MAD

Deliverable D15.1MDS1: Analysis of current audiovisual documentation models, Mapping of current standards

DOCUMENT IDENTIFIER	PS_WP15_ORF_D15.1_Documentation_Models_v1.6
DATE	16/3/2005
ABSTRACT	Analysis of the most popular audiovisual and currently used data-schemas within the consortium with the intention to provide support for the decision upon the PrestoSpace-Datamodel. Mapping of the analysed standards
AUTHOR, COMPANY	Christoph Bauer, Florian Rosensprung, Stefan Lajtos (ORF) Laurent Boch (RAI) Philippe Poncin (INA) Conny Herben-Leffring (B&G)
KEYWORDS	Documentation-model, Data-schema, mapping
RELATED ITEMS	D15.2
NATURE	Report
DISSEMINATION	Public

DOCUMENT HISTORY

Release	Date	Reason of change	Status	Distribution
0.1	2004-02-29	First Draft	Living	Confidential
0.2	2004-12-13	Exc.Sum. and Overview altered	Living	Confidential
1.0	2005-01-11	Schema-Descriptions inserted	Living	Confidential
1.1	2005-01-24	Mapping added	Living	Confidential
1.2	2005-01-26	Minor changes	Living	Confidential
1.3	2005-01-31	Rearrangement of chapters	Living	Confidential
1.4	2005-02-08	Changes in Conclusions	Living RC	Confidential
1.5	2005-03-16	IMMIX added	Living	Public
1.6	2005-03-16	Final adjusments	Published	

1 Executive Summary

This deliverable will be a report containing an analysis of representative audiovisual documentation models adopted by the major organisations participating in the project plus a set of standard-models with notable prevalence. Include is an overall mapping of entities and attributes into a common denominator. The focus is on the practical aspects of the different models in reference to the main focus of the overall workgroup as well as on the migration of legacy metadata.

2 Overview

To be able to decide on a documentation-framework (see D15.2) for PrestoSpace it is vital to analyse currently adopted audiovisual documentation models in the face of the main tasks of the future “PrestoSpace-model” as well as legacy documentation models awaiting the task to migrate those legacy metadata into a future PrestoSpace-Data-model:

- Providing a “minimal set” of universally valid descriptors.
- Having customisable structures to meet user common practice and retrieval/exploitation objectives of the different future user-companies
- Being able to accommodate existing standards and practices for the task of migrating legacy metadata as well as
  - Preservation technical Metadata
  - Restoration technical Metadata
  - Workflow Metadata
  - Descriptive Metadata (s.a. D15.3)
  - etc.

The main focus will be on models adopted by members of the PrestoSpace-Consortium; additional models/standards will be analysed if their market-penetration and technical potential is high enough to be considered.

An overall matching of the current adopted and utilized models will lead to a list of entities and a minimal dataset (common denominator), which can be taken as a starting-point for the decisions on a future PrestoSpace-Data-model, to be enhanced with subsets fulfilling the special needs of the PrestoSpace-factories and the workflow within.

3 Data Models (Customized/Legacy)

3.1 RAI-Documentation Model

3.1.1 Introduction

The current RAI[1] documentation model has been defined in order to populate the Rai archives **Multimedia Catalogue**, which is the system providing to the internal users the information about the huge amount of archived programmes.

The system has been in service since 1997, and today the process of documenting with the current model the older programmes is at a very good stage.

The user of Multimedia Catalogue can perform several activities (also depending on his/her user profile):

- Search

  Through a single web-based front-end application, the user can search an archived programme within all databases (including legacy ones) either with a simple interface or with a specific one

- Navigation

  Once that a single item is selected, the user can explore the documentation layers following the main entity relations.

- Browse Material

  The Multimedia Catalogue includes key frames, audio clips and movies in order to represent the described entities and complete their description with a direct inspection of a browsing quality Material

- Get for re-use (if enabled)

  The user can select, through the key frames representation, Material segments of which he/she asks to get a higher quality copy for re-use.

The current documentation model clearly depends on the features and the objectives of the Multimedia Catalogue and thus it happens to be quite complex.

This section tries to describe its most important features, by giving a broad synthesis accompanied by a few UML diagrams. Many minor details are omitted and a few characteristics could have been reported slightly imprecisely. The names of entities, attributes, and relations have been translated from Italian language trying to use a terminology already known in the PrestoSpace MAD area, in some cases taken from EBU Tech. 3295.

The first section deals with the documentation process, which is actually part of the ingestion of the catalogue system.

The second section provides information which is specific of the identification issue, which is considered the most important one.

The third section in effect demonstrates this fact by describing the core documentation profile, named “Simple Registry”

The fourth section eventually gives a fairly itemised illustration of the descriptive part of the documentation model.

[1] http://www.rai.it/

3.1.1.1 The documentation process

Jobs organisation

The documentation work flow starts with the preparation of jobs to work through, which can be divided into two different cases:

Working currently broadcasted programmes

Working programmes from archives

In both cases there is the assumption that it must be solved the issue of linking Editorial Objects (the programmes, logically) and Material (copies/tapes/location, physically).

In the first case the starting point is the “on air report” which lists the Editorial Objects related to their publication events. The published Material is used for the first link to a physical copy, while links to other Material are derived afterwards.

In the second case the Archive department prepares the jobs list where the Editorial Objects are already linked to the Material copies to be used in the continuation.

Multimedia acquisition

The next step is called “multimedia acquisition”, which consists in the recording Material for the documentation and for the ingestion into the Multimedia Catalogue.

The recorded Material consists of:

- Key frames, one for each scene change detected
- Low quality audio
- Low quality video

If required, such Material is then cut and packaged in order to match Editorial Objects (necessary for 24 hours acquisition from on air channels).

Documentation

The documentation itself is typically done by people belonging to other organizations (working outside Rai) who can use the key-frames segmentation and the low quality movie during their work.

The external organization has to adopt a validation practice in order to guarantee a defined quality level.

Not all programmes are documented as for some of them only the identification is required, considering the availability of key frames and movie, and a ***Simple Registry*** profile is adopted.

Documentation quality control

When the documentation is complete the work done is checked by professionals within Rai. If the documentation organization is considered reliable, this process is performed on a sample basis otherwise (new organization or got problems) it is done for each item.

Publication on Multimedia Catalogue

The *Simple Registry* profile information is provided to the users as soon as ready, even if the documentation process is not complete.

When the documentation is available with the OK flag from the quality control team, it is added to the to *Simple Registry* information.

3.1.1.2 The identification scheme

The entity which is the object of the identification effort is the ***Program***, the main fields of which are:

- Title
- Identifiers (archive number and accounting reference number)
- duration
- Category and genre (organisational classification scheme)

Program instances are related to *Material* - in turn related to the *Storages* - and to the *Publication Events* – dates, times and services.

The Model defines that a *Program Version* (related to the *Program* of which it is version) can be related to a single ***Series***, which is a collection of *Program Version* instances with own title and identifiers set.

For the Program there is a ***credits list*** collecting records with the following information:

- Role (organisational enumeration, pmeta:ROLE_TYPE_NAME), required
- Person, required
- “Qualification” (job title)
- Character name (role played in a fiction, pmeta:ROLE_NAME)

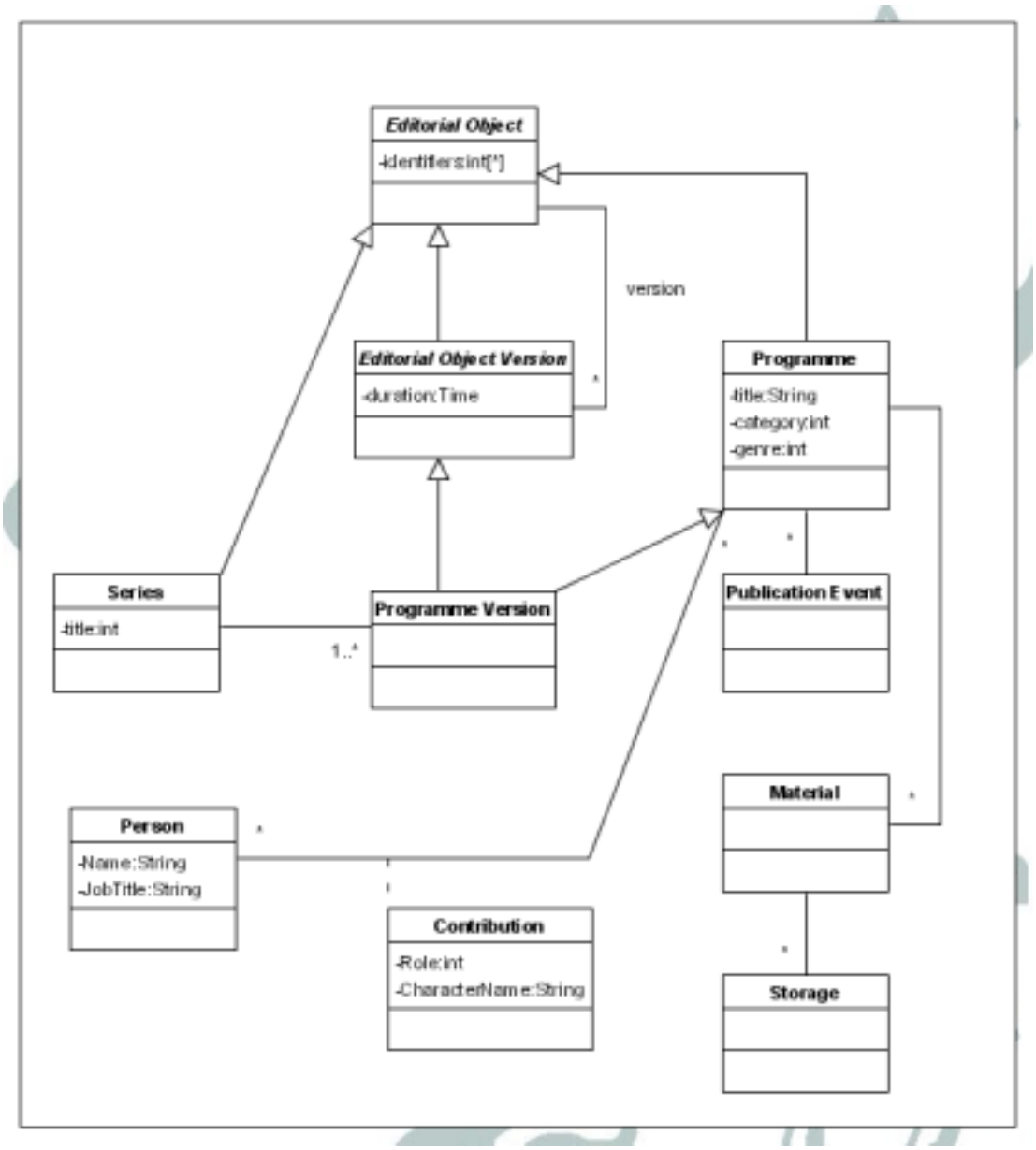


Figure 1 - identification entities

As an example, the Program instance X could have been broadcasted five times, three of which it was X1 version X and episode of Series Y1 and two other times it was X2 again version of X but episode of Series Y2.

