PROPOSED SMPTE STANDARD

for Television —

Metadata Dictionary Structure

1 Scope

The metadata dictionary structure defined in this standard covers the use of metadata for all types of essence (video, audio, and data in their various forms). Applications of individual dictionary entries will vary but, when used, metadata shall conform to the definitions and formats in this metadata dictionary structure standard and the associated metadata dictionary recommended practice (SMPTE RP 210). SMPTE RP 210 defines a registered set of metadata element descriptions for association with essence or other metadata and this standard and the contents practice shall be used together as a pair — neither shall be used in isolation.

2 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.

ANSI/SMPTE 298M-1997, Television — Universal Labels for Unique Identification of Digital Data

SMPTE 336M-2001, Television — Data Encoding Protocol using Key-Length-Value

SMPTE 359M-2001, Television and Motion Pictures — Dynamic Documents

SMPTE RP 210.1-2001, Metadata Dictionary

3 Metadata dictionary structure

The metadata dictionary structure provides flexibility in capturing metadata and exchanging it among applications through a standardized hierarchy of universal labels for the metadata elements, grouped to aid their management within a small but comprehensive number of classes. Metadata classes are collections of metadata elements with common characteristics or attributes. Additional classes are provided for user-defined metadata.

SMPTE RP 210 references an individual item or element of metadata using a two-part 16-byte universal label that is numerical (and hence language independent) and unique. The first eight bytes label the second eight bytes as a tag in a specific version of a designated metadata dictionary (tags are defined in SMPTE 336M). This tag is used to index the meaning or definition of the metadata element.

The actual metadata information described by the metadata element is the metadata value. The dictionary also contains information on the required format of metadata values and the allowable range of values (if applicable) either as a list or as a bounded range.

Individual data element values can frequently be represented in more than one way — for instance, it is possible to represent a textual value as ASCII or Unicode, where the value is identical but the particular representation different. It is important both that the representation is known and that as new representations are registered they can be accommodated. In this case, the last active word of the tag defines the representation in use — the default being 00_h.

The metadata dictionary is organized into nodes and leaves. The dictionary classes just described form class nodes and below these are further nodes at each subclass. To aid the management of the dictionary, these nodes and subnodes are assigned tags to which no value is assigned, so as to give clear breaks in the structure. Entries within a subclass form leaves, which are the data elements themselves.

Lower levels of the dictionary structure can be derived from the tag structure in SMPTE RP 210 and are detailed in SMPTE EG 37.

3.1 Compatibility with other metadata structures

The metadata dictionary structure is a framework that supports global interoperability by defining metadata tags in a way that enables the interchange of SMPTE metadata with metadata from different sources and originated by other bodies.

Many different cataloging conventions are used by communities who focus on a specific domain or subject or who have specific needs for archive and retrieval of multimedia data including, for example, intellectual rights. The metadata dictionary is not intended to replace conventions already in use, for example in textual naming or keywords. Within the framework of the metadata dictionary structure, different content creation communities, media indexing professionals, or metadata extractors and users can develop metadata conventions that meet their specific requirements.

3.2 Individual metadata classes

Within the metadata dictionary, metadata elements are organized into a hierarchical structure, where each is assigned to a metadata class as shown in the overview of figure 1. The initial set of metadata classes in this standard consists of:

Class 1: Identification and location
Class 2: Administration
Class 3: Interpretive
Class 4: Parametric
Class 5: Process
Class 6: Relational
Class 7: Spatio-temporal
Class 13: Organizationally registered for public use
Class 14: Organizationally registered as private
Class 15: Experimental

The number of metadata classes can be extended in the future to a maximum of 127. Although dictionary classes can be populated with any metadata (such as that associated with still images, audio, graphics, etc.), additional new classes may be created up to that limit to deal with specific metadata characteristics or attributes.

