

SMPTE STANDARD



Format for Non-PCM Audio and Data in AES3 — Audio Metadata

Page 1 of 21

Table of Contents	Page
1 Scope	4
2 Conformance Notation	4
3 Normative References	5
4 Terms and Definitions	5
4.1 Section	5
4.2 Payload	5
4.3 Utility Payload	5
4.4 Audio Metadata Payload	5
4.5 Audio Metadata Pack	5
4.6 Static payload tag	5
4.7 Dynamic payload tag	6
4.8 ASN.1 OID BER SubId	6
4.9 BER	6
4.10 BER short form	6
4.11 BER long form	6
4.12 UL	6
5 Notations	6
6 Audio Metadata	7
7 Encoding of Audio Metadata Pack	9
7.1 Introduction	9
7.2 Audio Metadata Pack Key	9
7.3 Audio Metadata Pack Length	9
7.4 Audio Metadata Pack Value	9
7.5 Audio Metadata Pack Payload Tag Assignment	12
8 Utility Payloads	12
8.1 Introduction	12
8.2 Config Payload	12
8.3 Sync Payload	14

8.4	CRC Payload	16
9	Audio Metadata Payloads	17
9.1	Audio Metadata Payload Overview	17
9.2	Audio Metadata Payload Tag	17
9.3	Audio Metadata Payload Length	17
9.4	Audio Metadata Payload Value	17
10	Payload Static Tag Assignments	17
10.1	ST 2109 Administrative Register	17
10.2	Private Payloads	18
11	Mapping of Audio Metadata Packs to AES3	18
11.1	Introduction	18
11.2	burst_preamble	18
11.3	Reference Point	19
11.4	Reference Position	19
11.5	Repetition Rate	19
	Bibliography (informative)	21

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 32NF.

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.

This standard describes an open and extensible framework for transport of Audio Metadata over an AES3 interface. This framework is based on the KLV data encoding protocol (SMPTE ST 336) and the SMPTE ST 337 format for Non-PCM Audio and Data in AES3.

Audio Metadata is defined as metadata that is associated with audio essence which is present in other AES3 channels, either as one or more PCM audio channels, or as coded audio essence (in ST 337 format). This audio metadata might require time synchronization with the audio essence, in some cases down to a resolution of an audio sample. This standard includes signaling that supports this synchronization.

1 Scope

This standard specifies format requirements for real-time Audio Metadata carried within an AES3 interface according to SMPTE ST 337. Audio Metadata is grouped into *payloads* for transport. This standard covers wrapping and transport of the Audio Metadata payloads only. The contents of audio metadata payloads are defined elsewhere and are beyond the scope of this standard, however Utility payloads that contain information that supports carriage of the Audio Metadata payloads are within scope and are defined in this standard. This standard specifies KLV encoding for wrapping and transport of Audio Metadata payloads, but places no requirement on the format of the metadata carried within Audio Metadata payloads.

This standard applies to Audio Metadata payloads only. Transport of payloads other than Audio Metadata payloads and associated Utility payloads is beyond the scope of this standard.

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; then 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 engineering document. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this engineering document are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

SMPTE ST 337:2015, Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface

SMPTE ST 355:2001, Format for Non-PCM Audio and Data in AES3 - KLV Data Type

SMPTE ST 336:2017, Data Encoding Protocol Using Key Length Value

ISO/IEC 8825-1:2008 Information Technology — ASN.1 Encoding Rules: Specification of Basic Encoding Rules (BER), Canonical Encoding Rules (CER) and Distinguished Encoding Rules (DER)

ETSI EN 300 468 v1.15.1 - Digital Video Broadcasting (DVB); Specification for Service Information (SI) in DVB systems

4 Terms and Definitions

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

4.1 Section

group of payloads within the Audio Metadata Pack

4.2 Payload

item consisting of a payload tag, payload length, and payload value

4.3 Utility Payload

payload whose value field contains information other than audio metadata

4.4 Audio Metadata Payload

payload whose value field contains audio metadata.

4.5 Audio Metadata Pack

ST 336 Defined-Length Pack, as defined in this document, with a single element, ST 2109 Audio Metadata, consisting of Utility Payloads and/or Audio Metadata Payloads

4.6 Static payload tag

ASN.1 OID BER SubId whose UL mapping is permanently assigned via the ST 2109 Administrative Register and is valid for all Audio Metadata Packs

4.7 Dynamic payload tag

ASN.1 OID BER SubId whose UL mapping is assigned within an Audio Metadata Pack, and whose assignment scope is limited to that pack

4.8 ASN.1 OID BER SubId

BER encoding of subidentifiers as defined in clause 8.1.2.4.2 of ISO/IEC 8825-1. The first 127 nonzero values are encoded in a single byte, values 128 to 16383 are encoded in 2 bytes, greater values are encoded in 3 or more bytes as necessary.

4.9 BER

Basic Encoding Rules, defined in ISO/IEC 8825-1

4.10 BER short form

BER encoding as defined in clause 8.1.3.4 of ISO/IEC 8825-1

4.11 BER long form

BER encoding as defined in clause 8.1.3.5 of ISO/IEC 8825-1

4.12 UL

SMPTE administered Universal Label as defined in SMPTE ST 298

5 Notations

Table columns shall have the meanings as defined in Table 1.

Table 1 - Definition of Table Headings

Table heading	Definition
Item Name	Name of this pack or metadata item as indicated in the defining document
Type	The data type of the entry
Len	Length of the value field for this entry, in bytes, according to the data type
Payload Tag	Value of the payload tag that identifies this entry
Req	Req indicates that the item is always present. This indicator maps to the isOptional entry in the SMPTE Groups Register, where isOptional corresponds to table entry != "Req". Opt indicates that the item can be present or omitted.

Table heading	Definition
	Opt* indicates that this item is present or omitted depending on the presence or value of other items.
Definition	Description of this entry
Symbol	The registry symbol that is associated with this item
Kind	The kind of the registry entry
Item UL	The SMPTE Universal Label that shall be associated with this item

6 Audio Metadata

In this standard, Audio Metadata is segmented into *payloads* that are intended to combine related audio metadata together (e.g. audio loudness metadata, object position metadata, etc.). Multiple payloads, grouped into sections, can be combined together into a single *Audio Metadata Pack*. For AES3 transport Audio Metadata Packs are wrapped within SMPTE ST 337 data bursts with no more than one Audio Metadata Pack per data burst. In addition to audio metadata payloads, packs can include one or more utility payloads that contain metadata that support the transport format. These payloads can include information that supports synchronization with associated audio essence, or identification of associated audio essence, for example.

Audio Metadata Packs are encoded in KLV format per SMPTE ST 336. Payloads are not required to be KLV encoded. The pack encoding is defined as a KLV Defined-Length Pack as specified in this standard and identified by a SMPTE Universal Label (Audio Metadata Pack). The Audio Metadata Pack consists of a single element, ST 2109 Audio Metadata, which is defined as a sequence of payloads. Each type of payload within an Audio Metadata Pack is identified by a unique SMPTE Universal Label (UL). The ST 2109 Administrative Register defines the mapping of payload ULs to static payload tags. Dynamic mapping of payload ULs to payload tags is also supported for payload types not defined in the ST 2109 Administrative Register. Consistent with the SMPTE metadata dictionary, Audio Metadata Packs can contain payloads with keys defined as private use ULs. Data within payloads identified as private use does not have to be publicly disclosed.