and Definitions	7	
4 XML Schema Definitions	7	
5 The LMT Controlled Vocabulary Register and its Terms	8	
5.1. General (Informative)	8	
5.2. Schema	8	
5.3. Elements and Attributes in the Register	8	
5.3.1. term	8	
5.3.2. termId	8	
5.3.3. termUpdate	8	
5.3.4. termName	8	
5.3.5. termVocabulary	8	
5.3.6. termNote	8	
5.3.7. relation	11	
6 Register updates	12	
6.1. SMPTE LMT Repository and Submission Package definition	12	
6.2. SMPTE repository structure	12	
6.3. QC of the submission package	13	
6.4. SMPTE Process	13	
6.5. SMPTE Headquarters (HQ) publishing	14	
7 Register Maintenance	14	
7.1. Official Maintainers	14	
7.2. Becoming a maintainer	14	
7.3. Removal of maintainers	15	
Annex A LMT XML Schema (normative)	16	
Annex B Change Control – to be deleted before submission	19	

Foreword

See AG-16 3.2 (Foreword), and ISO Directive Part 2 clause 12 (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 its Standards Operations Manual. This SMPTE Engineering Document was prepared by Technology Committee <TC-number-and-name>.

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; then formal languages; then figures; and then any other language forms.

If this is a revision, a topical list of changes [should/shall] be included here.

Introduction (Optional/Conditional)

The introduction provides specific information or commentary about the technical content of the document, and about the reasons prompting its preparation. See AG-16 clause 3.3 (Introduction), AG-16 clause 4.2 (Conformance Terms), and ISO Directive Part 2 clause 13 (Introduction).

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

The Language Metadata Table (LMT) is an expandable mapping resource that is used to organize language metadata via locations and dialects. It was created to provide a unified source of reference for language codes for use throughout the media and entertainment industries.

The LMT is a controlled vocabulary of codes from the larger IETF BCP 47 dictionary. The goal of LMT is to encourage interoperability in code usage by restricting the options to those in practical use within the industry.

The contributions to the register are delivered to SMPTE by Maintainer organizations. The SMPTE process is used to scrutinize the submissions prior to publication.

[Editors notes: The following paragraph will be replaced with the appropriate patent information during the SMPTE Headquarters publication process.]

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.

1 Scope

The scope clearly defines the subject of the document and the aspects covered, thereby indicating the limits of applicability of the document. See AG-16 clause 3.4 (Scope), and ISO Directive Part 2 clause 14 (Scope).

This SMPTE Standard defines a method for maintaining and validating a published register of Language Metadata Identifier Records.

2 Normative References

The normative references clause lists, for information, those documents which are cited normatively in the document. See AG-16 clause 3.5 (Normative References), AG-16 clause

4.3 (Normative References to Standards and Recommended Practices), and the ISO Directives Clause 15 (Normative References).

The following <document-type> contains provisions that, through reference in this text, constitute provisions of this standard. Dated references require that the specific edition cited shall be used as the reference. Undated citations refer to the edition of the referenced document (including any amendments) current at the date of publication of this document. All <document-type>s are subject to revision, and users of this engineering document are encouraged to investigate the possibility of applying the most recent edition of any undated reference.

SMPTE ST 2029:202x Uniform Resource Names for SMPTE Resources

IETF BCP 47 Tags for Identifying Languages, https://tools.ietf.org/html/bcp47

IETF RFC 5119 SMPTE URN Definition

SMPTE Registration Authority, https://smpte-ra.org.

W3C XML Schema Part 1: Structures

W3C XML Schema Part 2: Datatypes

3 Terms and Definitions

The terms and definitions clause provide definitions necessary for the understanding of certain terms used in the document. See AG-16 clause 3.6 (Terms and Definitions). AG-16 clause 4.4 (Terms and Definitions), and ISO Directive Part 2 clause 16 (Terms and Definitions).

For the purposes of this document, the terms and definitions given in SMPTE ST 2029 and IETF BCP 47 apply.

4 XML Schema Definitions

This section shall apply whenever a data structure is specified using XML schema definitions as specified in W3C XML Schema Part 1: Structures and W3C XML Schema Part 2: Datatypes.

In order to avoid duplication between text and schema, the cardinality and default values of elements are specified in the schema definitions only.

In the event of a conflict between schema definitions and the prose, the prose shall take precedence

5 The LMT Controlled Vocabulary Register and its Terms

5.1. General (Informative)

The register is represented as an XML Document complying with the XML Schema defined in Annex A. Several views of this data will be presented on the SMPTE registers website along with the XML document. One or more maintainers are responsible for providing updates to SMPTE.

