

SMPTE STANDARD

D-Cinema Packaging — Timed Text Track File



Page 1 of 14 pages

Table of Contents	Page
Foreword	2
Intellectual Property	2
1 Scope	3
2 Conformance Notation	3
3 Normative References	3
4 Glossary of Acronyms and Terms.....	4
5 Overview (Informative).....	4
6 Ancillary Resource Wrapping.....	6
6.1 Generic Stream Repetition.....	6
6.2 Indexing Generic Stream Data.....	6
7 Timed Text Track File Structure.....	7
7.1 Essence Encoding	7
7.2 Index Table	7
7.3 Timed Text Resource Constraints	7
7.4 Header MetadataConstruction	7
7.5 Essence Descriptors	8
7.6 Header MetadataConstraints	8
7.6.1 General	8
8 Essence Encryption	9
9 Synchronization.....	10
10 Random Access to Ancillary Resources (Informative).....	10
Annex A Labels and Descriptor Sets (Normative)	11
A.1 UL Values	11
A.2 TimedText Descriptor Set.....	12
A.3 TimedText resource SubDescriptor Set	12
Annex B 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 Standard 429-5 was prepared by Technology Committee DC28.

Intellectual Property

SMPTE draws attention to the fact that it is claimed that compliance with this Standard may involve the use of one or more patents or other intellectual property rights (collectively, "IPR"). The Society takes no position concerning the evidence, validity, or scope of this IPR.

Each holder of claimed IPR has assured the Society that it is willing to License all IPR it owns, and any third party IPR it has the right to sublicense, that is essential to the implementation of this Standard to those (Members and non-Members alike) desiring to implement this Standard under reasonable terms and conditions, demonstrably free of discrimination. Each holder of claimed IPR has filed a statement to such effect with SMPTE. Information may be obtained from the Director, Standards & Engineering at SMPTE Headquarters.

Attention is also drawn to the possibility that elements of this Standard may be subject to IPR other than those identified above. The Society shall not be responsible for identifying any or all such IPR.

1 Scope

This standard specifies the format of a Timed Text Track File for the distribution of D-Cinema timed text content using the Material Exchange Format (MXF). The Track File is a container for an XML document and optional support resources such as images or fonts. Encryption is optionally available for protecting against unauthorized disclosure of the file contents.

The standard defines data structures for interchange at the signal interfaces of networks or storage media, but does not define internal storage formats for compliant devices.

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.

[SMPTE 377M-2004] Television — Material Exchange Format (MXF) — File Format Specification

[SMPTE 379M-2004] Television — Material Exchange Format (MXF) — MXF Generic Container

[SMPTE 410-2008] Material Exchange Format — Generic Stream Partition

[SMPTE 429-3-2007] D-Cinema Packaging — Sound and Picture Track File

[SMPTE 429-6-2006] D-Cinema Packaging — MXF Track File Essence Encryption

Internet Engineering Task Force (IETF) (1996, November). [RFC 2045] Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies

Internet Engineering Task Force (IETF) (1996, November). [RFC 2396] Uniform Resource Identifiers (URI): Generic Syntax

Internet Engineering Task Force (IETF) (2005, July). [RFC 4122] A Universally Unique Identifier (UUID) URN Namespace

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

4 Glossary of Acronyms and Terms

Note: See also the list of acronyms and terms in SMPTE 377M.

Closed Caption: Timed-text intended for display on a device other than the theater screen.

Composition: A complete artistic or informational motion picture work, such as a feature, a trailer, or an advertisement, etc.

MXF: Material Exchange Format (see SMPTE 377M).

Open Caption: Timed-text intended for display on the theater screen, usually written in the same language as the sound essence.

Reel: A unit of essence comprising all or part of a composition.

Resource: An integral unit of data, such as an XML document, a font, or a sub-picture image.

Sub-picture: An ancillary image intended for display over a larger main image.

Subtitle: Timed-text intended for display on the theater screen, usually written in a language other than the language of the sound essence.

Timed-text: Text intended for display over a timeline, in synchronization with image and sound essence.

URI: A Uniform Resource Identifier, defined by [RFC 2396], is a text string which uniquely identifies a resource, such as an XML Schema. URI values often identify objects not accessible via a computer network.

UUID: Universally Unique Identifier, defined by [RFC 4122].

XML: eXtensible Markup Language, an abstract syntax for structured text with metadata.