Notice that X1 and X2 are not considered editorially equivalent, because they belong to different Series and they may have differences in title, duration and, slightly, in content.

3.1.1.3 The "Simple Registry" documentation profile

This is the minimum documentation profile the concepts of which are valid for all the other more complex profiles too.

It consists basically of the information provided by the ***Identification Scheme*** and the Shot segmentation.

Shot segmentation

The Editorial Object Version is automatically segmented into ***Shots***, which are defined to be *a continuous part of a scene taken from a single point of view*, and can be considered as Editorial Object as well.

For each Shot it is required to know the start position within the Editorial Object of which it is part and its duration.

A key frame has to be available in the Multimedia Catalogue in order to represent a Shot instance.

A simple colour description (BW/Colour) of the Shot is obtained automatically.

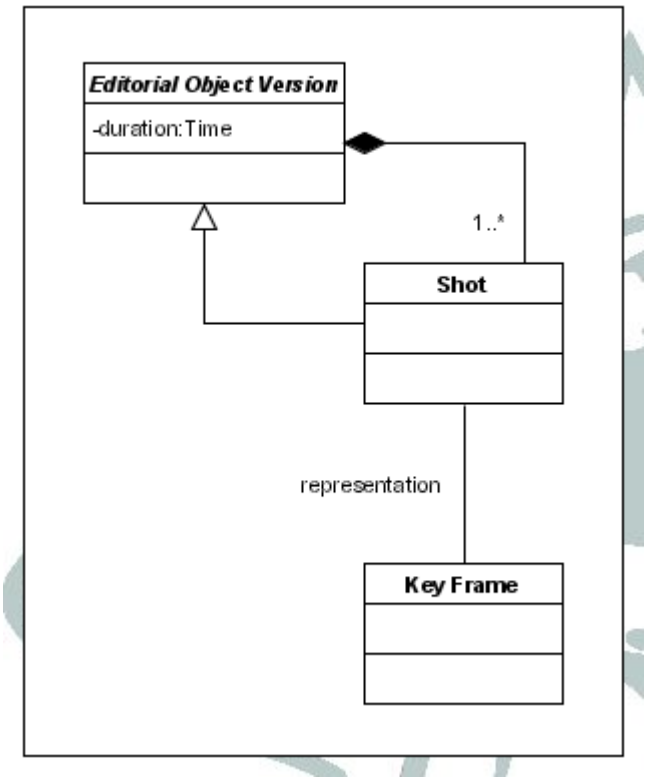


Figure 2 - the Shot segmentation

The Shot becomes the access unit of the Multimedia Catalogue and all the entities identified in the following processes of documentation inherits the results of the Shot segmentation (they can be seen as composed of Shots and they have links to their Shots).

3.1.1.4 The descriptive documentation model

The documentation process applies to the ***Program Version*** entity, which is discriminated into several specializations according to the most suitable documentation profile.

Generally we can find two different approaches which can be mixed with a variable balance in order to get the desired result:

- Synthetical documentation

  It tends to collect information which is related to the whole object

- Analytical documentation

  It tends to identify more precisely the entities to which the pieces of information are related

One of these entities is named ***Pointed Out Item***, which is an identified segment of a whatever *Editorial Object Version* and that has been pointed out to give a relevant description. The identification is given by the relative position (start and duration) in the timeline of the object of which it is a segment.

A similar concept is the ***Item*** which is specifically a part of a *Program Version* and that is further specialized. Besides the *Item* may be related to *Real Life Events* and to a list of ***Participants*** the records of which are exactly as those of the *Credits list* of the Identification Scheme.

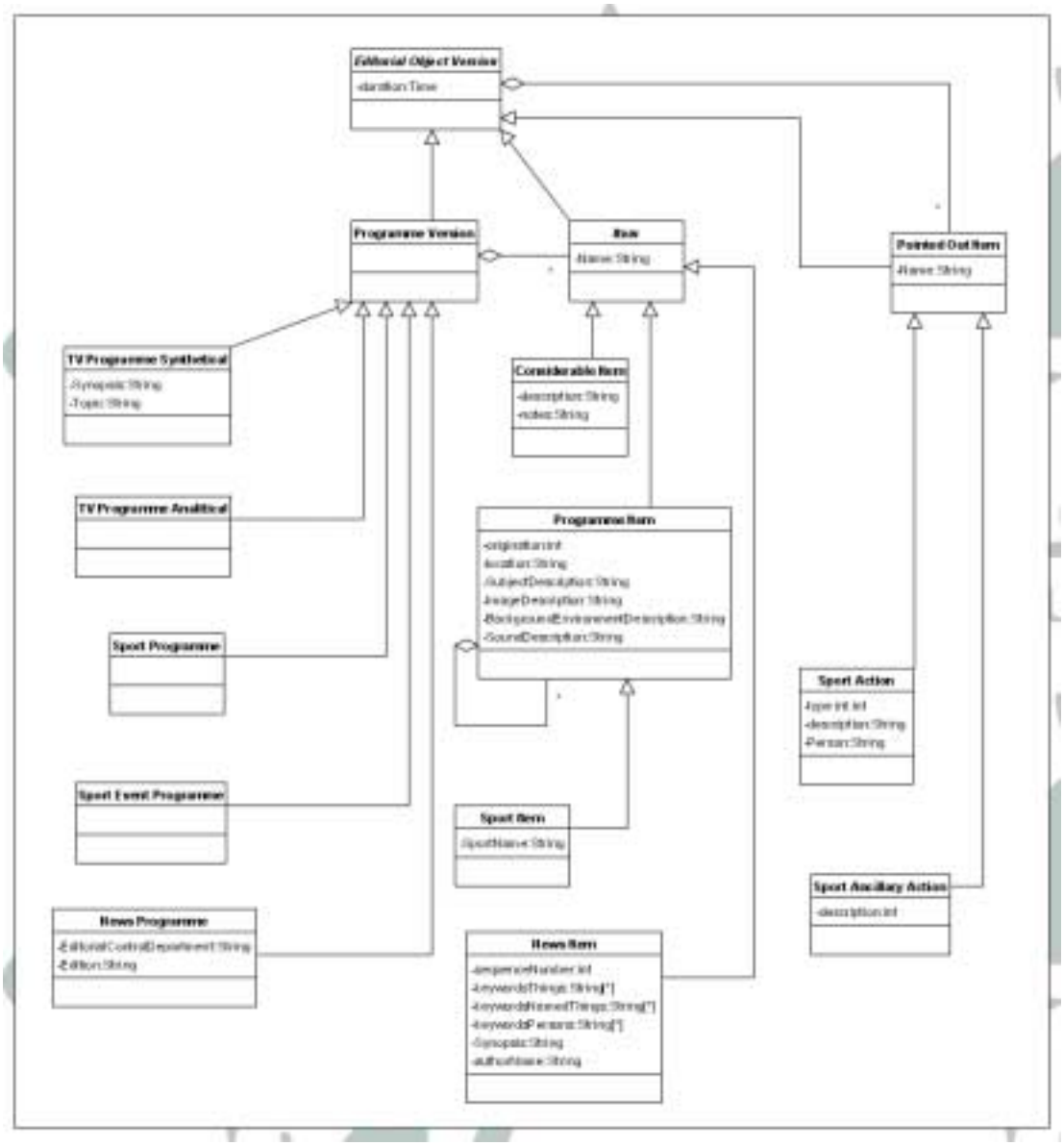


Figure 3 - broad documentation model diagram

The diagrams showed in Figure 3 and 4 provide the basis of the description of several documentation profiles.