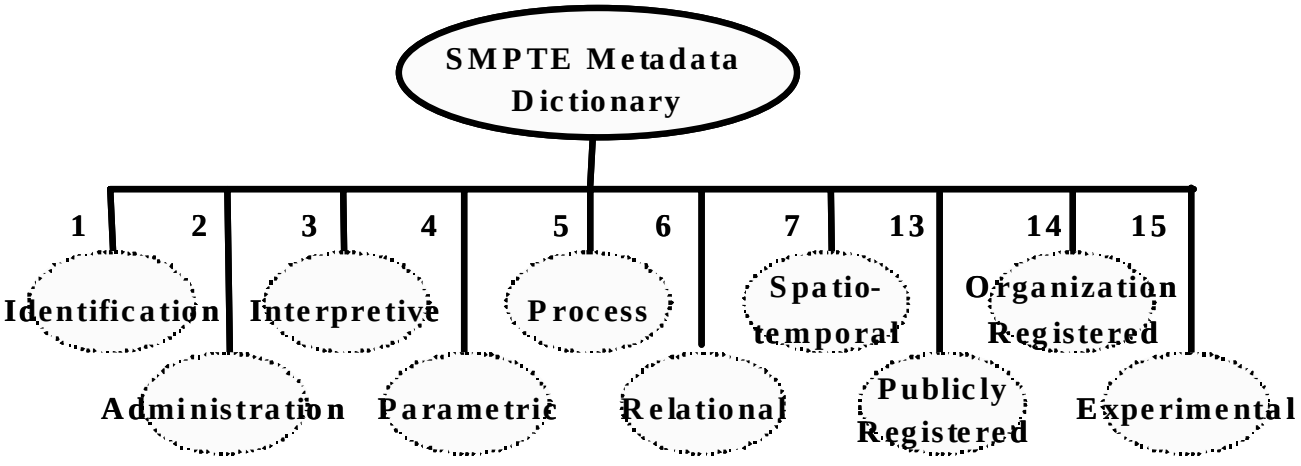

Figure 1 – Metadata class structure

3.2.1 Class 1: Identification and location

Metadata in this class shall consist of identifying information (IDs) that describes the essence of the overall bit stream or file. A critical part of class 1 metadata is unambiguous identification of the essence using a single, recognized number or label such as the SMPTE unique material identifier (SMPTE 330M). Information in this class shall include global and local identifiers as well as identifying information about the metadata elements themselves (so-called meta-metadata). Examples of subclass titles in this class are:

– Globally unique identifiers
– ISO identifiers
– Object identifiers
– Device identifiers
– Unique IPR identifiers
– Local locators
– Titles

3.2.2 Class 2: Administration

Metadata in this class shall consist of administrative or business data that describe information about the essence or metadata that are relevant to its application. Information on authorized use and restrictions on use, and encryption are in this metadata class. Cost information and information needed to protect intellectual property or to protect ownership shall also be contained in class 2. Examples of subclass titles in this class are:

– Supplier
– Rights
– Financial information
– Security
– Publication outlet
– Participating parties
– Broadcast and repeat statistics

3.2.3 Class 3: Interpretive

Metadata in this class shall consist of descriptive information which is normally considered either a subjective or a human-generated description of the essence or a computational result from machine examination of the essence. Interpretive information shall consist of, but not be limited to, textual terms (e.g., keywords, narrative summary, titles, genre categories, scripts, etc.), or computational metrics (e.g., color histograms, texture maps, object shapes, facial features, etc.). Information in class 3 shall be principally used for indexing, cataloging, administering, searching, and retrieving the content of essence. Examples of subclass titles in this class are:

– Fundamental (such as ISO language code, length and time systems)
– Descriptive (human-assigned)
– Categorization
– Assessments
– Descriptors (machine-assigned or computed)

3.2.4 Class 4: Parametric

Metadata in this class shall consist of information that describes the technical characteristics of the camera, sensor, or system that originates the essence or metadata. Information about the technical characteristics of the essence or metadata is also provided, including but not limited to its creation parameters and the configuration of the originating system. Examples of subclass titles in this class are:

– Video essence encoding characteristics
– Audio essence encoding characteristics
– Data essence encoding characteristics
– Metadata encoding characteristics
– Audio test parameters
– Film pulldown characteristics
– Fundamental sequencing and scanning
– MPEG coding characteristics
– Time code characteristics

3.2.5 Class 5: Process

Metadata in this class shall consist of information that describes how the essence was processed or otherwise changed or enhanced after its origination. This class shall include, but not be limited to, many of the parameters in an edit decision list. Additional information in class 5 shall be an audit trail (heritage) of all changes to the original content over time. Also included shall be a record of compression/decompression steps and any changes in storage media or format. Examples of subclass titles in this class are:

– Process indicators
– Manipulation
– Downstream processing history
– Enhancement or modification
– Audio processor settings (device-specific)
– Editing information