5 Overview (Informative)

D-Cinema compositions may contain subtitles, closed-captions or other forms of textual information in addition to sound and picture essence. The display of textual information varies according to purpose, but the essence encoding can be generalized as a timed-text resource (an XML document) that provides content, position and timing information (i.e., essence plus metadata), plus optional ancillary resources such as fonts and sub-pictures.

In some cases the essence is entirely contained in one or more sub-pictures. In these cases the timed-text resource contains only timing and position information.

Like sound and picture essence for D-Cinema, timed-text essence is encoded reel-by-reel. Each reel of timed-text essence consists of one XML timed-text resource plus optional supporting resources. To simplify the delivery of what may potentially be many dozens of files for a given reel, all of the resources for a reel are wrapped in a single MXF Timed Text Track File as described in this document. To protect against disclosure of the essence data to unauthorized parties, the essence may be encrypted.

The Timed Text Track File is based on the D-Cinema Sound and Picture Track File [SMPTE 429-3] with the exceptions stated in Section 7 of this document, with the additional use of Generic Stream Partitions [SMPTE 410] as required to carry Ancillary Resource data.

Figure 1 below illustrates a simple Timed Text Track File containing text-based subtitle essence. The timed-text resource is contained in the Track File along with a font resource used to render the characters when the text is reproduced on the theater screen.

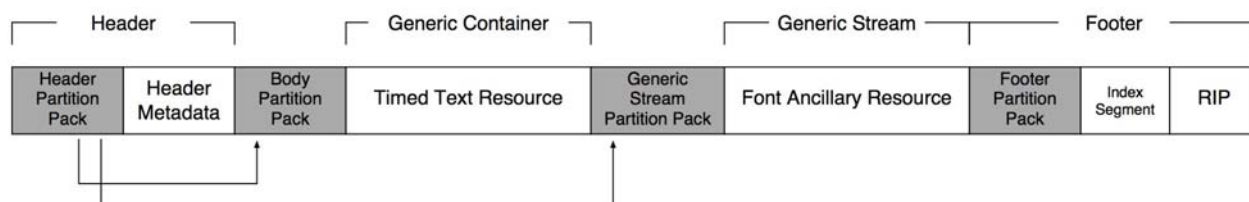


Figure 1 – Example Timed Text Track File Structure using a Font

Figure 2 illustrates a more complex Timed Text Track File containing sub-picture-based subtitles. The timed-text resource is contained in the Track File along with a number of sub-picture resources.

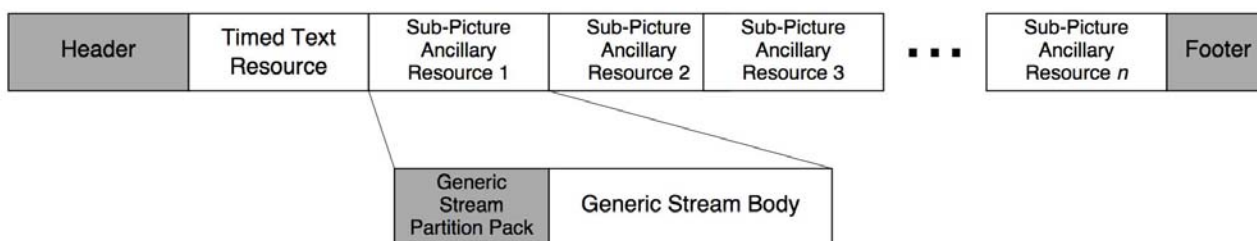


Figure 2 – Example Timed Text Track File Structure using Sub-Picture

6 Ancillary Resource Wrapping

Each Ancillary Resource in a Timed Text Track File shall be entirely contained within an MXF Generic Stream Partition per SMPTE 410. Each Generic Stream Partition in the file shall contain exactly one Ancillary Resource. Each Generic Stream Partition shall have a distinct BodySID per SMPTE 410. The Generic Stream Partition shall consist of a Generic Stream Partition Pack immediately followed by a single KLV packet containing all of the resource data. KLV Fill packets shall not be permitted between the Generic Stream Partition Pack and the resource KLV. The actual format of the resource data is beyond the scope of this document. Consult the defining document for the Timed Text Resource for more information.