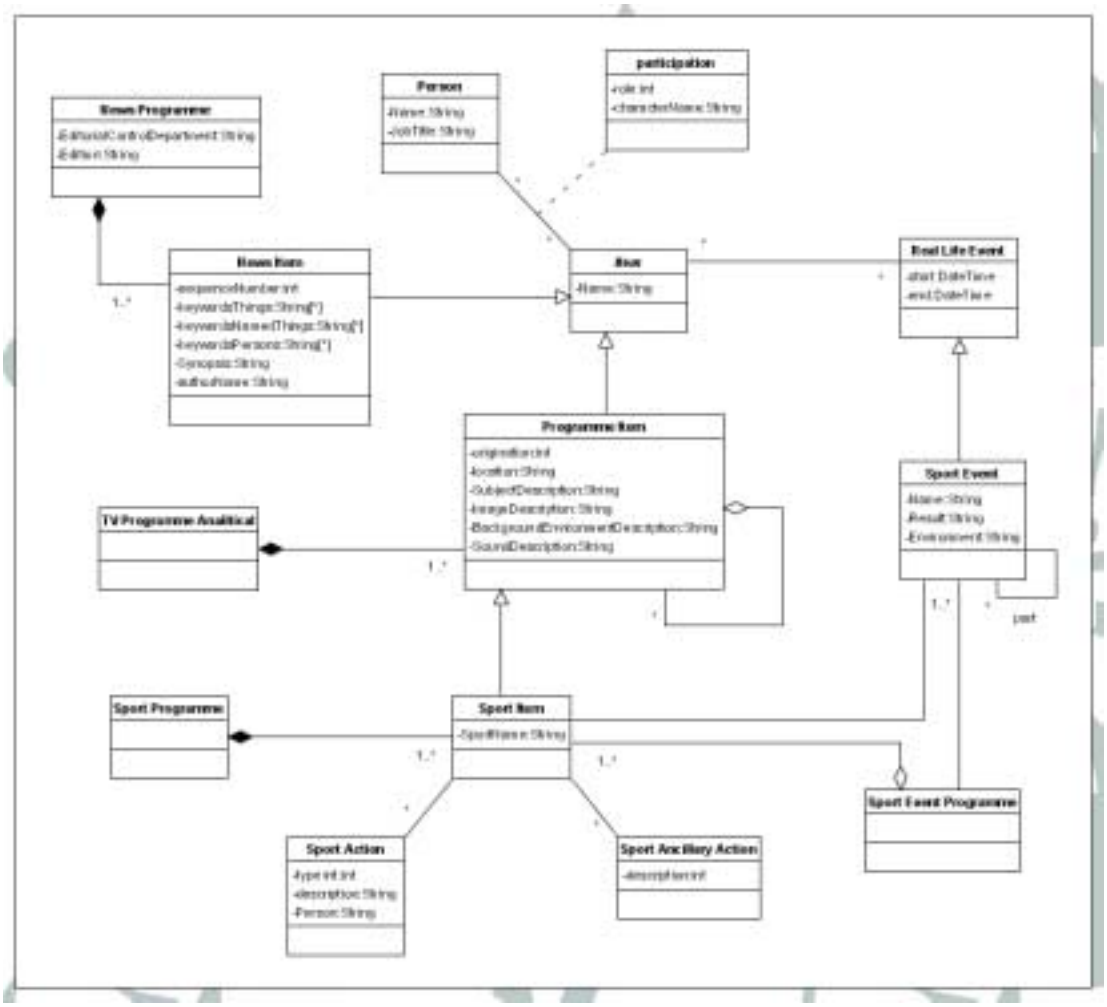


Figure 4 - documentation details

TV Programme Synthetical

This documentation profile has all the features of the *simple registry* one with the following additions:

A synopsis/description for the whole Program Version instance

A further classification field “topic/subject” based on an organizational classification scheme

However in some cases it is required that a few ***Considerable Items*** - which are part of the *Program Version* – have to be identified in order to get a specific description, together with a specific *participants* list and the dates of the possibly related *Real Life Events*.

TV Programme Analytical

This documentation profile differs from the *synthetical* one because it hasn’t any synopsis description for the whole editorial object instance, but it has a complete segmentation into ***Program items***.

The default grouping of *Program Items* is the chronological sequence along the editorial timeline, however other kinds of grouping are used according to the subject (items sharing subject) or the program structure (e.g. studio/discussion vs. edited service).

The *Program Item* can be further segmented into *Program Items* in order to provide an increasing specificity to the documentation.

The information fields of the *Program Item* are:

- Name (or title)
- Duration (inherited from *Editorial Object Version*)
- Origination (definition similar to pmeta: ORIGINATION_CODE but different values set)
- Location
- Subject (free text)
- Image description (free text)
- Background/environment description
- Sound description

Furthermore it inherits from *Item* the reference to *Real Life Events* and the specific *participants list*.

Sport Programme

This documentation profile is intended for Programmes dealing specifically with Sport without having the goal of covering a single particular sport event.

The ***Sport Program*** is completely segmented into ***Sport Items***, which are a specialization of *Program Item*.

In particular the *Sport Item* has to indicate the name of the Sport which is the subject of the item and it has to be related to a set of ***Sport Events***, considered as in a hierarchy as it follows:

- the sport main container event (e.g. champions league)
- the sport sub container event (e.g. second turn)
- the sport actual event (e.g. match Juventus-Milan)

For the sport actual event it is significant to have:

- the Event result: score of sport event (e.g. 1-0)
- a brief description of the environment

Moreover the *Sport Item* can be related to some *Pointed Out Items* which are part of it and are discriminated between:

- ***Sport Action***

  The *Sport Action* is an item where a *Person*, briefly identified, performs an action characteristic of the particular sport. The type of action is classified according to an organisational scheme including values such as “goal”, “penalty”, “red card” for football, etc. The action is described textually.

- ***Sport Ancillary Action***

  The *Sport Ancillary Action* is an item where someone, briefly identified, does something related to the *Sport Item*, but not technically related to the sport practice. Examples are: “a football player interviewed”, “teams listening the hymns”, “viewers and fans showing banners”. The action is described textually.

Sport Event Programme

The main difference from the *Sport Program* documentation is that here the ***Sport Event Program*** is deemed to cover completely a single particular *Sport Event*.

An example of this is the transmission of a Formula One Grand Prix.

The documentation profile uses the *Sport Item*, the *Sport Action* and the *Sport Ancillary Action* entities, where the balance is heavily towards the two *Pointed Out Item* specializations.

As an example the items of a football match could be:

- Interviews before the match
- First half
- Interviews and comments during the break (can include replays)
- Second half
- Interviews and comments after the match

Within all these items, especially the two halves, there can be many *Sport Actions* pointed out and described.

Of course the Real Life Event context is always the same, so it is shared.

News Programme

The News documentation profile is a “legacy” one, compared to those described before, and it is going to be modified in the future.

The ***News Program***, about which there have to be known the organization department having the editorial control and the name of the edition (properly part of the identification scheme), has to be completely segmented into ***News Items***, the names of which constitutes the headlines list of the *News Program*.

The *News Item* has also the following fields:

- a sequence number
- a synopsis description
- dates of Real Life Events (inherited from *Item*)
- author name – equivalent to pmeta: ROLE_TYPE_CODE where value is ”Broadcast Journalist” – explicit although inheriting from *Item* also the *participants list*
- Keywords related to “things”
- Keywords related to “named things”
- Keywords related to “persons” – again although inheriting from *Item* the *participants list*

An important known limitation is that a field about the location is missing and in practice this information is typically given as free text within the synopsis.

3.1.2 Conclusion

RAI introduced their DM in 1997, it allows both synthetic and analytical documentation (sub-items can be identified and described).

3.1.2.1 Main entities of RAI-DM

- program
- collection (of programs)

- item (inside a program)

External relations allow e.g. the identification of people with a given role, real-life-events, etc…

The material itself is linked to the program-entity

3.1.2.2 Standards and Exchange

Currently Rai archives are prepared for exporting (and importing) according to EBU Tech. 3295 (P_META), having developed a generic P_META filter application, which implements any configured export mapping from Rai organisational information sources, such as databases, xml files, and web-services, to P_META over XML.

This approach is generally adoptable by any organisation and doesn't rely on any particular feature of the Rai model.

3.2 INA – Documentation model

3.2.1 Introduction and Overview

The current INA[2] documentation model has been defined in order to populate the Ina archives data base, which is now called the *Totem* system providing to the internal users the information about the huge amount of archived programmes.

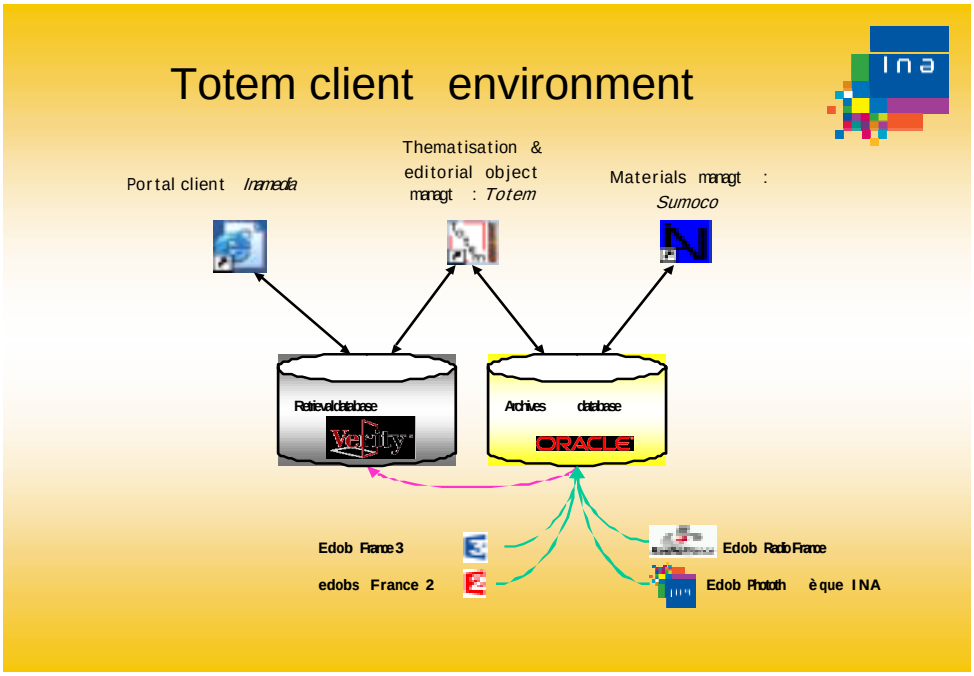
The system has been in service since 2004, and today the process of documenting with the current model the older and the current programmes is at a very good stage. INA's data model does not rely on any metadata standard, it has been defined trough an evolving process since 20 years in order to meet professional requirements for managing contents and materials in the audiovisual domain.

The developments during the last years intend to describe new entities (from abstracts to thematic collections), and to increase access and reuse for clients :

- Defining new entities to manage abstracts which are often selected, and to gather thematic collections.
- Providing retrieval and viewing facilities via INAmedia application on the Web.
- Including essence management of MPEG files.

The technical environment has been shared into three parts :

1. An Oracle data base for creation and modification of editorial objects and for management of the different materials.
2. A Verity engine for full text indexing and retrieval.
3. The Totem applications for users either internally to modify Oracle data, or externally for request and access through Verity.