3.2.6 Class 6: Relational

Metadata in this class shall consist of information that describes relationships among objects in the content or among any combination of essence, objects, and metadata. Examples of subclass titles in this class are:

- Generic relationships
- Relatives
- Essence-to-essence relationship
- Metadata-to-essence relationship
- Metadata-to-metadata relationship
- Object-to-object relationship
- Metadata-to-object relationship
- Related production material
- Numerical sequence
- Relationship structures

3.2.7 Class 7: Spatio-temporal

Metadata in this class shall consist of information about aspects of the content or the originating camera, sensor, or system relating to time, place, or space. Geospatial information in class 7 shall be any information that defines the positions or places (either absolute or relative) of objects, scenes, individuals, or any other component of the essence. Temporal elements such as dates, time codes, synchronization marks, temporal keywords, and motion (vector) parameters shall also be part of class 7 metadata. Examples of subclass titles in this class are:

- Position and space vectors
- Absolute position
- Image positional information
- Rate and direction of positional change
- Abstract locations
- Angular specifications
- Distance measurements
- Operational date and time information
- Absolute date and time
- Relative durations
- Rights date and time
- Setting date and time (characterized time period)
- Delay
- Latency

3.2.8 Class 13: Organizationally registered for public use

Metadata in this class shall consist of individual elements of metadata that have been registered by a specific user organization or individual and are therefore reserved and managed separately from the other classes (1 -7) of metadata. Information about publicly registered metadata shall appear in the appropriate clauses of the published metadata dictionary. Class 13 metadata shall be managed by the SMPTE Registration Authority and its approval shall be consistent with SMPTE Administrative Practices.

3.2.9 Class 14: Organizationally registered as private

Metadata in this class shall consist of individual elements or groups of elements of metadata that have been registered by a specific user organization or individual and are therefore reserved and managed separately from the other classes (1 -7) of metadata. Information about organizationally registered metadata shall not be made public but the metadata tags shall be publicly identified in the metadata dictionary and shall be reserved for use by the registered organization. Class 14 metadata shall be managed by the SMPTE Registration Authority and its approval shall be consistent with SMPTE Administrative Practices.

3.2.10 Class 15: Experimental

Metadata in this class shall consist of information whose definition and use does not need to conform to the definitions in the metadata dictionary. Class 15 metadata is intended for use in multimedia research or other limited access, experimental environments where experimentation with new metadata elements and applications does not depend on strict conformance to approved standards and which remain within a test or laboratory environment.

3.3 Dictionary element structure and format

The metadata dictionary shall consist of the following fields for each metadata element:

3.3.1 Data element tag

This has eight fields representing the eight octets or bytes of the data element tag. These uniquely identify the specific metadata element in the dictionary in an hierarchical fashion. Classes are designated with the first bye in the data element tag and subsequent bytes enable the hierarchical identification of subclasses and/or individual leaves.

3.3.2 Data element name

This entry is the English language name for the element represented numerically by the data element tag.

3.3.3 Dictionary version at introduction

This entry records the version number of the dictionary standard (i.e., the structure standard/recommended practice pair) which first recorded the allocation of a data element against a data element tag.

3.3.4 Data element definition

This entry is the detailed and unambiguous English language definition of what is represented by the data element tag and element name.

3.3.5 Type

This entry identifies the representation category of metadata value allowed for this element in order to permit correct interpretation of that value; for example, in the case of a text string or an SMPTE time code. A compound type entry is a class or subclass node or split in the metadata dictionary hierarchy and cannot have a value associated with it. A document is under development which defines the types for each element in more detail.

3.3.6 Value length

This entry defines the permitted length in bytes or characters of the value of the data element. In some cases, such as a text string, the length is not defined or limited and the value length is described as variable. However, in practice, a variable length may be limited by the application specification.

3.3.7 Value range

This entry defines any limitations on the permitted values of a data element.

3.3.8 Node/leaf

This entry defines whether the dictionary entry is a node in the classification structure or a leaf as defined in 3.1.

3.3.9 Defining document

In cases where the data element type or other parameter is defined in another document, this entry references that standard or the authoritative source of the information.