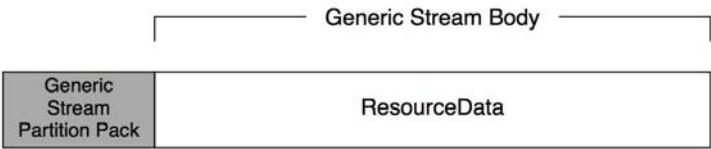


Figure 3 – Timed Text Resource Partition Structure

The Ancillary Resource KLV packet shall be identified by the Default Generic Stream Data Element key (see Generic Stream Data Element coding in SMPTE 410). Data Arrangement bit 1 shall be zero (the KL pair shall not be considered an intrinsic part of the Ancillary Resource data) and bits 3 and 2 shall be one and zero, respectively (the Generic Stream payload is a byte string). Wrapping Signaling bits 1-3 shall be zero (there are no internal access units). The following UL value gives the Default Generic Stream Data Element key, set per the above constraints:

Table 1 – Ancillary Resource Key (hexadecimal)

06.0e.2b.34 01.01.01.0c 0d.01.05.05 01.00.00.00

6.1 Generic Stream Repetition

Generic Stream Repetition, as defined by SMPTE 410, shall not be used.

6.2 Indexing Generic Stream Data

Generic Stream Data shall not be indexed using SMPTE 377M Index Tables. Generic Stream Partitions shall be included in the RIP. See Section 10 for an informative description of locating an Ancillary Resource using its Resource ID as a lookup value.

7 Timed Text Track File Structure

Timed Text Track Files shall conform to the provisions of SMPTE 429-3 with the following exceptions:

- a. The Timed-Text Resource shall be clip-wrapped, not frame-wrapped.
- b. The file shall contain more than three (3) partitions when Ancillary Resources are present.
- c. Essence Constraints shall be determined by this document.

7.1 Essence Encoding

The essence container in the Timed Text Track File shall contain the Timed Text Resource, an XML document that contains all of the timing and position information for the timed-text instances. A Timed Text Track File shall contain exactly one Timed Text Resource.

The Timed Text Resource shall be clip wrapped as a single Data Element in a single Data Essence Item of a Generic Container [SMPTE 379M].

The Timed Text Resource may refer to Ancillary Resources such as fonts and sub-pictures. All Ancillary Resources referenced by the Timed Text Resource shall be contained within the Timed Text Track File in separate Generic Stream Partitions (see Section 6 above). The Timed Text Track File shall not contain resources not referenced by the Timed Text Resource.

7.2 Index Table

The Timed Text Track File shall contain in the Footer Partition a standard MXF Index Table as defined in SMPTE 377M and as constrained by SMPTE 429-3. The Index Table shall contain one entry, pointing to the beginning of the single Data Element in the GC that holds the clip-wrapped Timed Text Resource. Because the Index Table will have only one entry, the table shall consist of a single IndexTableSegment. Within the segment, the DeltaEntryArray shall be empty and the value of EditUnitByteCount shall be 0 (zero).

7.3 Timed Text Resource Constraints

While this specification does not define or reference a specific standard for the format of the Timed Text Resource, the following requirements must be met by the resource format for the resource to be used in a Timed Text Track File:

1. The resource shall be encoded as an XML document.
2. The resource shall be identifiable using an embedded UUID value [RFC 4122].
3. The resource shall refer to external resources using UUID values.

For example, SMPTE 428-7 meets these requirements.

7.4 Header Metadata Construction

The Timed Text Resource shall be described by a top-level File Package [SMPTE 377M]. The File Package shall contain one Data Essence Track (SMPTE 337M, B.14 and B.15) with a single Data Source Clip (SMPTE 377M, B.16). A single Material Package shall be present which shall contain one Data Essence Track with a single Data Source Clip referencing the File Package.

If the Timed Text Track file contains encrypted essence (see Section 8 below), the header shall contain a Cryptographic Framework [SMPTE 429-6].

7.5 Essence Descriptors

The primary File Package in the header metadata shall have a strong reference to a TimedText Descriptor, which shall describe the Timed Text Resource (see Annex A.2).

If the Timed Text Resource references one or more Ancillary Resources, the TimedText Descriptor shall contain the same number of strong references to TimedTextResource Descriptors (one for each Ancillary Resource, see Annex A.3). A TimedTextResource Descriptor contains the ResourceID and MIME Type [RFC 2045] of the respective resource, and also the BodySID of the Generic Stream Partition containing the resource data. Figure 4 illustrates the metadata descriptors for a Timed Text Track File containing a Timed Text Resource and two Ancillary Resources (a font and an image).