The user of *Totem* can perform several activities (also depending on his/her user profile):

[2] http://www.ina.fr/

- Search

  Through a single web-based front-end application, the user can search an archived program within all databases (including legacy ones) either with a simple interface or with a specific one

- Navigation

  Once that a single item is selected, the user can explore the documentation layers following the main entity relations.

- Browse Material

  The Multimedia Catalogue includes key frames in order to represent the described entities and complete their description with a direct inspection of a browsing quality Material.

- Get for re-use (if enabled)

  The user can select, through the key frames representation, Material segments of which he/she asks to get a higher quality copy for re-use.

The current documentation model clearly depends on the features and the objectives of either professional or legal deposit uses and thus it happens to be quite complex.

This document tries to describe its most important features, by giving a broad synthesis accompanied by a few diagrams. The names of entities, attributes, and relations have been translated from French language.

The first section deals with the documentation process, which is actually part of the ingestion of the catalogue system.

The second section provides information which is specific of the identification issue, which is considered the most important one.

The third section eventually gives a fairly itemized illustration of the descriptive part of the documentation model.

3.2.1.1 The documentation process

Jobs organisation

The documentation work flow starts with the preparation of jobs to work through, which can be divided into two different cases:

- Working currently broadcasted programmes
- Working programmes from archives

In both cases there is the assumption that it must be solved the issue of linking Editorial Objects (the programmes, logically) and Material (copies/tapes/location, physically).

In the first case the starting point is the “on air report” which lists the Editorial Objects related to their publication events. The recording of full broadcast programmes is done “on air” directly to MPEG files (both Mpeg2 -8Mbits and Mpeg1 -1,5Mbits). These Materials are used for the first link to a physical media, while links to other Material are derived afterwards when the professional elements are delivered by the broadcasters.

In the second case the Archive department prepares the jobs list where the Editorial Objects are already linked to the Material copies to be used in the continuation (after digitisation of previous analogue materials).

Multimedia acquisition

The next step is using the “Media Index” application which consists in import of programmes segmentation made on Mpeg1 files and for their ingestion into the Documentation database.

The recorded Material consists of:

- Key frames, one for each scene change detected
- List of sequences time codes.

If required, such Material is then cut and packaged in order to match Editorial Objects (necessary for 24 hours acquisition from on air channels).

Documentation

The documentation itself is typically done by Ina’s librarians who can use the key-frames segmentation and the low quality file during their work.

Not all programmes are documented as for some of them only the identification is required, considering the availability of key frames and movie, and a *Cataloguing* profile is adopted.

Documentation can apply on different levels, for a whole programming session, or for a specific program (single or composite) or for programmes items if necessary (e.g. news and documentaries).

Publication on Media Index

The *Media index* profile information is provided to the users as soon as ready, even if the documentation process is not complete.

When the documentation is available with the OK flag from the validation team, it is added to the to *Data base* information.

3.2.1.2 The identification scheme

The entity which is the object of the identification effort is the *Program*, the main fields of which are:

- Title
- Identifiers (archive number and accounting reference number)
- Date of transmission
- Duration
- Production code and Producers
- Authors
- Genre and Form (organisational classification scheme)

The Model defines that a *Program* can be related to a single *Collection*, which is a collection of *Program* instances with own title and identifiers set.

For the Program there is a *Cast list* collecting records with the following information:

- Person, required
- Role, required
- “Qualification” (job title)
- Character name (role played in a fiction)

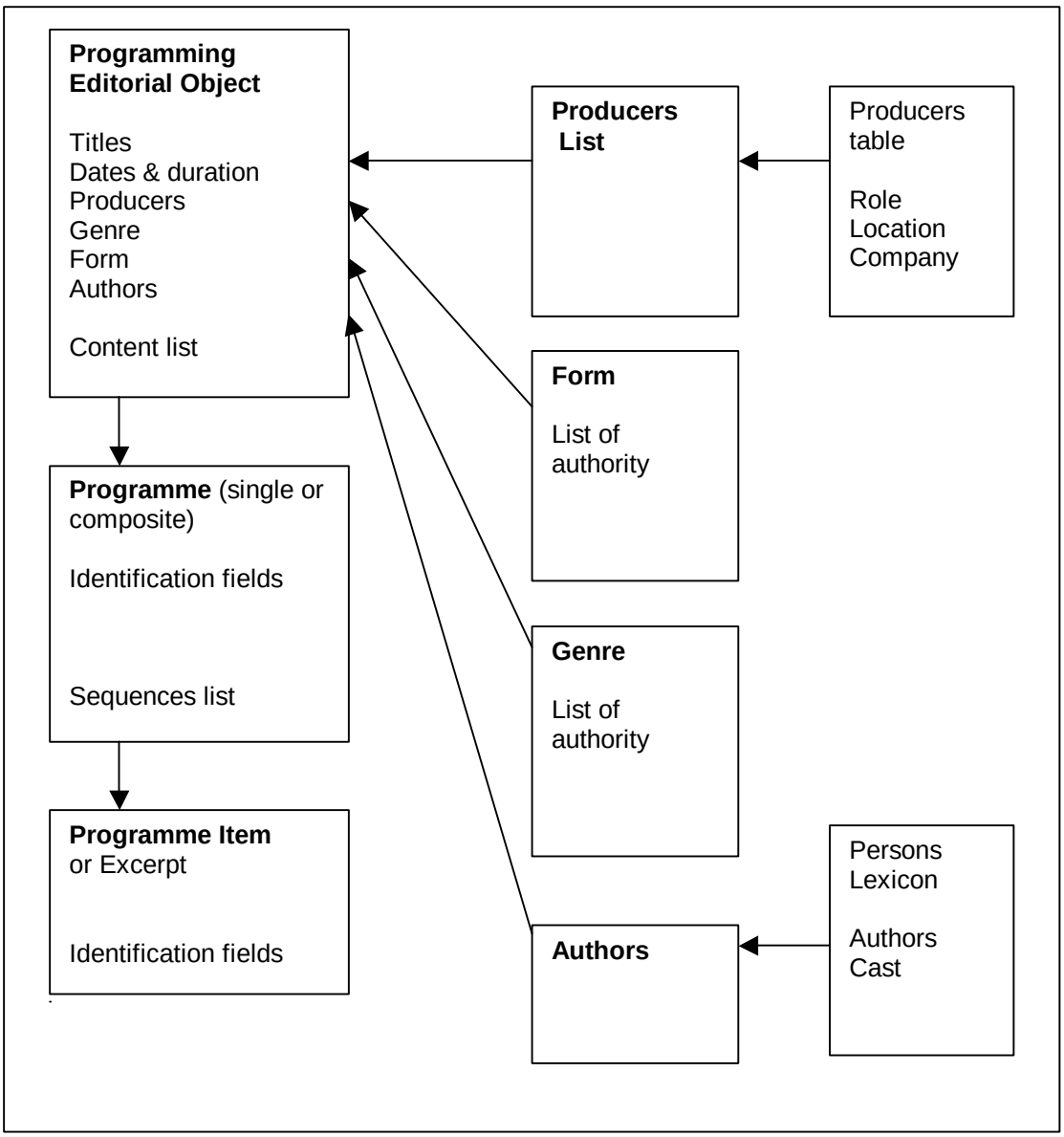


Figure 5 - identification entities

The *Programme* instances are related to professional materials and elements of materials - in turn related to the Storage locations. These are described in another database and the relations are made through a *Link Table* as represented in figure2.

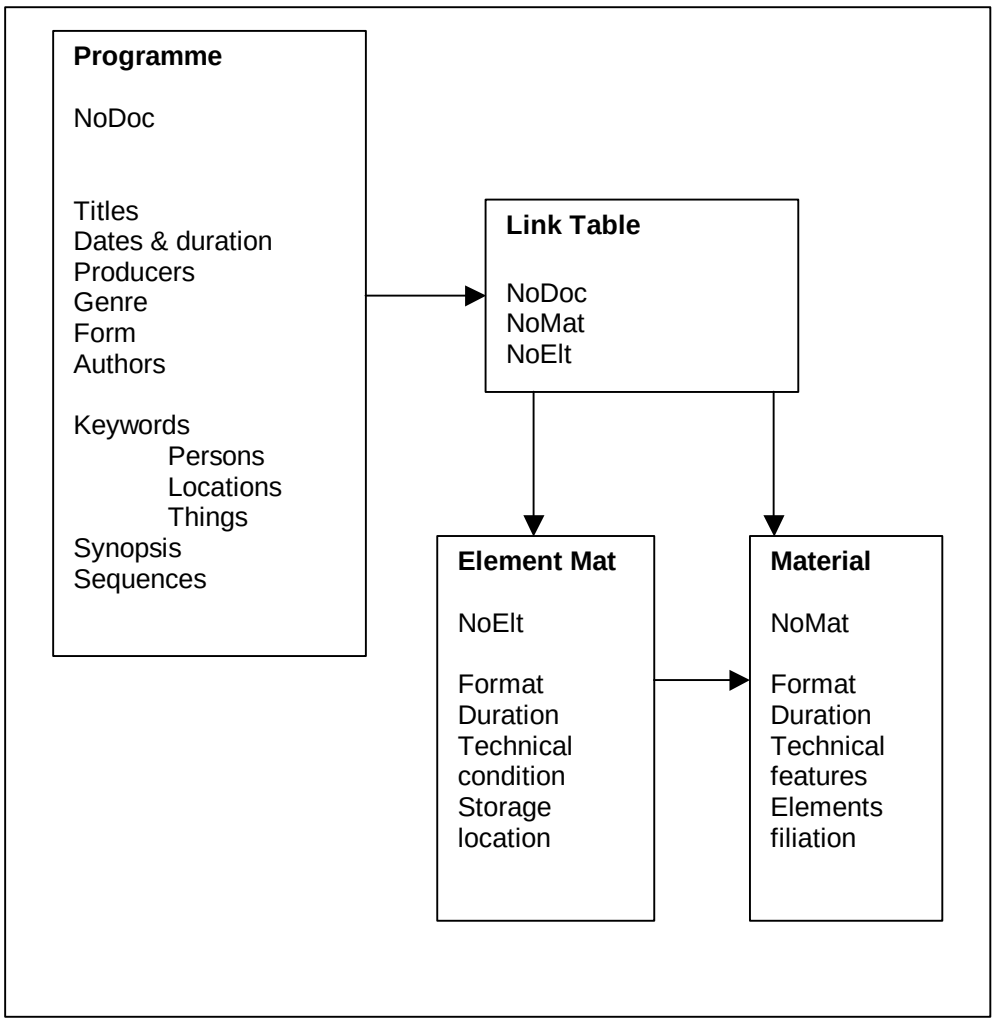


Figure 6 – description of materials

For professional uses the elements are the physical entities that are really moved and will go to the workflow to be processed until final delivery. The material describes the conceptual entity for each kind of material (format and standard) and includes the filiation to all existing elements.

