EBU – TECH 3295

P_META

Metadata Library

Source: ECM

Status: Specification 2.2

Geneva
September 2011

[*] Page intentionally left blank. This document is paginated for two sided printing

Introduction

This is version 2.2 of the P_META Semantic Metadata Schema. EBU project group ECM MAG has been working since 1999 to create a standard vocabulary for information relating to programme information in the professional broadcasting industry.

In comparison to previous versions, this version of P_META has a significant number of amendments including the attributions of qualifiers names to previously unnamed structures, an adaptation to the most recent timecode recommendations from SMPTE, the optimisation of the data structure as supported by XML, the separation of the original P_META metadata set into a common toolkit (this document) and separate application specifications (e.g. for programme exchange). B

The rules for defining, editing, managing and handling Classification Schemes are now defined in a separate EBU Technical Specification (Tech 3336).

P_META 2.2 has been developed in full awareness of other metadata sets developed by various industry bodies in addition to EBUCore (Tech 3293), for example:

- SMPTE (http://www.smpte.org),
- ISO MPEG7 (http://www.chiariglione.org/mpeg/),
- Dublin Core (http://dublincore.org/),
- PBCore (http://www.pbcore.org/),

AES (http://www.aes.org) using EBUCore for descriptive metadata,TV Anytime (http://tech.ebu.ch/tvanytime, Business to Consumer).

P_META has varying degrees of overlap with these other metadata sets.

With P_META 2.2, the EBU has collated and defined a basic catalogue of data items that can be used as the building blocks for common information exchanges between content creators, distributors, archives and systems. These data items have been validated against real business transactions undertaken by EBU members.

P_META is managed and maintained by the EBU. It will be updated in response to user feedback.

How may P_META be used?

Although P-META was originally designed to support business to business content exchanges, it has also been implemented for other purposes like for exchange between production systems or as high level descriptive semantic metadata.

Media Producers are welcome to learn from P_META as a semantic and grammatical basis for interface specifications and to implement P_META in combination with specific engineering guidelines per application and use case.

P_META can simply be implemented 'as is' or extended to cover specific needs.

P_META has been designed to be as flexible as possible in implementation while retaining consistency of meaning. It is technology-independent, and can be used in applications to create XML documents, to embed metadata in file formats such as MXF, BWF or in simple Word templates.

It is not intended for use in databases (although it could be used as a starting point) but it is designed to be the "language on the wire" in implementation. Its use of controlled terms defined in classification lists favours interoperability and allows data exchanges to be translated into different human languages without corruption.

P_META Terms and Conditions of Use

P_META is freely available for all to use, but you should take note of the following:

© EBU 2011.

REDISTRIBUTION AND USE OF THIS SPECIFICATION AND ASSOCIATED RESOURCES IS PERMITTED PROVIDED THAT THE FOLLOWING CONDITIONS ARE MET:

- REDISTRIBUTIONS MUST RETAIN THE ABOVE COPYRIGHT NOTICE, THIS LIST OF CONDITIONS AND THE FOLLOWING DISCLAIMER IN THE DOCUMENTATION AND/OR OTHER MATERIALS PROVIDED WITH THE DISTRIBUTION;

- NEITHER THE NAME OF THE EBU NOR THE NAMES OF ITS CONTRIBUTOR(S) MAY BE USED TO ENDORSE OR PROMOTE PRODUCTS DERIVED FROM THIS SPECIFICATION AND ASSOCIATED RESOURCES WITHOUT SPECIFIC PRIOR WRITTEN PERMISSION.

DISCLAIMER: THIS SPECIFICATION AND ASSOCIATED RESOURCES IS PROVIDED BY THE COPYRIGHT OWNER "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS [SOFTWARE], EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Contents

P_META

Metadata Library

EBU Committee	*First Issued*	*Revised*	*Re-issued*
PMC	2006	September 2011 (v2.2)	

Keywords: Metadata, Schema, XML

1. Overview

1.1 The EBU P_META Schema

1.1.1 Scope and Goals

Metadata is essential to broadcasting.

Originally, the EBU P/Meta project, and more recently, the Metadata Expert Community (ECM) MAG group have played key roles in the development of P_META. In so doing they have been faithful to the EBU's credo of promoting the use of existing standards, taking into particular account the point of view of broadcasters.