3.4 Metadata dictionary structure maintenance

The principles for maintenance and administration of the metadata dictionary structure are defined in the following clauses:

3.4.1 Dictionary version information

The following information shall be provided by the SMPTE Registration Authority with each update to the metadata dictionary structure and contents:

Standard name: Metadata dictionary structure standard

Standard designator: One-byte unsigned integer in the range of 1 to 127

Version number: One-byte unsigned integer in the range of 1 to 127

Effective date: TBD

Database format: Text

Summary of changes: Text

Database administrator: Text with URL or e-mail

Contact information: Text with URL or e-mail

3.4.2 Dictionary management and compatibility requirements

To ensure reliable and correct interpretation of legacy material in the future, changes to SMPTE RP 210 shall be additions only. Deletions or changes to entries (other than purely editorial) are not permitted. This addition process shall be speedily carried out and documented in accordance with SMPTE 359M by the SMPTE Registration Authority. It shall occur immediately on request from the appropriate SMPTE technology committee and shall be administered and scrutinized with a light but formal touch to ensure minimal delay in the availability for use of a newly required registered tag. The version number of the dictionary shall be incremented each and every time an addition (or group of additions) is made since this is critical to ensuring the operational compatibility of metadata decoders. The incrementing of the version number shall not prevent use of unaffected tags,

structure, or contents by a decoder that conforms to the prior version.

It is inevitable, given the above addition process, that eventually the dictionary will become cluttered with legacy entries to the point where the responsible SMPTE technology committee determines it has reached the limit of its usefulness. At this stage, or when other changes to the dictionary contents, to an existing approved dictionary structure, or to relationships among metadata classes that prevent backward compatibility are necessary, a new dictionary consisting of both structure standard and contents recommended practice shall be created and the structure and contents of the new dictionary made readily accessible on-line by the SMPTE Registration Authority to allow upgrades to decoders. The superseded standard shall then undergo no further revision unless essential under the SMPTE five-year rule.

3.4.3 Dictionary availability

The dictionary documents shall be available in electronic form in a defined electronic publishing format, such as XML with an accompanying document type definition.

The latest version of SMPTE RP 210 shall be made available on the SMPTE website in a noneditable format (pdf or equivalent). It is preferred that a minimum of the two immediate previous versions be also available in a clearly indicated archive.

Annex A (normative)
Glossary of terms

A.1 content: The program content will consist of the sum total of the essence (video, audio, data, etc.) and the metadata.

A.2 data element: An individual unit of data or metadata.

A.3 data element definition: The detailed and unambiguous definition of what is represented by the key and element name.

A.4 data element name: The English language name for the element represented numerically by the data element tag.

A.5 data element tag: The 8-byte tag that uniquely identifies the data element in the metadata dictionary.

A.6 essence: Identified by the SMPTE/EBU Task Force for Harmonized Standards for the Exchange of Program Material as Bitstreams (TFHS) as the digital representation of video, audio, and/or data information. Essence can therefore also be graphics, telemetry, photographs, or other information.

A.7 label: A 16-byte key that explicitly identifies a predefined value or group of values.

A.8 metadata: Generally referred to as data about data or data describing other data. More specifically, information that is considered ancillary to or otherwise directly complementary to the essence. Any information that a content provider considers useful or of value when associated with the essence being provided.

A.9 metadata class: The broad category of metadata that forms the first level of hierarchy for all registered metadata.

A.10 metadata dictionary: The standard database of approved, registered data element tags and their definitions.

A.11 metadata element: A broad term for a unit of metadata.

A.12 metadata tag: The 8-byte tag that uniquely identifies the data element in the metadata dictionary.

A.13 type: Information about the representation of the metadata or data value. A document is under development which defines the type for each element in more detail.

A.14 value: The instance of information described by the metadata tag. Equivalent to description in the terminology of MPEG-7.

Annex B (informative)
Organization of references

The organization of standards and guidelines for metadata are illustrated in figure B.1. No single standard can contain all of the information needed to describe and encode metadata. Hence, a layered approach is used to convey the information so the user can select the applicable standard(s) for the level of implementation needed. This standard, the metadata dictionary recommended practice (SMPTE RP 210), the standard for key-length-value (KLV) data encoding (SMPTE 336M), and the standard for the registration process (SMPTE 359M) form the SMPTE normative standards for metadata. Informative SMPTE documents supplement the standards for encoding with examples and administrative instructions on managing the data standardization and registration process via the SMPTE Registration Authority.