5.2. Schema

The schema for the published data is given in Annex A. The semantics of the schema elements and attributes are defined in section 5.3. The hierarchy and number of occurrences of each element are defined in the schema. Attributes that are required are defined in the schema.

5.3. Elements and Attributes in the Register

5.3.1. term

The term element defines a term in the register.

5.3.2. termId

A unique numerical identifier required for every term.

5.3.3. termUpdate

Indicates the status of the term. It shall have the value add or deprecate or delete?????.

5.3.4. termName

Shall be a valid IETF BCP-47 langtag value

5.3.5. termVocabulary

Shall take one of the enumerated values from the LMTtermVocabularyType in the Schema.

- Language Metadata Table LMT shall be used for individual LMT Language Codes
- Language Groupings LMT shall be used to define a Language Group.

5.3.6. termNote

Each term element shall have the required termNote child elements defined in Table 1. The label attribute contains the name of the termNote field, ex: Language Group Name, Language Group Tag, Language Group Code.

5.3.6.1. label attribute

The label attribute shall take one of the enumerated values from the LMTtermNoteLabelType as defined in the schema.

 $Table\ 1\ LMT term Note Label Type\ semantics$

Enumerated Value	Definition
Language Group Name	The Group name is equivalent to the generic language name. Language dialects are subordinate to their language grouping. Ex: Armenian - Western falls under Armenian Family. Required when termVocabulary is Language Groupings LMT Forbidden when termVocabulary is Language Metadata Table LMT
Language Group Tag	IETF BCP 47 tag. Required when termVocabulary is Language Groupings LMT Forbidden when termVocabulary is Language Metadata Table LMT
Language Group Code	URN or URI for the language group term. See 5.3.6.2 Required when termVocabulary is Language Groupings LMT Forbidden when termVocabulary is Language Metadata Table LMT
Audio Language Tag	IETF BCP 47 language tag. Typically spoken/audio language. Forbidden when termVocabulary is Language Groupings LMT Required when termVocabulary is Language Metadata Table LMT
Long Description 1	Description of language name in Latin script following IETF BCP 47 standard Forbidden when termVocabulary is Language Groupings LMT Required when termVocabulary is Language Metadata Table LMT
Long Description 2	Alternate description of language name in Latin script following IETF BCP 47 standard Forbidden when termVocabulary is Language Groupings LMT Optional when termVocabulary is Language Metadata Table LMT
Audio Language Display Name 1	Endonym of audio language. Typically, the same as Visual Language Display Name 1 but not always. Forbidden when termVocabulary is Language Groupings LMT Optional when termVocabulary is Language Metadata Table LMT

Audio Language Display Name 2	Alternate endonym of audio language. Typically, the same as Visual Language Display Name 2 but not always. Forbidden when termVocabulary is Language Groupings LMT Optional when termVocabulary is Language Metadata Table LMT
Visual Language Tag 1	Script in which language is written following IETF BCP 47 standard (which calls for the tags to be presented in Latin Script). Forbidden when termVocabulary is Language Groupings LMT Optional when termVocabulary is Language Metadata Table LMT
Visual Language Tag 2	Alternate script in which language is written following IETF BCP 47 standard (which calls for the tags to be presented in Latin Script). Forbidden when termVocabulary is Language Groupings LMT Optional when termVocabulary is Language Metadata Table LMT
Visual Language Display Name 1	Endonym of written language. Typically the same as Audio Language Display Name 1 but not always. Forbidden when termVocabulary is Language Groupings LMT Optional when termVocabulary is Language Metadata Table LMT
Visual Language Display Name 2	Alternate written endonym. Typically, the same as Audio Language Display Name 1 but not always. Forbidden when termVocabulary is Language Groupings LMT Optional when termVocabulary is Language Metadata Table LMT
Code	URN or URI for each language tag term. See 5.3.6.2 Forbidden when termVocabulary is Language Groupings LMT Required when termVocabulary is Language Metadata Table LMT
Scopes	A list of comma separated values to restrict the scope of the usage of the tag. If not specified the scope is global Optional EXAMPLE: "ISDCF, MovieLabs"

5.3.6.2. urn encoding of termNote values IETF RFC 2141 defines the general syntax of URNs as:

```
<URN> ::= "urn:" <NID> ":" <NSS>
```

SMPTE URNs use the NID smpte, which was defined by IETF RFC 5119 for registration in the IANA registry of URN NIDs.

