

SMPTE RECOMMENDED PRACTICE

Text-Based Metadata
Carriage in MXF



Table of Contents	Page
Foreword	2
Intellectual Property	2
Introduction.....	2
1 Scope	3
2 Conformance Notation	3
3 Normative References	3
4 Definition of Acronyms, Terms and Data Types	4
5 General (Informative)	4
6 Generic Stream Partition for Text-Based Metadata.....	5
6.1 Generic Stream Partition Pack.....	5
6.2 Generic Stream Data Element Coding.....	6
6.3 Generic Stream Constraints.....	7
7 Descriptive Metadata Scheme and Sets for Text-Based Metadata.....	7
7.1 Universal Label for Descriptive Metadata Scheme.....	7
7.2 Descriptive Metadata Set Definitions	7
8 Guideline on Text-Based Metadata Implementation.....	10
8.1 Selection of Carriage Mechanism	10
8.2 Text-Based Metadata Identification	11
8.3 Header Metadata Implementation (Informative)	11
Annex A Bibliography (Informative)	14

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 RP 2057 was prepared by Technology Committee 31FS.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Standard. 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.

This document addresses a requirement for carrying text-based metadata with a specified text MIME type encoded using either Unicode UTF-8 or UTF-16 character encoding (such as XML) as it is in a MXF file without encoding each element using the KLV Packet. Text-based metadata can be encapsulated in an MXF file using the following methods as described in the rest of this introduction.

Text-based metadata can be carried using the Generic Stream Container defined in SMPTE ST 410. SMPTE ST 410 was developed to give a mechanism to carry either essence or metadata that cannot suitably be stored in the Header Metadata in a MXF file body using the Generic Stream Container. The essence is limited to 'lumpy' essence that cannot be carried by the Generic Essence Container. The metadata would be bulky text data that cannot be carried by the Header Metadata.

This document therefore specifies the mapping of text-based metadata into the Generic Stream Container along with how to identify the Generic Stream Payload. It also includes the specification of the following requirements described in SMPTE ST 410:

- Specify the format and semantics of the Payload bytes.
- Specify a Universal Label to identify each payload.
- Define the mechanism used to link from the Header Metadata to the Generic Stream Payload by use of the Generic Stream ID.
- Specify whether an index table is forbidden, optional or mandatory for the associated Generic Streams along with the format of the Index Table.

Text-based metadata can be owned by a DM Framework as defined in SMPTE ST 377-1. In this case, text-based metadata is encapsulated in the MXF Header Metadata using Descriptive Metadata Plug-In.

1 Scope

This document defines how to carry text-based metadata with a specified text MIME type encoded using either Unicode UTF-8 or UTF-16 character encoding (such as XML) in a MXF file. This standard defines two methods of text-based metadata carriage. The first method is to carry text-based metadata using the MXF Generic Stream Partition defined in SMPTE ST 410. The second method is to carry it in the MXF Header Metadata by defining a DM Framework which includes Text-based Sets.

This document specifies the Generic Stream Partition Pack for several text-based metadata formats including the KLV coding of the Generic Stream Data Element, and DM Schemes including DM Framework, Sets and Items for text-based metadata.

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

Note: All references in this document to other SMPTE documents use the current numbering style (e.g. SMPTE ST 336:2007) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 336M-2007). Documents with the same root number (e.g. 336) and publication year (e.g. 2007) are functionally identical.

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

SMPTE ST 336:2007, Data Encoding Protocol Using Key-Length-Value

SMPTE ST 377-1:2009, Material Exchange Format (MXF) — File Format Specification

SMPTE ST 395:2003, Television — Metadata Groups Registry Structure

SMPTE ST 400:2004, Television — SMPTE Labels Structure

SMPTE ST 410:2008, Material Exchange Format — Generic Stream Partition

RFC5646 Tags for Identifying Languages

<http://www.iana.org/assignments/media-types/text>, Public Registry of Text MIME Types

ISO/IEC 10646:2003, Information Technology — Universal Multiple-Octet Coded Character Set (UCS)

XML, Extensible Markup Language (XML) 1.0, W3C Recommendation

XML, Extensible Markup Language (XML) 1.1, W3C Recommendation

4 Definition of Acronyms, Terms and Data Types

For the purposes of this document, the terms and definitions given in SMPTE ST 377-1 and SMPTE ST 410 and the following apply.