The result is a standard metadata library and toolkit that will be used to develop applications including programme exchange between broadcasters and possibly between production processes. It offers a way to share the meaning of electronic information necessary or useful for the business-to-business exchange and processing of programme-related information and content.

1.1.2 What is P_META?

P_META represents the 'semantic layer', also known as the 'descriptive metadata layer', that is the exact definition and meaning of each element of description considered to be representative of common production practices. This definition is based on the business meaning and understanding of each concept and is expressed in terms of normal human language.

As a fundamental principle, the data analysis has been set to identify concepts or subjects which can be reused in a variety of applications, thus giving maximum precision in meaning, yet allowing maximum flexibility in the definition of the basic elements and data structures of the "P_META metadata schema".

P_META is a toolbox of common tools. The implementation of P_META is realised through application-based specifications using these common tools and also by defining contextually relevant additional tools (within each application namespace) when required.

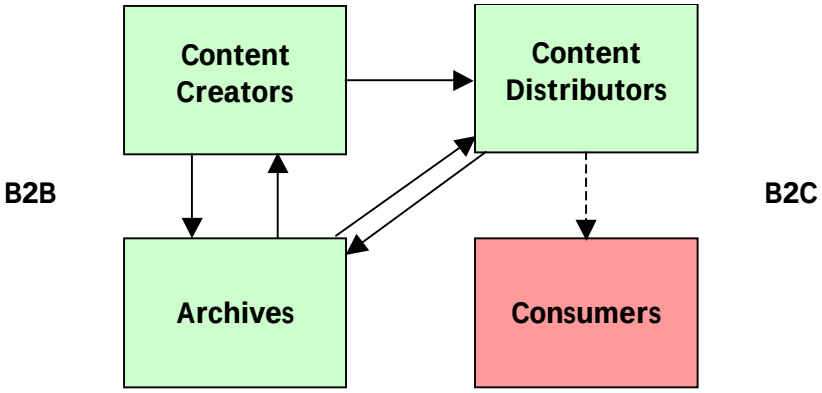
The 'technology layer' is concerned with the coding of information for storage and transmission of actual descriptions. The representation format of this version of P_META is XML (http://www.w3.org/XML/).

This English version of the specification is the reference.

1.2 The Context

P_META applies primarily to the business-to-business (B2B) environment although its interoperability with business-to-consumer (B2C) metadata is an essential feature.

As a broadcaster, if you can't find your content, you don't have it, and if users don't find your content, you simply do not exist for them!



1.2.1 Content Creators

These are involved in the Production of programmes and other content and make new material available for publication.

1.2.2 Archives

This addresses the preservation of existing material that can be re-used as originally produced, or may provide input material for new programmes. New techniques are being developed to generate metadata automatically while analysing content during the restoration and digitisation processes.

1.2.3 Content Distributors

These are involved in the publication and delivery of material to the consumers. They participate in interchanges in the business-to-consumer scenario. They are also known as Broadcasters or Content Aggregators.

1.2.4 Consumers

These are the users at the end of the media supply chain.

1.3 The Objective

P_META is a library of common terms, data types and data structures representative of what are being used in production, to editorially identify, or technically describe content and associated rights. Such a library is often referred to as a toolbox.

This toolbox is used to develop independent applications.

How may P_META be used?

1. *Broadcasters that have developed internal metadata solutions are invited to seek maximum commonalty when mapping to P_META and related applications.*