The first part of the Namespace Specific String for SMPTE Registers is defined in SMPTE ST 2029:

```
<NSS> ::= "ra" ":" <REGISTER_IDENTIFIER> ":" <REGISTER_VALUE>
```

The register specific identifier for this document shall be

```
<REGISTER IDENTIFIER> ::= "lmt"
```

The register specific value for this document shall be the value of the termNote XML element with the attribute Audio Language Tag.

```
<REGISTER VALUE> ::= "lmt"
```

EXAMPLE: The termNote XML element:

```
<termNote label="Audio Language Tag">es-419</termNote>
```

Will be urn encoded as

urn:smpte:ra:lmt:es-419

5.3.7. relation

5.3.7.1. Presence of a relation element

This element indicates that there are one or more related terms to this LMT code value.

If the termVocabulary child of a term element is set to Language Metadata Table LMT then the relation will link to a term with its termVocabulary child set to Language Grouping LMT.

If the termVocabulary child of a term element is set to Language Grouping LMT then there will be on or more relation elements that link to term elements with termVocabulary child set to Language Metadata Table LMT.

5.3.7.2. relationType

The schema allows for language relationships through the use of Language Grouping.

There is no requirement a specific term to be part of a Language Group. The use of IETF BCP 47 "Macrolanguage" and "Language Family" designations allow for alphabetical sorting by grouping, keeping languages like Chinese together. If not, languages like Mandarin and Cantonese would separate. A simple hierarchy allows for the maximum flexibility. Some language grouping examples are:

Greek: to account for ancient vs modern

English: British, Canadian, Australian, American, etc

Spanish: Latin American vs European, Mexican vs Argentinian

Special: for codes such as "und" (undetermined)

The following	enumerations of	of the relationT	'vne elemen	t are permitted.
THE TOHOWING	chamer actors (or the relations	y pe cicinen	t are permitted.

Value	Meaning	Notes
EQT	Equivalent to	This term is equivalent to the related term
ВТ	Broad Term	The related term is a broader term than this term EXAMPLE: the relation element for en-AU (Australian English) has a BT relation to en (Generic English)
NT	Narrow Term	The related term is a narrower term than this term EXAMPLE: the relation element for en (Generic English) has a NT relation to en-au (Australian English)
TT	Top Term	The related term is the Top Most Broad Term.

5.3.7.3. termId

Required. Shall be set to the termId value of the related element in the LMT XML document

5.3.7.4. termName

Required. Shall be an identical copy of the termName value of the related term referenced by termId.

5.3.7.5. termVocabulary

Required. Shall be an identical copy of the termVocabulary value of the related term referenced by termId.

6 Register updates

6.1. SMPTE LMT Repository and Submission Package definition

SMPTE manages its copy of the register using a Git repository. The repository contains the files detailed below. An update request to start the SMPTE process is generated by a maintainer issuing a pull request to SMPTE. Details are provided in Administrative Guideline AG-xx.

6.2. SMPTE repository structure

The repository shall contain one of each files listed below. All other files in the repository will be ignored and may be deleted at the discretion of SMPTE HQ.

- 1. lmt.xml An XML document of the new LMT version
- 2. lmt-ref.xml An XML document of the current published lmt.xml

- a. Line endings and indent style shall match the new document
- 3. lmt-diff.xml The output of the Unix diff -a -b lmt.xml lmt-ref.xml
- 4. lmt-release-summary.md A markdown narrative to be published with this version
- 5. lmt-control.pdf A PDF of the published version of this document for reference by maintainers.
- 6. README.md A markdown description of the repo referencing the PDF in the repo for contact information.

6.3. QC of the submission package

When the submission package is received, SMPTE will validate:

- 1. The request was received from one of the approved register maintainers.
- 2. The Zthes/LMTMetadata/version field has been updated
- 3. The Zthes/LMTMetadata/DatesIS08601/release field has been updated
- 4. The Zthes/LMTMetadata/DatesIS08601/release field is absent
- 5. Line endings shall be unix style (\n) and indenting shall be performed with spaces
- 6. lmt.xml validates against the schema defined in the latest version of this document
- 7. The web pages for the smpte-ra website render correctly
- 8. The provided documents meet SMPTE process rules

6.4. SMPTE Process

The submission package shall be treated as an incoming SMPTE Standard and subject to the procedures of the SMPTE Standards Operations Manual. Specific Instructions are given below.

NOTE: The SMPTE process means that individual submissions might be accepted, rejected or modified by consensus.