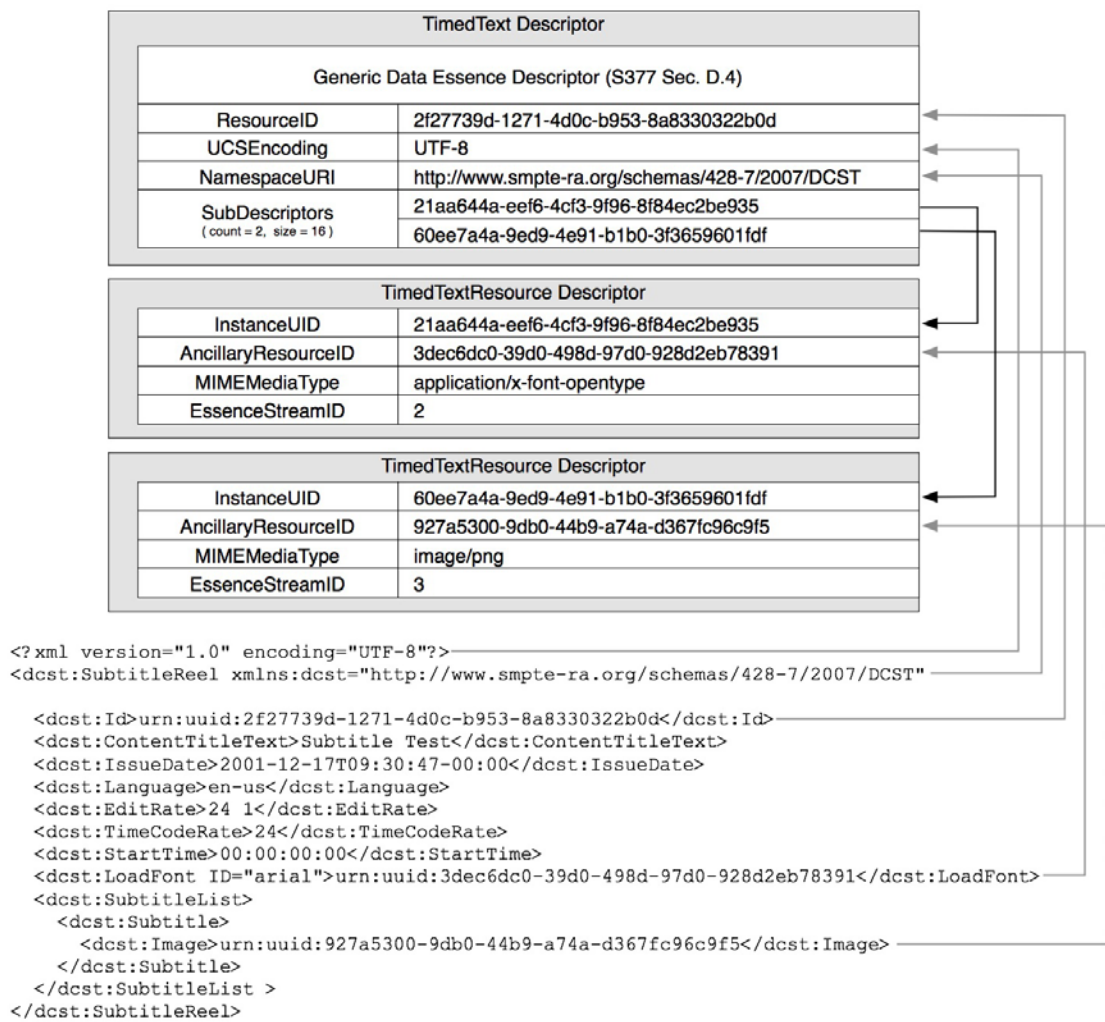


Figure 4 – Essence Descriptor example

7.6 Header Metadata Constraints

7.6.1 General

The MXF Header Metadata of Timed Text Track Files shall conform to the section “Header Metadata Constraints” in SMPTE 429-3.