The same architecture is used to link programmes to the MPEG files (both Mpeg2 -8Mbits and Mpeg1 -1,5Mbits). Therefore the viewing file can be directly accessed from the Program instance because they are available on remote disks servers.

For a Program item or excerpt, it is even simpler because the Mpeg file relates only to a single entity and the link is made directly in the item description.

3.2.1.3 The descriptive documentation model

The documentation process applies to the ***Program*** entity, which is discriminated into several specialisations according to the most suitable documentation profile.

Generally we can find two different approaches which can be mixed with a variable balance in order to get the desired result:

- Synthetical documentation

  It tends to collect information which is related to the whole object

- Analytical documentation

  It tends to identify more precisely the entities to which the pieces of information are related

One of these entities is named ***Program Item***, which is an identified segment of a whatever *Editorial Object* and that has been pointed out to give a relevant description. The identification is given by the relative position (start and duration) in the timeline of the program of which it is a segment.

Besides the *Item* may be related to an action, an event or an interview and to a list of *Participants* the records of which include those of the *Credits list* of the Identification Scheme and other less important participants related to a second Person lexicon (not referenced in the Person Thesaurus).

TV Programme Synthetical

This documentation profile has all the features of the *Identification catalogue* with the following additions:

- A Synopsis/description for the whole Program Version instance
- A further classification using fields for topics, persons, locations based on a Keywords index.

This applies to telefilms, feature series, magazines and variety shows.

TV Programme Analytical

This documentation profile differs from the *synthetical* one because it hasn't any synopsis description for the whole editorial object instance, but it has a complete segmentation into ***Program items***.

The default grouping of *Program Items* is the *Content list* of sequence along the editorial timeline, however other kinds of grouping are used according to the subject (items sharing subject) or the program structure (e.g. studio/discussion vs. edited service).

The *Program* description includes an abstract of each sequence and principal descriptors related to them. The *Program Items* include *Secondary Keywords* related to the sequences in order to provide an increasing specificity to the documentation.

The information fields of the *Program Item* are:

- Name (or title)
- Duration (inherited from *Program*)
- Origination and production type
- Person, Location and Topics Keywords
- Summary (free text)
- Shoots description (free text)
- Background/environment description

This applies to commentated reports, documentaries and elaborated studio recordings.

Documentary Programme

The ***Documentary Program*** is completely segmented into Program *Items*.

In particular the *Documentary Item* has to indicate the precise locations, the date of shooting, the persons involved and the event environment.

The actions are described textually and the purpose is to allow precise retrieval for excerpts reuse in production process.

News Programme

The ***News Program***, about which there have to be known the organization department having the editorial control and the name of the edition (properly part of the identification scheme), has to be completely segmented into ***News Items***, the names of which constitutes the headlines list of the *News Program*.

The News description finally includes a program entity for the whole editorial object and a Program entity for each News item.

The *News Item* has also the following fields:

- a sequence number
- a synopsis description
- dates of Real Life Events
- author name –where value is ”Broadcast Journalist” – but also the whole crew
- Keywords related to “things”
- Keywords related to “locations”
- Keywords related to “persons”

One of the key information is the origin of the sequences to allow rights management and clearance for reuse.

3.2.1.4 Elements

A table of the main elements with their format and content is provided in the Annex.

3.2.2 Conclusion

The INA Documentation model, introduced in 2004, allows both synthetic and analytical documentation (sub-items can be identified and described).

3.2.2.1 Main entities of INA-DM:

- programme (single or composite)
- programme-item (or excerpt)
- programming editorial object

Material and elements of material are described in a separate database and related via a link-table to the programme instances.

External relations are given to lists of producers, second persons lexicon (those not in the person thesaurus), etc.

3.2.2.2 Standards and exchange

INA's data model does not rely on any metadata standard, it is a system directly "grown" out of the professional requirements for managing contents and materials in the audiovisual domain.

For data exchange and system interoperability, INA implemented an export in XML format which gives a standard and structured form for further uses.

3.3 FARAO (ORF)

3.3.1 Introduction

FARAO is the TV-Database-schema used by the Austrian Broadcasting Corporation (ORF) [3]. Since 1984 this host-based system has been online, so in this paper FARAO is the exponent of the very old and completely specialised schemas. Beneath this quite antique part of FARAO all employees of ORF are able to call up “multimedia-elements” (keyframes) of all news-programs since August 1998.

FARAO started as a simple “storage and material management system” and has been widened very soon to provide synthetical and analytical documentation. There is a subdivision of the 62 elements of FARAO which are “Structural Data”, “Content Date”, “Rights”, “Transmission Data”, “Recording Data”, “Carrier Data”, “Technical Data”, “Reference Data”, “Acquisition Data”, “Loan Data”, “Storage Data” and “Other Data”. FARAO is strictly index-list-based and works with “controlled vocabulary” in some elements.

3.3.1.1 Elements / Fields

Field Nr.	Fieldname	Description
Structural Data		
1	Programme Counter	Counter for programmes on carrier
2	Contribution Counter	Counter for contributions (within programme) on carrier
3	Record Date Of Creation	
Content Date		
4	Programme Title	Programme-title and additional information in order to distinguish carriers in list-view. Additional information may be: type of generation (original, duplicate), episode number, aspect ratio, TV-system, production year etc. program-title on the "content-page" may differ from the programme-title on the "carrier-page".
5	Subtitle	May also contain original and alternative title
6	Contribution Title	(or chapter-title.) May also contain warning about blocking notes and missing material.
7	Content (story)	Short summary of content
8	(Motives) Description (visual content)	Additional rules for segmentation of description into "sequences" in practice, additional vocabulary for standardized aspects (day/night, season, good-shot, clippings of archive material, style of editing, tricks & animation etc.) in use.
9	Person	Both production staff, authors/creators and persons or ensembles appearing on screen as well as indicator on used photographs
10	Keywords	Accumulative selection from approx. 60 keywords describing general thematic fields
11	Geography	real geographical place of action
12	Timeframe Content	real temporal boundaries of action
13	Programme Format	Accumulative selection from approx. 80 programme formats, target-groups and genres.
14	Production Country	for movies
15	Production Date	for movies
16	Programme Length	
17	Contribution Length	

[3] http://www.orf.at/

Rights		
18	Creator	Company responsible for the creation / production according to credits
19	Basic Rights Qualifier	See summary
20	Production Number	Internal ID for productions, which may encompass one or several contributions or programmes. Allows to link into the ORFEUS system.
21	Production Subnumber	Counter for production numbers, which encompass more than one "title".
22	Rights data Available Flag	
23	Note of blocking	
Transmission Data		
24	Date	May also contain flags for "working-material" and "collection tapes" and may represent recording date instead of transmission date.
25	Channel	
26	Date	
27	Channel	
Recording Data		
28	Tape-ID	
29	Length on Tape	
30	Reduced Length	
31	Timecode Start	of the programme / contribution
Carrier Data		
32	Tape Format	
33	Tape Delivery Date	
34	Tape Length	
35	Tape Manufacturer	
36	Tape Specification	
37	Tape Charge Number	
38	Subtype	e.g. high-band, low-band
Technical Data		
39	S/W / Colour	
40	Roll Counter	Counter if programme is split onto several carriers
41	Technical Ok Flag	
42	Parameters Pictures	
43	Parameters Sound	multiple entries + free text possible
44	Flag Assign as Transmission Tape	
45	MAZ-PASS-ID	Allows to link into the MAZ-PASS system
Reference Data		
46	Department	
47	Production Number	
48	Production Subnumber	
Acquisition Data		
49	Tape of Date of Entry	
50	Tape Origin	
51	Tape Date of Notification	

52	Tape Transfer Necessary Flag	
53	Tape Transfer Done Flag	
54	Tape Responsible Department	
Loan Data		loan-history stays available
55	On Loan Flag	
56	Loan-ID	Links into FARAO-loan-module
57	Secondary Loan Flag	
58	Secondary Loan -ID	Links into FARAO-loan-module
59	Number of Loans	
Storage Data		
60	Storage Number	
61	Old Storage Number	
Other Data		
62	Remarks	Remarks include information about related material, copies, rights, blocking notes, special visual quality, availability of digital preview media (keyframes), technical defects etc.

3.3.2 Conclusion

FARAO is a very old and strictly “material-oriented” model, the elements are very television-specific and some of them may not be used for common audio-visual-data. Although both synthetical and analytical documentation is possible and done, the implementation is rather coarse and not part of the core-model.

3.3.2.1 FARAO entities

- Program
- Program item (“contribution”)

Program items are linked to the overlying program only via the carrier/material (Archive-No.); so documentation is always linked to only ONE instance of material. Links to other carriers/´material is only given by a non-standardized “remark”, mostly on the material with coarse synthetical data, pointing to the fully documented item.

3.3.2.2 Standards and Exchange

Since no standards has been taken into account during development of FARAO (1984) and exchanging data with other systems is not supported or implemented, FARAO is a “paradigm” for legacy systems and isolated applications.