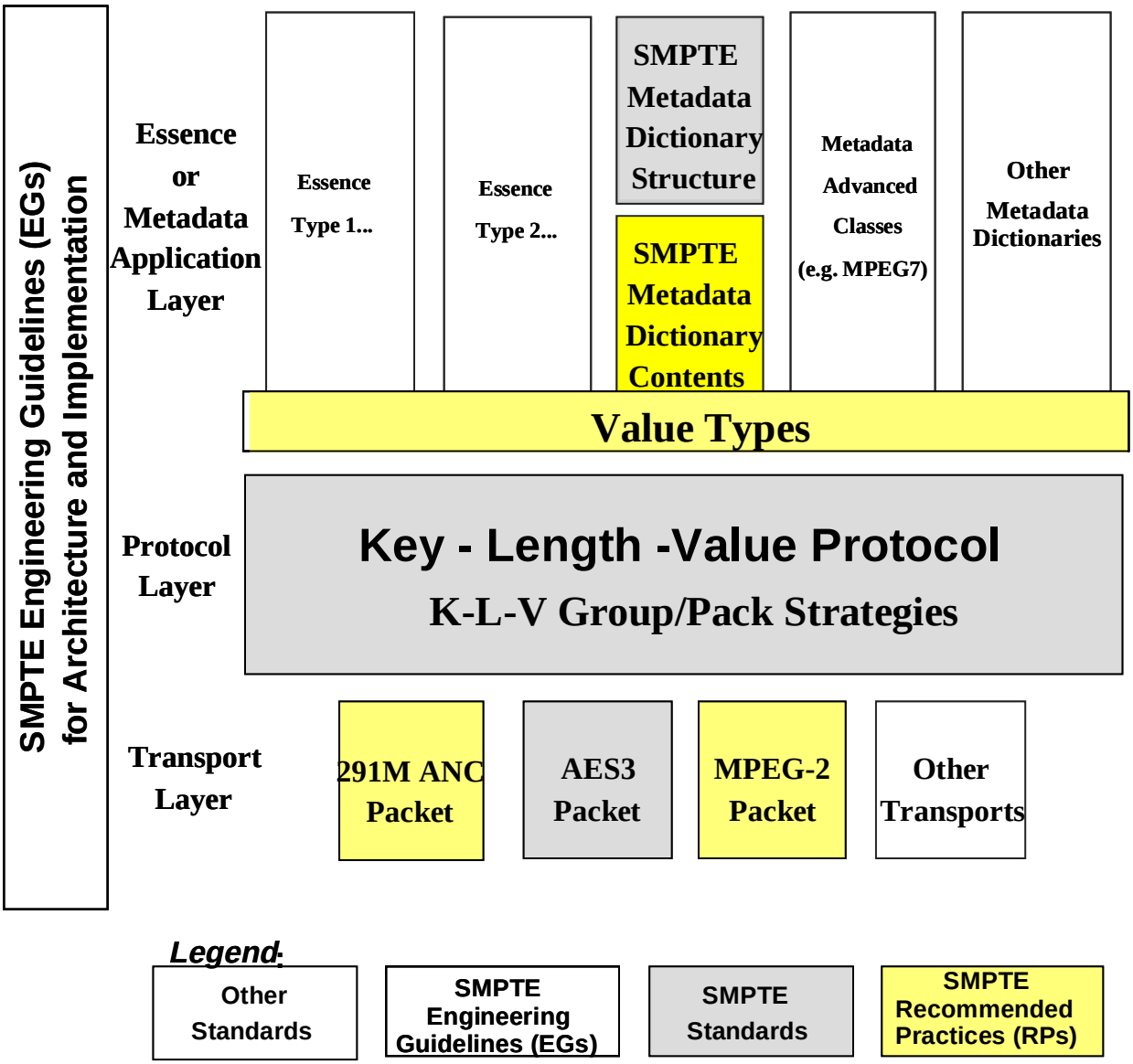

Figure B.1 – Organization of references

Annex C (informative)
Bibliography

SMPTE 330M-2000, Television — Unique Material Identifier (UMID)

SMPTE 355M-2001, Television — Format for Non-PCM Audio and Data in AES3 — KLV Data Type

SMPTE RP 217, Nonsynchronized Mapping of KLV Packets into MPEG-2 Systems Streams

SMPTE EG 37, Node Structure for the SMPTE Metadata Dictionary