8 Essence Encryption

Timed Text Track Files may be encrypted. For this purpose, the Timed Text Resource shall be contained in an Encrypted Triplet per SMPTE 429-6. Ancillary Resources may also be encrypted. Because an Ancillary Resource is a component of the file's essence track, it shall be encrypted using the same Cryptographic Context used to encrypt the Timed Text Resource. Ancillary Resources shall not be encrypted unless the Timed Text Resource is also encrypted.

When encrypting Ancillary Resources, the Generic Stream Data Element KLV packet in the Generic Stream Partition that contains the resource data shall be contained in an Encrypted Triplet per SMPTE 429-6. If the optional MIC value is present in the EKLK packet, the Sequence Number value shall increment with each successive encrypted Ancillary Resource (i.e., the first Ancillary Resource EKLK packet shall have a Sequence Number that is one greater than that of the Timed Text Resource EKLK packet, the second Ancillary Resource EKLK packet shall have a Sequence Number that is two greater than that of the Timed Text Resource EKLK packet, etc).

Figure 5 is a modified form of the Cryptographic Framework diagram from SMPTE 429-6. This diagram shows the relationship of the Framework instance to the EKLK packets in the Generic Stream Partitions. As specified in SMPTE 429-6, the EKLK packet contains a weak reference to the Cryptographic Framework. Decoders can trace from the Cryptographic Framework to the Generic Stream Partitions by using the set of TimedTextResource sub-descriptors in the File Package that references the Framework.

A Timed Text Track File that contains Encrypted Triplets shall have a Cryptographic Framework and single Cryptographic Context (i.e., all Encrypted Triplet packets in a Timed Text Track File shall be encrypted using the same symmetric key).