Summarizing the mapping-experiment of the FARAO-DM showed the future needs in handling old legacy data during import into the PrestoSpace-factories as well as the advantages (simple relations, mostly quick mapping solutions, etc.) and disadvantages (poor access to model-structure and technical documentation alike) of mapping an old (and rather simple) legacy model with a elementary modern one. But the core problem will be the linking of descriptive metadata and ALL related material as this is not supported by the FARAO-model.

3.4 IMMIX (B&G)

3.4.1 Introduction

This chapter gives an outline of the metadata model of the iMMix system, the new information system of the Netherlands Institute for Sound and Vision (Beeld en Geluid), which is still being further developed within the iMMix project programme.

The basis of the iMMix system is the catalogue. The legacy catalogues are unsatisfactory because they have not been designed with a relational structure, which means that new indexing methodologies (the new Metadata model of Beeld en Geluid) could not be implemented in the existing systems.

Much effort has been put into translating the Beeld en Geluid vision on the subject of metadata to computerised solutions. A metadata model and functional model have been developed and in order to make the archive material accessible in the right forward-looking way, a new indexing methodology will be implemented as well as a newly built thesaurus which will support both the catalogue input and search processes.

This iMMix system will in the end support both the back office (acquisition, preservation, and indexing processes) and the front office (lending and customer administration). An inventory module and a rights management module will be added or linked later on as well.

Migration from legacy systems to iMMix will take place in several steps. The basic system initially concentrates on handling the flow of new television material from the Dutch public broadcasters. This material will in future be imported directly from the MAM-system that is being developed as part of the Digital Platform infrastructure (the central facility DDV). This cooperation bond between the Dutch public broadcasters and their umbrella organisation Publieke Omroep, NOB and Beeld en Geluid, is realising in several projects a complete digital workflow for television production, transmission and archiving. This system is expected to be operational this summer. The packages to be exported from this central facility to iMMix will consist of high res files, low res files (MPEG1), keyframes, subtitle files and metadata (technical and descriptive) as XML.

3.4.2 Working Processes

The organisational structure of Beeld en Geluid reflects the aim to achieve a product-oriented and customer-oriented organisation, and is designed around the most important processes within Beeld en Geluid. The two distinct functions, namely the operational archive function and the cultural history function, can also be recognised in the processes. Both primary processes are broken down into acquisition, conservation, indexing, and making (information about) archive materials available to different types of customers.

During the acquisition of the collection, the materials are listed, selected, and sometimes provided with short provisional descriptions.

Conservation includes the repository management for various types of collections, the preservation process, and the duplication of materials for internal reasons or external use.

In the indexing process, the formal description and the description of the content is prepared for the catalogue on the basis of guidelines and indexing techniques.

The reuse and delivery of the Beeld en Geluid collections refers to all forms of service and provision of information to internal and external users; not only broadcast professionals, but also i.e. museums, publishers and the educational field.

The following types of data therefore have a role in the primary Beeld en Geluid operational processes.

- Data concerning the materials (the collections)
- Data concerning people and organisations

- Data for indexing the content, that is, the subject
- Data concerning the use of the materials
- Data about planning, production, and selection
- Data for management
- Data needed for executing transactions

3.4.3 The IMMIX Metadatamodel

3.4.3.1 Basic Principles

The iMMix metadata model is open and flexible and thus can be extended, whenever necessary. Some basic principles were defined on how the entire collection should be structured in the new system:

Write once, read many

Different parts of the collections which belong to one production should be identified as such. The metadata that covers those different parts should only be written once. For example in case of a documentary divided into three programmes it is more efficient to fill in the main title, a summary and some index terms that cover every programme in the documentary in stead of doing so per instance.

Inheritance of metadata

Data that is input at higher description levels must be able to 'flow' automatically to all linked, lower description levels and units.

Functional granularity

(Groups of) metadata should be accessed (catalogued, altered, deleted, retrieved) according to the functionality in the process. This basic principle includes the principle of appropriate access: users should have access to only the (groups of) data that they need for the execution of a specified task.

Unique identification

All distinguishable layers and elements of the descriptions must have a unique identifier.

Extensibility

The metadata model must be arranged to be so open and flexible that it can be extended, whenever necessary, with additional data and working processes. The way that the metadata are structured in iMMix should be applicable to:

- o different collections (i.e. museum objects and bibliographic files)
- o different media (radio, television, film, music, internet pages, etc.), as well as
- o different genres (i.e. documentaries, news programmes and fiction series)
- o additional functionality like restoration, preservation and making inventories, rights management and lending
- o new forms of distribution and use of audiovisual material (specifically Internet and multimedia presentations and applications).
- o different purposes; it must be possible to define specific target groups and make different descriptions for each of those groups (see: Intentions).

3.4.3.2 METADATA – means to meet requirements

To meet these basic requirements, next to the metadata model which defines the way the metadata are to be structured, two other means were developed:

- Formats: these define in which way the fields and metadata are to be presented to the documentalist.

- Intentions: these define which metadata should be available to fulfil the specific information needs of a specific target group. The idea is that a broadcast professional prefers objective annotations and for a visitor of the future Beeld en Geluid Media experience i.e. a more catchy description is needed. Also this principle enables the wish to add domain specific and other additional data to a description.

3.4.4 Metadatamodel

The metadata model defines the way the metadata should be structured. It is roughly divided in four stages: concept, actual realisation, physical embodiment, and carrier. Those four stages represent different layers in the model.

Work: the name of the intellectual and artistic concept or idea which is the foundation of one or more realisations.

Realisation: a realisation is an elaboration of a concept: a specific single or multiple productions. Every realisation has a clear structure and form of content compared to other realisations of the same work. In the case of several productions the realisation contains all data that are valid for the underlying expressions.

Series: a series is a group of expressions, usually decided by the makers or producers. A series has a beginning and an ending.

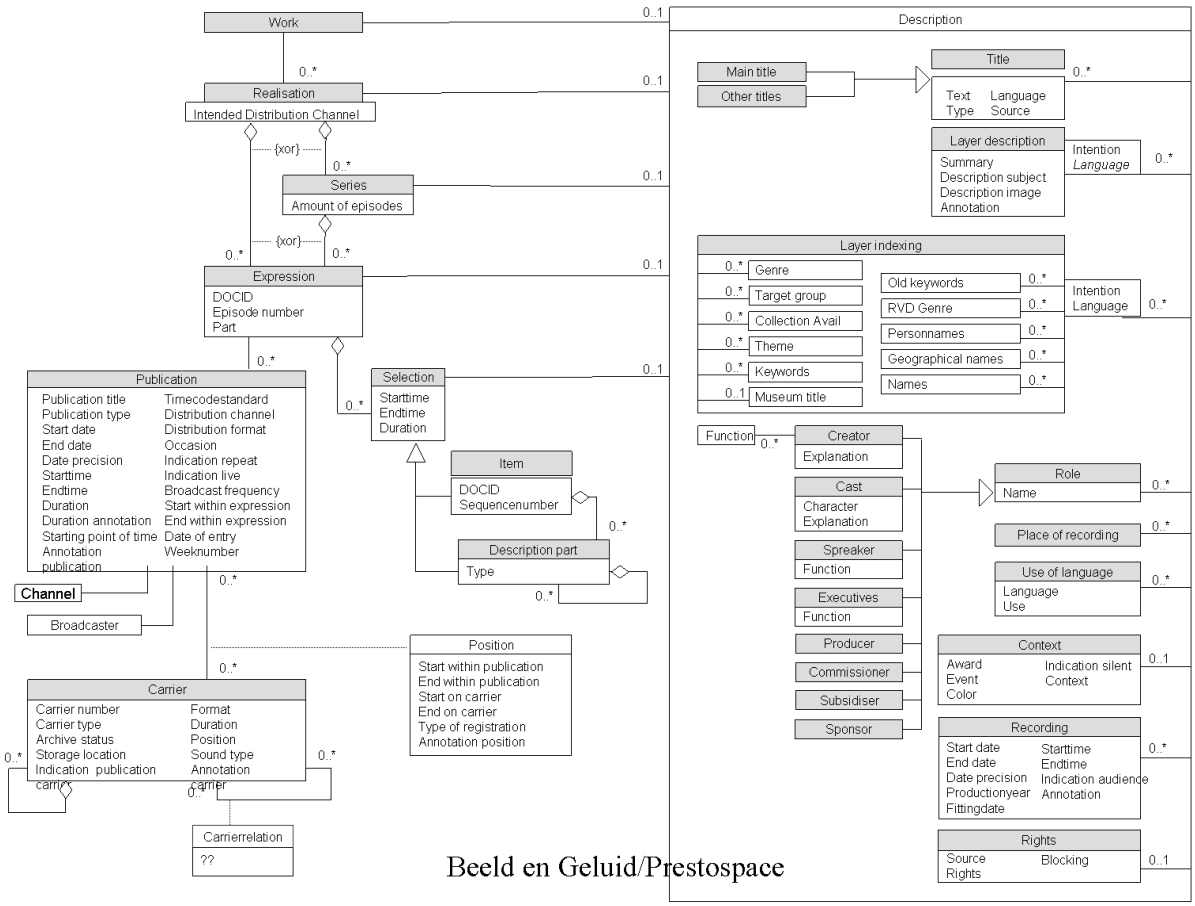
Expression: an expression is a concrete result of a realisation or series. It is the actual production.

Selection: a selection is a part of an expression. For example a news item.

Publication: a publication of an expression; i.e. a television broadcast.

Position: the position of the publication on a carrier.

Carrier: description of a carrier.



UML Diagram of the iMMix metadata model that actually is implemented.

3.4.5 Conclusion

The Beeld en Geluid new metadatamodel as it will be introduced in the next months is a very rich and flexible model which allows for elaborated as well as minimal descriptions and can be easily extended.

3.4.5.1 Standards

The principles as described above and the desire to promote compatibility led to the use of standards within iMMix where possible. For the form and content of the relevant metadata references are i.e. Dublin Core, SMPTE and P-Meta. For modelling the metadata the IFLA-model (International Federation of Library Associations) has been the most important reference and for the exchange format this is XML (also AXF).