IANA: Internet Assigned Numbers Authority

MIME: Multipurpose Internet Mail Extensions

Text-based Metadata: Text data with a specified text MIME type registered in Public Registry of Text MIME Types, encoded using either Unicode UTF-8 or UTF-16 character encoding

XML: Extensible Markup Language according to versions 1.0 or 1.1 of the W3C Recommendation

5 General (Informative)

The MXF File Format specification standard, SMPTE ST 377-1, provides the Descriptive Metadata Plug-in mechanism to carry additional Descriptive Metadata elements in the MXF Header Metadata. This document defines the DM Scheme including DM Framework, Sets and Items for Text-based Metadata with a specified text MIME type encoded using either Unicode UTF-8 or UTF-16 character encoding. The Text-based Metadata may include Descriptive metadata or Application metadata.

The Text-based Metadata is referenced from DM Tracks. The DM Tracks may be static, timeline or dynamic DM Tracks depending on applications. Each DM Track owns one or more DM Segments each of which strongly references the DM Framework for Text-based Metadata.

The Text-based DM Framework specifies the Generic Stream Text-based Sets and the UTF-8 or UTF-16 Text based Set as a Subclass of the Text-based Object to support the following two carriage mechanisms:

1. Generic Stream Partitions specified by SMPTE ST 410 are used for "bulky" text-based documents where the Generic Streams Payload is referenced through the Generic Stream Text-Based Set. The Generic Stream Text-based Set properties do not carry the text data, but include the linkage information (Stream SID) of the Generic Stream that contains the Text-based Metadata document.
2. Text-based Metadata is carried as payload of the- UTF-8 or UTF-16 Text-based Set.

Figure 1 illustrates the inheritance hierarchy of text-based Descriptive Metadata Sets by a UML diagram. Examples of the Header Metadata implementation and the linkage to the Text-based Metadata are given in Section 8.3.

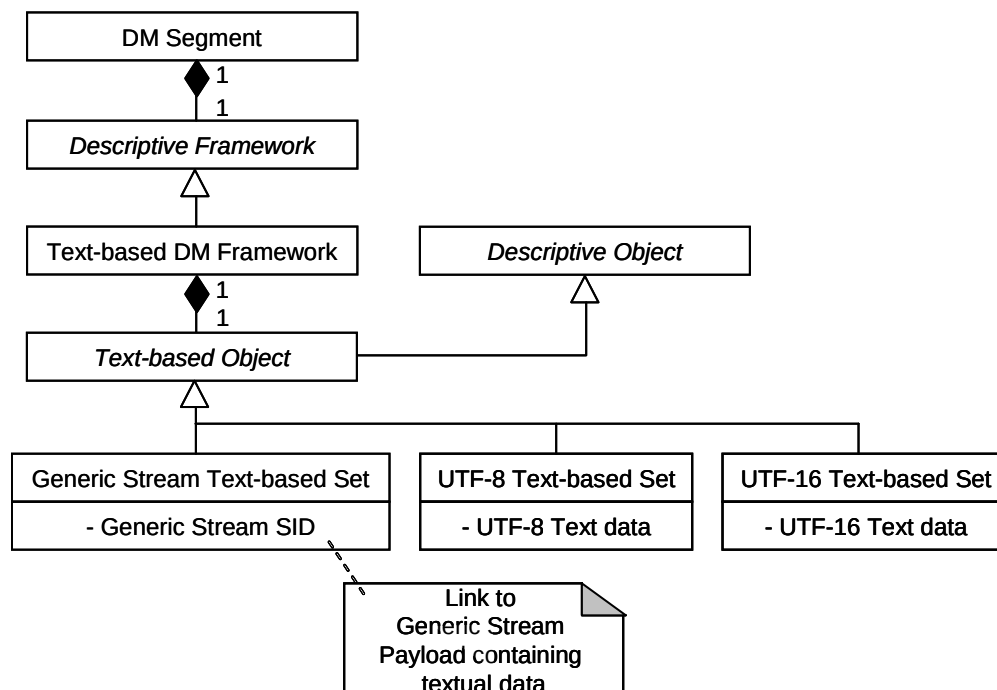


Figure 1 – Inheritance hierarchy of text-based Descriptive Metadata Sets

6 Generic Stream Partition for Text-based Metadata

6.1 Generic Stream Partition Pack

The Generic Stream Partition Pack comprises a Generic Stream Partition Pack Key, a Length and a Value as defined in SMPTE ST 410.

6.1.1 Generic Stream Partition Pack Key

The Key of the Generic Stream Partition Pack shall be as defined in SMPTE ST 410.

6.1.2 Generic Stream Partition Pack Length

The length field shall be as specified in SMPTE ST 410.

6.1.3 Generic Stream Partition Pack Values

The value of each item within the Generic Stream Partition Pack shall be as specified in SMPTE ST 410.

Note: The Generic Stream Payload represents Metadata and not Essence. Consequently, according to SMPTE ST 410, there is not an entry in the Essence Containers Batches of the Partition Packs and the Preface Set to identify the Generic Streams Payload.