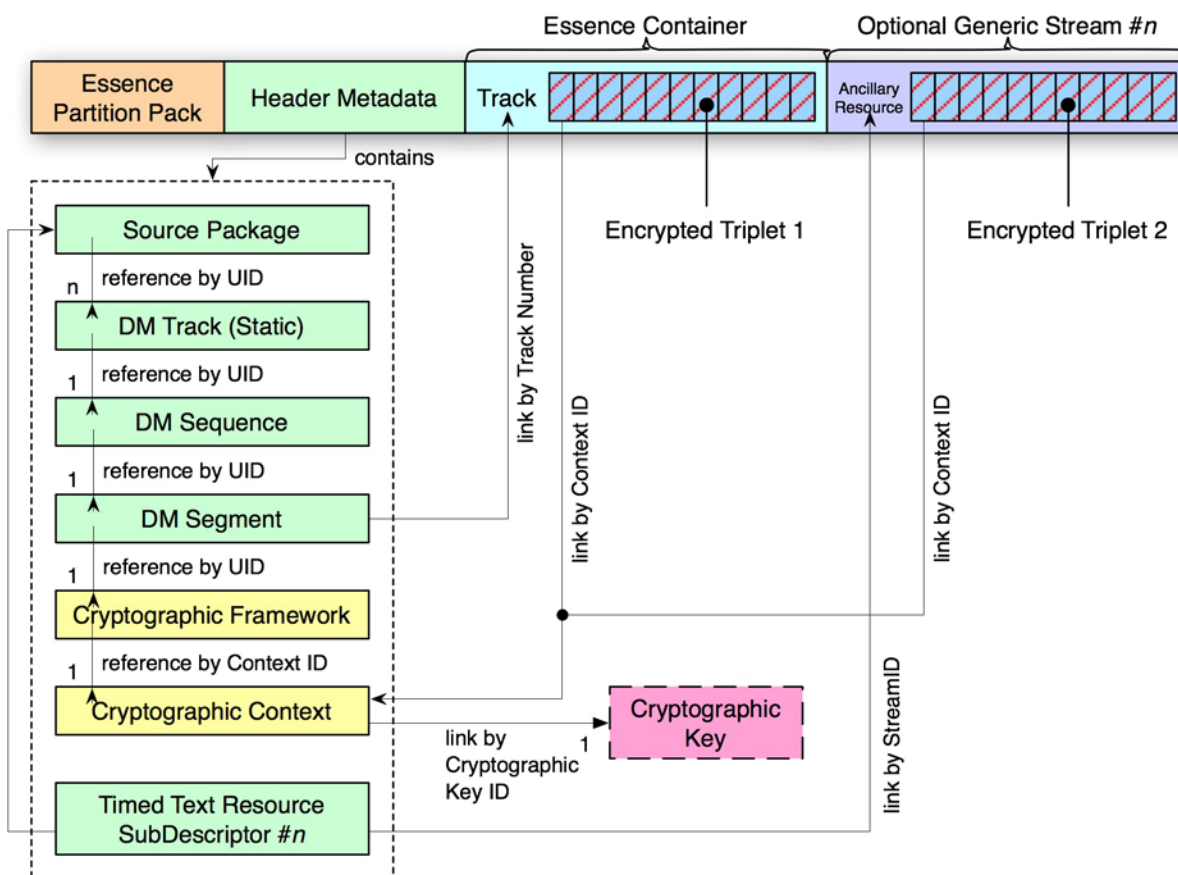


Figure 5 – Cryptographic Framework

9 Synchronization

Synchronization information is contained in the Timed Text Resource. The normative definition of the Timed Text Resource shall specify synchronization with other essence. Timed Text Track Files shall not contain synchronization information other than what is required to create valid MXF header metadata. When present in a Timed Text Track File, MXF header synchronization metadata shall be ignored by the decoder.

10 Random Access to Ancillary Resources (Informative)

During reproduction of the essence encoded in the Timed Text Resource, the decoder will have to retrieve from the Track File any Ancillary Resources referenced by the Timed Text Resource. This section provides an informative method of efficiently performing this retrieval.

It should be noted that each Generic Stream Partition has a distinct BodySID value. This value is given in the respective TimedTextAncillaryResource sub-descriptor that describes the Ancillary Resource contained in a Generic Stream Partition, and also in the Random Index Pack (RIP) required by SMPTE 429-3. Given this information, and a UUID value identifying an Ancillary Resource, the following algorithm can be used to seek to the location of the Ancillary Resource in the Track File:

1. Search for the UUID value in the AncillaryResourceID property in the set of TimedTextAncillaryResource sub-descriptors referenced by the TimedTextResource descriptor. If no match is found, the resource does not exist in the file.
2. Using the value of the EssenceStreamID property of the matching sub-descriptor, locate in the Random Index Pack (RIP) the Pair entry having a matching BodySID value.
3. The ByteOffset property in the Pair located in step 2 above gives the location of the Partition Pack at the start of the Generic Stream Partition. The Ancillary Resource data will be contained in the KLV packet immediately following the Partition Pack.

Annex A (Normative)

Labels and Descriptor Sets

With the exception of InstanceID and GenerationUID, which are already defined in SMPTE 377M, all Local Tag values for the descriptors shall be dynamically allocated as defined in SMPTE 377M. The translation from each dynamically allocated local tag value to its full UL value can be found using the Primer Pack mechanism defined in SMPTE 377M.

A.1 Key UL Values

Table A.1 – Specification of the Timed Text Essence Container Label

Byte No.	Description	Value (hex)	Meaning
1-12	Defined by Generic Container		See SMPTE 379M Table 2
13	Essence Container Kind	02h	MXF Generic Container
14	Mapping Kind	13h	Timed Text
15	reserved	01h	
16	reserved	01h	

Table A.2 – Key Value for a Timed Text Essence Element

Byte No.	Description	Value (hex)	Meaning
1-12	See SMPTE 379M Table 3		
13	Item Type Identifier	17h	Timed text Item
14	Essence Element Count	01h	Count of XML Resource Elements in the file (always 1)
15	Essence Element Type	0Bh	Clip Wrapped Element
16	Essence Element Number	01h	

A.2 TimedText Descriptor Set

Table A.3 – TimedTextDescriptor

Item Name	Type	Len	UL Designator	Req ?	Meaning	Default
TimedText Descriptor	Set UL	16	See table below	Req	Defines the TimedText Descriptor Set	
Length	BER Length	var		Req	Set length (See SMPTE 377M, "Header metadata coding")	
All items from the MXF Generic Data Essence Descriptor						
ResourceID	UUID	16	06.0E.2B.34 01.01.01.0C 01.01.15.12 00.00.00.00	Req	A UUID value that identifies this Resource	
UCSEncoding	UTF16 String	var	06.0E.2B.34 01.01.01.0C 04.09.05.00 00.00.00.00	Req	A text string giving the ISO/IEC 10646-1 encoding of the essence data	"UTF-8"
NamespaceURI	UTF16 String	var	06.0E.2B.34 01.01.01.08 01.02.01.05 01.00.00.00	Req	A URI value giving the XML namespace name of the top-level XML element in the essence data	