- 1. Once a submission is received and has passed validation, a meeting of the LMT working group is convened to elevate the package as a WD to the TC.
- 2. The TC chairs initiate a 2-week pre-FCD review to become a CD. [there may be comments]

- 3. The TC chair asks the SVP to put the package put on the SMPTE-RA website in the public-CD for comment area.
- 4. The TC Chairs commence a 3-week FCD ballot. With the question "Do the changes to the Language Metadata Table Register fulfil the register requirements in the latest version of SMPTE ST xxxx"
- 5. If there are comments from the Public CD process or from the Ballot, they shall be addressed.
- 6. A DP vote follows if there were substantive comments to be addressed
- 7. An ST Audit takes place to verify process was followed.
- 8. The LMT is published in the live area of SMPTE-RA and the public CD version is removed

6.5. SMPTE Headquarters (HQ) publishing

The resulting table shall be made available on the SMPTE-RA website at a permanent URL SMPTE may optionally make available other views of the canonical XML available. Tooling and processing of these alternate views is at the discretion of SMPTE HQ and outside the scope of this document.

7 Register Maintenance

7.1. Official Maintainers

The Automatic update procedure in section 6 is available to SMPTE Standards Members who are identified as official Maintainers by the SMPTE Technology Committee. The list of maintainers shall be made available on the SMPTE Register site.

The role of a maintainer is to represent a group that is active in the maintenance of the controlled vocabulary. A maintainer has the same role as a proponent as defined in the SMPTE Standards Operations Manual.

7.2. Becoming a maintainer

A maintainer shall be a proponent of an approved project to update the register. The project shall state which of the proponents shall be maintainers and whether the proponent is stay a maintainer when the project is complete.

7.3. Removal of maintainers

A maintainer may revoke their status by communicating the desire to do so to the chair of the controlling SMPTE Technology Committee. The controlling SMPTE Technology Committee may remove maintainer status by consensus.

Annex A LMT XML Schema (normative)

This specification is accompanied by the following element, which is an XML schema document as specified in XML Schema Part 1: Structures.

rddXX-202x.xsd

This element collects the XML schema definitions defined in this document. In case of conflict, this document takes precedence.

The schema is presented informatively in the table below

Table A.1 LMT XML Schema

```
s:schema attributeFormDefault="unqualified" elementFormDefault="qualified"
<xs:simpleType name="LMTtermNoteLabelType"
<xs:restriction base="xs:string">
       <xs:enumeration value="Language Group Tag"
<xs:enumeration value="Language Group Code"</pre>
       <xs:enumeration value="Long Description 2" />
<xs:enumeration value="Audio Language Display Name 1" />
      <xs:enumeration value="Audio Language Display Name 2" />
<xs:enumeration value="Visual Language Tag 1" />
<xs:enumeration value="Visual Language Tag 2" />
       <xs:enumeration value="Visual Language Display Name 1" />
<xs:enumeration value="Visual Language Display Name 2" />
<xs:simpleType name="relationTypeType">
  <xs:restriction base="xs:string">
      <xs:lestification base= xototring /
<xs:enumeration value="EQT" /><!-- Equivalent To -->
<xs:enumeration value="BT" /><!-- Broad Type -->
<xs:enumeration value="NT" /><!-- Narrow Type -->
                       <xs:element type="xs:string" name="version" maxOccurs="1" minOccurs="1"/>
<xs:element type="xs:string" name="maintainer" maxOccurs="1" minOccurs="1"/>
<xs:element name="DatesISO8601" maxOccurs="1" minOccurs="1">
                                    <xs:element name="publish" type="xs:date" maxOccurs="1" minOccurs="0"/>
```

An informative picture of the Schema is given in Figure 1.

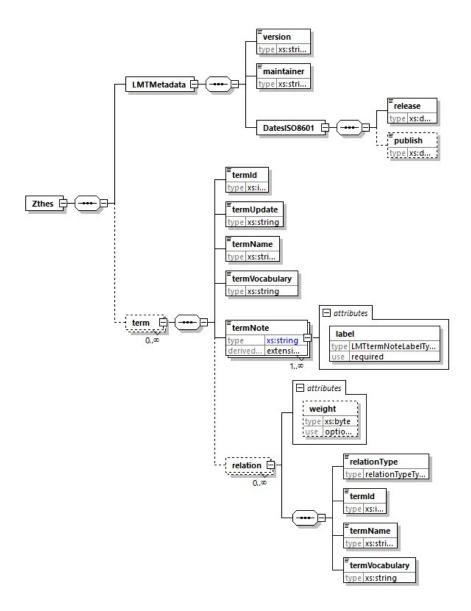


Figure 1 Pictorial View of LMT Schema

Annex B Change Control – to be deleted before submission

- R7 incorporated Movielabs Schema picture (Figure 1) and presented to harmonization meeting #1
- R6 added maintainer text & distributed to harmonization team for comment
- R5 simplified process description post HPA