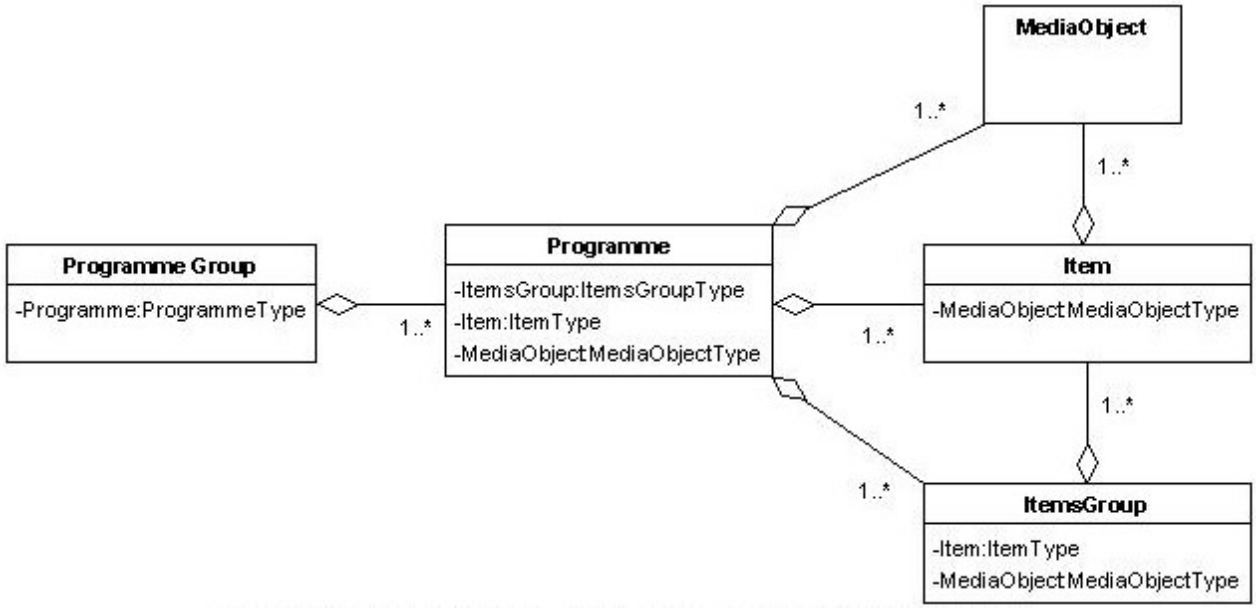
2. *Broadcasters that plan to implement metadata are invited to inspire themselves from P_META to define their requirements and to prepare for discussing metadata with manufacturers, software developers and in particular asset management system providers.*

3. *P_META can simply be implemented 'as is' or extended to cover specific needs.*

As already mentioned, finding content is vital. Search and query plays an essential role that is entirely supported by metadata and by P_META.

1.4 The P_META Model

The conceptual data model around which P_META has been built can be summarised as follows.



1.4.1 Programme Group

This is a collection of programmes created by a commissioning or scheduling decision and joined by a common editorial concept.

1.4.2 Programme

This is an audiovisual work that has been defined and created by a commissioning decision.

1.4.3 Item Group

This is a collection of items grouped as part of a programme, or otherwise editorially related.

1.4.4 Item

This is a constituent part of a programme. It is identifiable as a separate entity or as part of a programme. It is, for example, delimited in time.

1.4.5 Media Object

This is a single component of a programme or item, of a single media type. It is continuous in time. The level of granularity for the definition of a Media Object is a ‘segment’ or a ‘shot’.

1.4.6 Other entities

Other entities are contextually defined in the definition layer. Such an important entity is the brand, which is a collection of assets with a recognisable collective identity. Other important entities are the persons, organisations involved in the creation, management and control of content or events and locations related to the content.

2. Schema Documentation

2.1 Introduction

The P_META metadata toolbox is a library of common terms (a reference vocabulary) based on data types and data structures documented in the following sections.

This specification complies with XML 1.0 (http://www.w3.org/2001/XMLSchema)

2.2 Authoring conventions

The following mapping rules were agreed to convert original P_META 1.2 notation into XML.

2.2.1 From 'attribute names' to 'XML elements and attributes'

All P_META attributes have been converted into XML elements or attributes. The naming convention of elements names shall be to start each word with an uppercase (capital) letter and not to use separators such as '_'. Attributes names shall start with lowercase letters.

2.2.2 From 'alternative' to 'choice'

The original P_META (versions 1.1 and 1.2) notation used '|' as a means to signal that 'alternative' values (other options) could be provided.

For example, i|...|...|...|j is analysed as follows:

- 'i' indicates the minimum number of information elements that must be chosen from the alternatives ('i' must not exceed the total number of offered elements);
- 'j' indicates the maximum number of information elements that can be chosen from the alternatives ('j' must not exceed the total number of offered elements).
- An information element can be chosen once only.

If two or more information elements are included between two vertical bars separated by a colon ':', these information elements have to be considered as a single information element, as if they were included within a repeating group bounded to one, that is "0{..}1".