Table A.4 – Key for TimedTextDescriptor

Byte No.	Description	Value (hex)	Meaning
1-13	Defined in the Structural Header Metadata Implementation section of SMPTE377M (File Format Specification)		
14	Set Kind (1)	01h	TimedTextDescriptor
15	Set Kind (2)	64h	
16	Reserved	00h	Reserved

A.3 TimedText Resource SubDescriptor Set

The TimedTextResource Sub-Descriptor is a supplementary Essence Descriptor that can be strongly referenced by the TimedText Descriptor. In order that the strong reference can be made, the MXF Generic Descriptor (as defined in SMPTE 377M) has an additional optional property as defined in Table A.5.

The Local Tag value associated with this additional optional property (called "Sub Descriptors") shall be dynamically allocated (dynamic) as defined in SMPTE 377M.

Table A.5 – Additional Optional Property for the MXF Generic Descriptor

Item Name	Type	Len	Local Tag	Item Designator	Req ?	Meaning	Default
All elements from the Generic Descriptor defined in SMPTE 377M							
Sub Descriptors	Array of StrongRef (Sub Descriptors)	8+ 16n	dynamic	06.01.01.04.06.10.00.00	Opt	Array of strong references to sub Descriptor sets	

Note: The method of adding a Sub-Descriptor was first used in SMPTE 422M.

Table A.6 – TimeTextResourceSubDescriptor

Item Name	Type	Len	UL Designator	Req ?	Meaning	Default
TimedTextResourceSubDescriptor	Set UL	16	See table below	Req	Defines the TimedTextResourceSubDescriptor Set	
Length	BER Length	Var		Req	Set length (See SMPTE 377M, "Header metadata coding")	
Instance UID	UUID	16	01.01.15.02.00.00.00.00	Req	Unique ID of this instance [The ISO/IEC 11578 (Annex A) 16 byte Globally Unique Identifier]	
Generation UID	UUID	16	05.20.07.01.08.00.00.00	Opt	Generation Identifier [RP210 Specifies the reference to an overall modification]	
AncillaryResourceID	UUID	16	06.0E.2B.34 01.01.01.0C 01.01.15.13 00.00.00.00	Req	A UUID value that identifies this Ancillary Resource (copied from the set of resource ids in the Timed Text Resource)	
MIMEMediaType	UTF16 String	var	06.0E.2B.34 01.01.01.07 04.09.02.01 00.00.00.00	Req	A MIME Type identifier (see Content-Type Header Field in [RFC 2045]) which identifies the resource data type	
EssenceStreamID	UINT32	4	06.0E.2B.34 01.01.01.04 01.03.04.04 00.00.00.00	Req	The BodySID of the partition that contains the resource data	

Table A.7 – Key for TimedTextResourceSubDescriptor

Byte No.	Description	Value (hex)	Meaning
1-13	Defined in the Structural Header Metadata Implementation section of SMPTE 377M (File Format Specification)		
14	Set Kind (1)	01h	TimedTextResourceSubDescriptor
15	Set Kind (2)	65h	
16	Reserved	00h	Reserved

Annex B (Informative)
Bibliography

[SMPTE 336M-2007] Data Encoding Protocol using Key-Length-Value

[SMPTE 422M-2006] Material Exchange Format — Mapping JPEG 2000 Codestreams into the MXF Generic Container

[SMPTE 428-7-2007] D-Cinema Distribution Master — Subtitle

[SMPTE 429-7-2006] D-Cinema Packaging – Composition Playlist

[SMPTE RP 210] Metadata Dictionary Registry of Metadata Element Descriptions

[SMPTE RP 224] SMPTE Labels Register

[SMPTE EG 41-2004] Material Exchange Format (MXF) — Engineering Guideline

[SMPTE EG 42-2004] Material Exchange Format (MXF) — MXF Descriptive Metadata

World Wide Web Consortium (W3C) (2004 February 4). Extensible Markup Language (XML) 1.0 (Third Edition)

World Wide Web Consortium (W3C) (2004 October 28). XML Schema Part 1: Structures (Second Edition)

World Wide Web Consortium (W3C) (2004 October 28). XML Schema Part 2: Datatypes (Second Edition)

Internet Engineering Task Force (IETF) [RFC2046] (November 1996) Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types