3.5 DR Metadata standard

3.5.1 Introduction and Overview

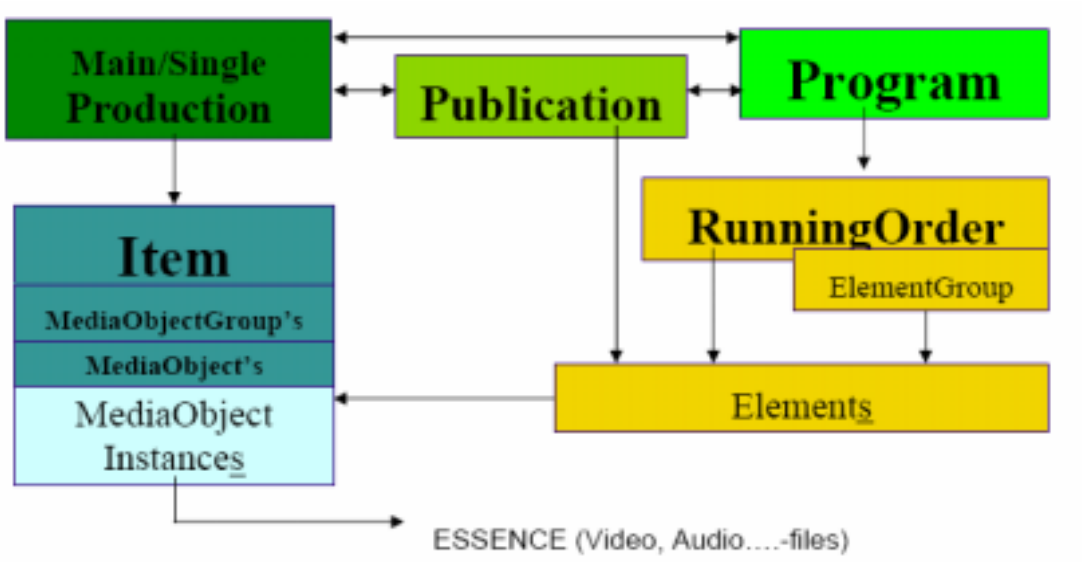
With the implementation of the DR Metadata standard[4] the DR – Danish Broadcasting Corporation[5] wants to improve and facilitate production and storage. For the DR it was very important, that all systems work in concert and that the exchange of information will be transparently.

To achieve this DR defined a data model and standards for system-to-system (S2S), business-to-consumer (B2C) and business-to-business (B2B) exchange.

This metadata schema is internal defined for the DRAMS (DR Asset Management System) and it should also have relations to other international standards. DRAMS should be used for archiving and also for production. The DRAMS specification describes multimedia content like productions, items, programs and articles (contributions).

The DR Metadata model consists of entities, attributes and relations. The DR Metadata standard is intended to serve as a framework for DR asset management with a strong focus on core descriptive metadata. The standard is intentionally restricted and does not cover all system- and implementation specific descriptions. The motivation for these restrictions was to enable cost-effective implementations and allow the use of standard products to the highest possible degree.

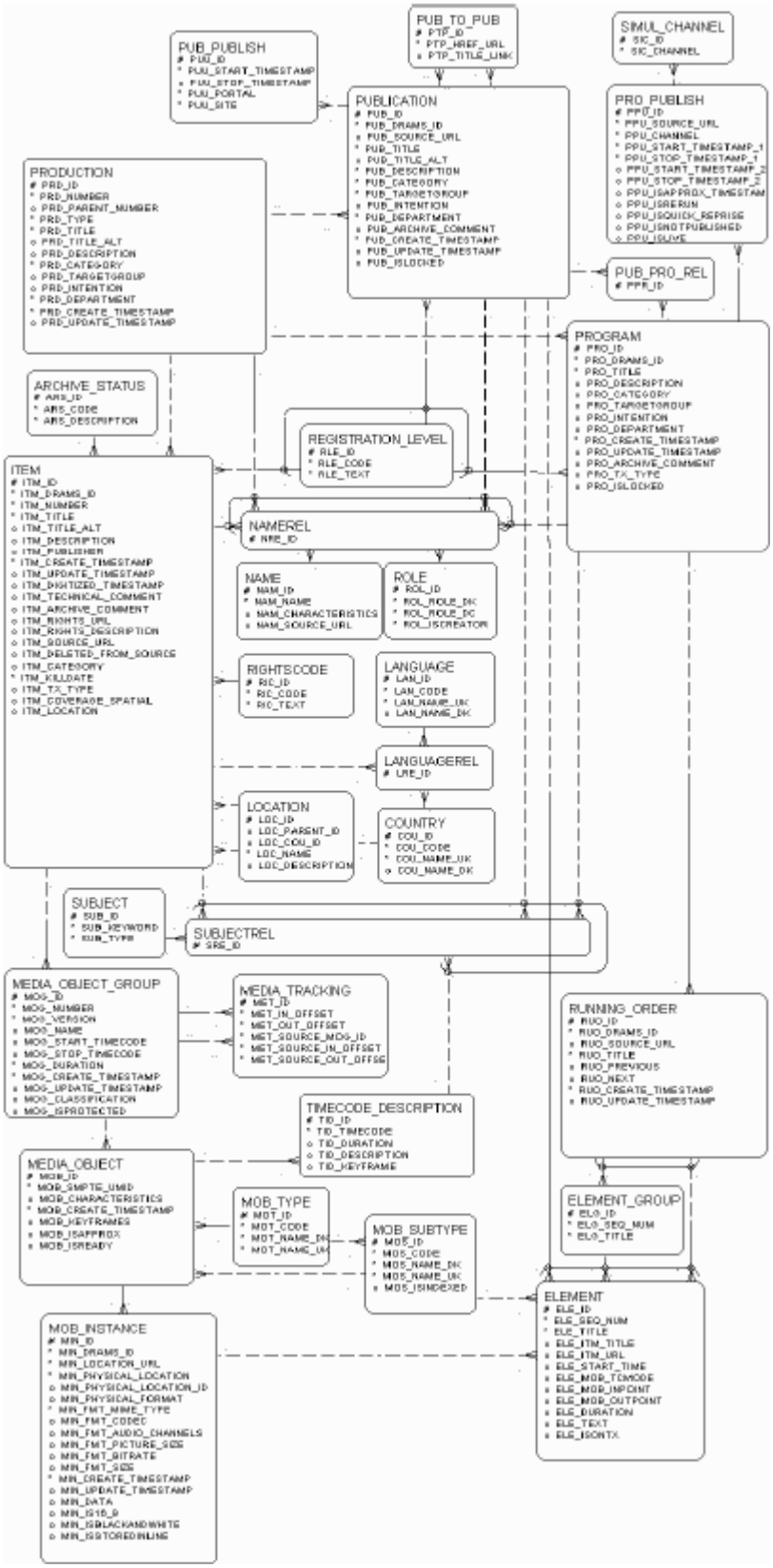
3.5.1.1 The main entity compositions of DR-DM



[4] http://www.dr.dk/omdr/pdf/metadata/DR metadata standard ver12 m logo.pdf

[5] http://www.dr.dk/

3.5.1.2 The logical data model



3.5.1.3 Taxonomies and thesaurus

DR-DM uses these kinds of reference tables:

- based on scheduled configuration update jobs (e.g. channel. department, target-group, etc.
- based on International standards (e.g. country, language, role)
- based on System administrator and/or user update (e.g. location, publisher, rights-code, etc.)
- based on system provider update (MOB[6] types)

3.5.1.4 Elements (Data Dictionary)

Fieldnr.	Fieldname	Description
1	ARCHIVE_STATUS	The Archive_status entity contains a controlled list of codes and descriptions.
1.1	ARC_CODE	Code for status.
1.2	ARC_DESCRIPTION	Description for the meaning of the actual status Code.
2	COUNTRY (COU)	The Country entity contains a controlled list of international country codes and names in Danish and English.
2.1	COU_CODE	International country code.
2.2	COU_NAME_UK	English name for country.
2.3	COU_NAME_DK	Danish name for country.
3	ELEMENT (ELE)	An Element is the smallest part of a Running Order. The instance description contains basic properties such as title, item URL and text.
3.1	ELE_SEQ_NUM	Element Group or Element sequence number.
3.2	ELE_TITLE	Displayed title for element in running order.DC field
3.3	ELE_ITM_TITLE	Copy of referenced Item title. DC field
3.4	ELE_ITM_URL	Reference to an external item. In case the element is imported from a NCS the content could be the MIN_LOCATION_URL.
3.5	ELE_START_TIME	Estimated start time
3.6	ELE_MOB_TCMODE	Meaning of time code can be one of W (wall clock) or Z (zero based).
3.6	ELE_MOB_INPOINT	Time code for in point
3.8	ELE_MOB_OUTPOINT	Time code for outpoint
3.9	ELE_DURATION	ELE_DURATION Estimated or exact duration in seconds
3.10	ELE_TEXT	Text DC field
3.11	ELE_ISONTX	Flag is set if the element is available on the transmission server.
4	ELEMENT_GROUP (ELG)	
4.1	ELG_SEQ_NUM	Element Group or Element sequence number.