A qualifier name can be attributed if contextually required for clarification.

```
                <element name=”Q” minOccurs=”i” maxOccurs=”j”>
                                <ComplexType>
                                        <Choice>
[Q]i|E1|E2|..|En|j                              <element ref= ”E1”/>
where i,j <= n      <==>                        <element ref= ”E2”/>
                                        ...
                                        <element ref= ”En”/>
                                        </Choice>
                                </ComplexType>
                </element>
```

- i=0 means that no selection is required, which makes the element optional with minOccurs = '0'
- i!=1 means that at least 1 (or more) choice shall be made by repeating the element 'i' times without repeating previous selections
- j!=1 means that a maximum of 1 (or more) choice shall be made by repeating the element 'j-i' times without repeating previous selections. j >= i

2.2.3 From 'Repeating groups' to ' Complex Elements'

The original P_META (versions 1.1 and 1.2) notation used '{...}' as a means to signal a repeating group, which ordered sequence of attributes should be repeated as defined by the parameter of the repeating group.

- In i{.....}j, 'i' and 'j' define respectively the minimum and maximum numbers of repetitions on the information between brackets.
- {.....}n means that the group can be repeated indefinitely.
- 0{....} means the group is optional.

```
                <element name=”Q” minOccurs=”i” maxOccurs=”j”>
                        <ComplexType>
                                <Sequence>
[Q]i{E1:E2:..}j    <==>                 <element ref= ”E1”/>
                                                <element ref= ”E2”/>
                                                ...
                                <Sequence>
                        </ComplexType>
                 </element>
```

The 'group' is an XML sequence within an XML element qualified by an element name, and 'repetition' of the group is defined by the element cardinality (minOccurs and maxOccurs) boundaries.

2.2.4 From 'Sets' to 'Complex Types'

Original P_META Sets clearly correspond to XML complex types.

```
                    <complexType name="SetName">
                                    <sequence>
                                                <element name="Q1" type ="E1Type"/>
SetName[                                        <element name="Q2" type ="E2Type"/>
[Q1]E1:[Q2]E2:...[Qn]En  <==>                   ...
]                                               <element name="Qn" type ="EnType"/>
                                    </sequence>
                    </complexType>
```

2.3 Reference Data

2.3.1 Enumerated Lists

In XML, enumerated lists like those used in P_META 1.2 can easily be defined as shown in the following example:

```
<simpleType name="TitleType">
    <restriction base="string">
            <enumeration value="Original"/>
            <enumeration value="Main"/>
            <enumeration value="Working"/>
            <enumeration value="Alternative"/>
    </restriction>
</simpleType>
```

However, in P_META 2.1 the systematic use of Classification Schemes has been preferred to support extensions independently of the schema.

2.3.2 Classification Schemes

The definition and use of classification schemes is defined in EBU Tech 3336.

2.4 Namespace conventions

Namespaces should comply with the following conventions using date as a versioning qualifier, based on RFC5174 (EBU namespace).

Schema namespace:

- urn:ebu:metadata-schema:SchemaName_YYYYMMDD (Month and Day are optional)

Example: urn:ebu:metadata-schema:EBU_PMETA_20090701

Reference Data and Classification Scheme (CS) namespaces:

- urn:ebu:metadata-cs:ClassificationSchemeName_YYYYMMDD (Month and Day are optional)

Example: urn:ebu:metadata-cs:RoleCS_20080701

The 'ebu:' Unified Resource Namespace (URN) is being registered by IANA (www.iana.org).

2.5 Type Formats: W3C/XML

P_META uses most key XML simple datatypes (http://www.w3.org/TR/xmlschema-2/) e.g. date (ISO 8601 and RFC 3339), time (ISO 8601 and RFC 3339), string, int (integer), boolean, nonNegativeInteger, float.

2.6 Date and time formats

2.6.1 ISO 8601 and IETF RFC 3339

It is particularly important to respect the syntax for date and time (http://www.w3.org/TR/NOTE-datetime and IETF RFC 3339), which can be summarised as follows:

Year:
YYYY (e.g. 1997)

Year and month:
YYYY-MM (e.g. 1997-07)

Complete date:
YYYY-MM-DD (e.g. 1997-07-16)

Complete date plus hours and minutes:
YYYY-MM-DDThh:mmTZD (e.g. 1997-07-16T19:20+01:00)

Complete date plus hours, minutes and seconds:
YYYY-MM-DDThh:mm:ssTZD (e.g. 1997-07-16T19:20:30+01:00)

Complete date plus hours, minutes, seconds and a decimal fraction of a second
YYYY-MM-DDThh:mm:ss.sTZD (e.g. 1997-07-16T19:20:30.45+01:00)

where:

YYYY = four-digit year
MM = two-digit month (01=January, etc.)
DD = two-digit day of month (01 through 31)
hh = two digits of hour (00 through 23) (am/pm NOT allowed)
mm = two digits of minute (00 through 59)
ss = two digits of second (00 through 59)
s = one or more digits representing a decimal fraction of a second
TZD = time zone designator (Z or +hh:mm or -hh:mm)

It is not specified how many digits may be used to represent the decimal fraction of a second. An adopting standard that permits fractions of a second must specify both the minimum number of digits (a number greater than or equal to one) and the maximum number of digits (the maximum may be stated to be "unlimited"). Common practice is to use 3 digits.

Times are expressed in UTC (Coordinated Universal Time), with a special UTC designator ("Z"), or times are expressed in local time, together with a time zone offset in hours and minutes. A time zone offset of "+hh:mm" indicates that the date/time uses a local time zone which is "hh" hours and "mm" minutes ahead of UTC. A time zone offset of "-hh:mm" indicates that the date/time uses a local time zone which is "hh" hours and "mm" minutes behind UTC. A standard referencing this profile should permit one or both of these ways of handling time zone offsets.

Durations are represented by the format:

P[n]Y[n]M[n]DT[n]H[n]M[n]S (or P[n]Y[n]W[n]DT[n]H[n]M[n]S to use the week format).

In this representation replace [n] with the appropriate number for the element that follows it (leading zeros are optional but may clarify ambiguous durations). The capital letters ('P', 'Y', 'M', 'W', 'D', 'T', 'H', 'M' and 'S') are used as they are and not replaced.

Example:

"P3Y6M4DT12H30M0S" defines "a period of three years, six months, four days, twelve hours, thirty minutes, and zero seconds".

Elements may be omitted if their value is zero. To resolve ambiguity, "P1M" is one month and "PT1M" is one minute. The smallest value used may also have a decimal fraction, as in "P0,5Y" to indicate half a year. The number of seconds can include decimal digits to arbitrary precision.

2.6.2 Video and Audio time point references

P_META uses three methods to identify video and audio time point references:

- a time duration according to ISO 8601 or IETF RFC 3339
- timecodes as defined by SMPTE in specification ST 12-1:2008 (SMPTE Time & Control Code)
- a number of edit units, which are the fraction of time calculated as the inverse of the frame rate for video, or the inverse of the sample rate for audio.

Audiovisual entities generally embed the property of having a “Timeline”, which comes from the fact that the AV work is conceived to be played for a defined “Duration”, and all the events characteristic of the AV work itself are located on the Timeline.

The Timeline concept applies to AV ‘editorial entities’ as well as to the ‘physical entities’, which are the sources providing the AV material for actual realisations.

A typical application of the timeline mechanism is for identifying the location of a given AV-entity A which is a part (in time) of another AV-entity B.

As B has got its own duration D(B), we can say that A, with its own duration D(A), is located at point S of the Timeline of B.

This means that if A is located on the Timeline of B, from S to E, then E=S+D(A).

In P_META there are two mechanisms for expressing a position on a Timeline:

- the “Elapsed Time”, which gives the time elapsed on the Timeline of the reference entity (B in the example above) from its beginning.
  - the data type for that is a ISO 8601 duration(e.g. PT1M5.0S) or IETF RFC 3339;
  - the reference point for the elapsed time is always the beginning of the reference entity.

- the “Elapsed Units” which give the same information in terms of the number of Edit Units (which are countable)
  - this is to be preferred because it ensures that Timeline markers fall on the boundary of the Edit Unit;
  - duration of the EditUnit must be known unambiguously and indicated, otherwise it is better to use the “Elapsed Time”.

The two mechanisms mentioned above can also be used to locate the position of an AV-entity on the Timeline of a material source.

However there are contexts, in terms of the type of source, where the information in those terms is not available or it’s possibly ambiguous. For instance, identifying the position on a clip within a video-tape in terms of “Elapsed Time” or “Elapsed Units” from the “BOT (Beginning of Tape) is very difficult in practice. The BOT position itself may be not precise enough. In those cases, typically, the position on material source (e.g. the tape) is indicated by the “TimeCode”, which is a label recorded together with the EditUnit.

Although the “TimeCode” mechanism doesn’t provide any certainty about the uniqueness of the point on the Timeline (the same TimeCode might be repeated) and neither it provides reliable information on Duration (“TimeCode” is not constrained to be continuous), this is the way on which legacy production systems rely for editing and for saving EDLs (Editing Decision Lists).

This is the reason why P_META also supports the indication of TimeCodes for all the cases where the Timeline positioning deals with material sources.

However it is recommended to also provide, if available, the information in terms of elapsed time or edit units. This is supported by the P_META schema.

2.7 P_META Documentation

2.7.1 Schema Document Properties

Target Namespace	urn:ebu:metadata-schema:pmeta_2011
Element and Attribute Namespaces	❖ Global element and attribute declarations belong to this schema's target namespace.
	❖ By default, local element declarations belong to this schema's target namespace. ❖ By default, local attribute declarations have no namespace.

2.7.2 Declared Namespaces

Prefix	Namespace
Default namespace	http://www.w3.org/2001/XMLSchema
pmeta	urn:ebu:metadata-schema:pmeta_2011
xml	http://www.w3.org/XML/1998/namespace

2.7.3 Elements, Simple Types and Complex Types: Definitions

All the definitions can be found in the P_META schema.

3. Implementation Guidelines

Several aspects of the specification are left to the appreciation of the implementer (e.g. regarding the implementation of classification schemes or mapping to pre-existing in-house metadata schemas). It is foreseen that feedback will be provided, which would be reflected in future implementation guidelines.

It is also expected that application specific guidelines will be provided as part of the different application-based specifications.

4. Maintenance

P_META will be maintained by the EBU and suggestions for corrections or additions can be made at (metadata@ebu.ch). EBU members can also provide feedback via the EBU Technical Department's website (http://tech.ebu.ch/MetadataMaintenanceSpecifications).

Contributions will be subject to peer review by the metadata experts participating in ECM/MAG (http://tech.ebu.ch/groups/pmag), a specialised Project Group of the Metadata Expert Community (ECM)

5. Download Zone

The following documents are available for download:

Filename	Document description	Contents
http://www.ebu.ch/metadata/schemas/pmeta/2011/EBU_PMETA_2011.zip	Schema	EBU_PMETA_2011.xsd xml.xsd

6. Other EBU Metadata Specifications

The following applications are using or are directly related to the use of P_META:

- Tech 3293 – EBU Core (http://tech.ebu.ch/docs/tech/tech3293v1_3.pdf)
- Tech 3331 - EBU Exchange Metadata Schema, which defines a common core set of metadata for B2B content exchange (http://tech.ebu.ch/docs/tech/tech3331v1_1.pdf).
- Tech 3332 - EBU Music Report Metadata Schema, which defines a common core set of metadata for reporting music broadcasts (http://tech.ebu.ch/docs/tech/tech3332v1_1.pdf).
- Tech 3336 - EBU Classification Schemes (http://tech.ebu.ch/docs/tech/tech3336v1_1.pdf).
- R 128 - EBU Recommendation on 'Loudness Normalisation and Permitted Maximum Level of Audio Signals' (http://tech.ebu.ch/docs/r/r128.pdf)

7. Useful links

- AES (http://www.aes.org)
- Dublin Core (http://www.dublincore.org)
- EBU Metadata (http://tech.ebu.ch/metadata)
- ETSI (http://www.etsi.org)
- IANA (http://www.iana.org)
- IETF (www.ietf.org)
  - RFC 3339 – Date and Time (http://www.ietf.org/rfc/rfc3339.txt)
  - RFC 5174 – EBU Namespace (http://www.ietf.org/rfc/rfc5174.txt)
- SMPTE (http://www.smpte.org)
  - Timecode : SMPTE 12M
  - MXF : SMPTE 377 M
- SMPTE ST 12-1:2008 (Time and Control Code) (http://store.smpte.org/category-s/22.htm)