[6] Media OBject

	4.2	ELG_TITLE	A common title for several elements. DC field
5		ITEM (ITM)	An item is the smallest unit that carries descriptive metadata (including rights). Each item contains zero or more media object groups.
	5.1	ITM_DRAMS_ID	Unique identifier for a DRAMS item
	5.2	ITM_NUMBER	Item number.
	5.3	ITM_TITLE	Primary item title. DC field
	5.4	ITM_TITLE_ALT	Alternative item title. DC field
	5.5	ITM_DESCRIPTION	An account of the content of the production. DC field
	5.6	ITM_PUBLISHER	Foreign Key to Publisher entity. An entity responsible for making the production available DC field
	5.7	ITM_CREATE_TIMESTAMP	Creation date and time. For imported material the create date refer to the date for formal issuance (e.g. publication) DC field
	5.8	ITM_UPDATE_TIMESTAMP	Update date and time DC field
	5.9	ITM_DIGITIZED_TIMESTAMP	Timestamp for when the material was digitized DC field
	5.10	ITM_TECHNICAL_COMMENT	Technical comments
	5.11	ITM_ARCHIVE_COMMENT	Archive comments
	5.12	ITM_RIGHTS_URL	URL pointing to detailed rights information.
	5.13	ITM_RIGHTS_DESCRIPTION	Should reflect and describe rights restrictions specified in the DR Rights system. DC field
	5.14	ITM_ SOURCE_URL	Unique source identifier for imported Item DC field
	5.15	ITM_DELETED_FROM_SOURCE	This field indicates that the Item is deleted in the source system.
	5.16	ITM_CATEGORY	Foreign Key to Category entity. Category for Item. This field is default inherited from PRODUCTION. DC field
	5.17	ITM_KILLDATE	Date for deletion of item
	5.18	ITM_TX_TYPE	Foreign Key to TX_type entity. Indicates the usability for Item.
	5.19	ITM_COVERAGE_SPATIAL	Foreign Key to Location entity. Location for content (Notre Dame, Paris, France) Location for content is picked from the location reference table. Users shall have the option of adding new locations to this structure. DC field
	5.20	ITM_LOCATION	Foreign Key to Location entity. Location for recording (Studio 4 in DR, Denmark). Location for recording is picked from the location reference table. Users shall have the option of adding new locations to this structure.
6		LANGUAGE (LAN)	The Language entity contains a controlled list of international Language codes and names in Danish and English.
	6.1	LAN_CODE	International language code. DC field.
	6.2	LAN_NAME_UK	English name for language
	6.3	LAN_NAME_DK	Danish name for language
7		LANGUAGEREL (LRE)	Relations between Language, Country and Item are stored in the Language Relation entity.
8		LOCATION (LOC)	The Location entity contains a list of used locations.
	8.1	LOC_COU_ID	Unique location id.
	8.2	LOC_NAME	Name for location.
	8.3	LOC_DESCRIPTION	Description associated to location

9		MEDIA_OBJECT (MOB)	A Media Object defines a piece of media in terms of media type and relations to zero or more instances. All instances defined by the Media Object should be semantically equivalent.
	9.1	MOB_SMPTE_UMID	Unique material identifier
	9.2	MOB_CHARACTERISTICS	Text describing the content
	9.3	MOB_CREATE_TIMESTAMP	Time for creation or recording. Default MOG_CREATE_TIMESTAMP
	9.4	MOB_ISAPPROX	The create timestamp is approximately
	9.5	MOB_ISREADY	Flag indicating the instance(s) is ready for publication.
	9.6	MOB_KEYFRAMES	A collection of automatic generated images (keyframes) associated to video media objects.
10		MEDIA_OBJECT_GROUP (MOG)	The Media Object Group is a container for one or more semantically equivalent Media Objects defined in time by start- and stop time code.
	10.1	MOG_NUMBER	Media object group number.
	10.2	MOG_VERSION	Media object group version
	10.3	MOG_NAME	Name for group like Camera 1, or scene 1
	10.4	MOG_START_TIMECODE	Start time code
	10.5	MOG_STOP_TIMECODE	End time code.
	10.6	MOG_DURATION	Duration
	10.7	MOG_CREATE_TIMESTAMP	Creation time stamp
	10.8	MOG_UPDATE_TIMESTAMP	Update time stamp
	10.9	MOG_CLASSIFICATION	The field can have none or one of the following values - (feed, udsendelseskopi, udsendelseskopi med skilte). The classification is intended to be filled in automatically form the feed scheduling system
	10.10	MOG_ISPROTECTED	Flag indicating that the instance is referenced by an element (program or publication). If this flag is set then it is anticipated that the Item must not be deleted.
11		MOB_INSTANCE (MIN)	The Mob Instance entity represents a reference to a media object instance. The instance description contains basic properties such as date, location, format, status and format-specific information.
	11.1	MIN_DRAMS_ID	Unique identifier for a DRAMS MOB instance.
	11.2	MIN_LOCATION_URL	Locator for this essence file. This can be a reference to a file in a production area or on a remote server. This may also be used as a system-specific identifier for archive systems or similar.
	11.3	MIN_PHYSICAL_LOCATION	Description of physical location. This attribute is only used when storing essence on physical media such as videotapes or optical discs.
	11.4	MIN_PHYSICAL_LOCATION_ID	ID describing identifier on physical media. This attribute is only used when storing essence on physical media such as videotapes or optical discs (i.e. tape number).
	11.5	MIN_PHYSICAL_FORMAT	Foreign Key to Media format entity. Audio/video format for physical media
	11.6	MIN_FMT_MIME_TYPE	MIME type describing the format (file extension)
	11.7	MIN_FMT_CODEC	Foreign Key to Media codec entity. Actual coding format.
	11.8	MIN_FMT_AUDIO_CHANNELS	Audio channel format (1,2,4,5.1)

	11.8	MIN_FMT_PICTURE_SIZE	Picture size pixels.
	11.9	MIN_FMT_BITRATE	Average bit rate for this instance in Bytes/s.
	11.10	MIN_FMT_SIZE	Size of referenced file in Bytes
	11.11	MIN_CREATE_TIMESTAMP	Creation time stamp
	11.12	MIN_UPDATE_TIMESTAMP	Update time stamp
	11.13	MIN_DATA	If Text Media Object is supposed to be indexed then data are stored in MIN_DATA
	11.14	MIN_IS16_9	Flag for 16:9 video
	11.15	MIN_ISBLACKANDWHITE	Flag for black and white video, still and Graphics
	11.16	MIN_ISSTOREDINLINE	Flag telling that text media object is stored in MIN_DATA.
12		MEDIA_TRACKING (MET)	The Media Tracking entity contains linkage between Media Object Groups. This linkage is used to save inheritance and reference information between Media Object Groups.
	12.1	MED_IN_OFFSET	Offset for start location for the source material within the actual video Media Object.
	12.2	MED_OUT_OFFSET	Offset for stop location for the source material within the actual video Media Object.
	12.3	MED_SOURCE_MOG_ID	Reference to source material
	12.4	MED_SOURCE_IN_OFFSET	Offset for start location within the source material
	12.5	MED_SOURCE_OUT_OFFSET	Offset for stop location within the source material
13		MOB_SUBTYPE (MOS)	The MOB Subtype contains a controlled list of media subtypes only related to MOB Types Text and Data.
	13.1	MOS_CODE	Short name for Subtype
	13.2	MOS_NAME_DK	Name for subtype in Danish
	13.3	MOS_NAME_UK	Name for subtype in English
	13.4	MOS_ISINDEXED	Flag indicates that the text shall be indexed
14		MOB_TYPE (MOT)	The MOB Type contains a controlled list of six media types.
	14.1	MOT_CODE	Short name for type: A (Audio), V (Video), S (Still), G (Graphic), D(Data), T (Text)
	14.2	MOT_NAME_UK	Audio, Video, Still, Graphic, Data, Text
	14.3	MOT_NAME_DK	Lyd, Video, Billede, Grafik, Data, Tekst
15		NAMEREL (NRE)	Relations between Names and Roles are stored in the Name Relation entity.
16		NAME (NAM)	The Person entity stores basic person information such as name and characteristics.
	16.1	NAM_NAME	Name for person, group, ensemble etch. DC field
	16.2	NAM_CHARACTERISTICS	Characteristics such as current title and salutation.
	16.3	NAM_SOURCE_URL	A reference to an external catalogue system.
17		PRODUCTION (PRD)	A production may be standalone or related to a main production (parent). A main production defines a group of productions such as a series. The entity description contains basic properties such as DR production number and title.
	17.1	PRD_NUMBER	The DR production number. DC field
	17.2	PRD_PARENT_NUMBER	If this production number exists then the actual production is a child production. Parent number referring to the parent production (main production). DC field

17.3	PRD_TYPE	Foreign Key to Production_type entity. Type as TV, Radio, Web.....
17.4	PRD_TITLE	Primary production title. DC field
17.5	PRD_TITLE_ALT	Alternate production title. DC field
17.6	PRD_DESCRIPTION	Description of the production. DC field
17.7	PRD_CATEGORY	Foreign Key to Category entity. Category for associated programs. DC field
17.8	PRD_TARGETGROUP	Foreign Key to Targetgroup entity. Subcategory to Category
17.9	PRD_INTENTION	Foreign Key to Intention entity. Intention for production (teach, entertain...)
17.10	PRD_DEPARTMENT	Foreign Key to Department entity. Department (descriptive name)
17.11	PRD_CREATE_TIMESTAMP	Creation timestamp DC field
17.12	PRD_UPDATE_TIMESTAMP	Update timestamp DC field
18	PROGRAM (PRO)	Material sequences from productions assembled to one sequence and transmitted as a program. Has metadata such as title, description etc. The Program is guaranteed to represent a single and unique sequence of material (list of Running Orders and Elements).
18.1	PRO_DRAMS_ID	Unique identifier for a DRAMS program.
18.2	PRO_TITLE	The program title. DC field
18.3	PRO_DESCRIPTION	The program description. DC field
18.4	PRO_CATEGORY	Foreign Key to Category entity. Category for associated programs DC field
18.5	PRO_TARGETGROUP	Foreign Key to Targetgroup entity. Subcategory to Category
18.6	PRO_INTENTION	Foreign Key to Intention entity. Intention for production (teach, entertain...)
18.7	PRO_DEPARTMENT	Foreign Key to Department entity. Department (descriptive name)
18.8	PRO_CREATE_TIMESTAMP	Creation timestamp DC field
18.9	PRO_UPDATE_TIMESTAMP	Update timestamp DC field
18.10	PRO_ARCHIVE_COMMENT	Archive comment
18.11	PRO_TX_TYPE	Foreign Key to TX_type entity. Indicates the usability for Program.
18.12	PRO_ISLOCKED	Locked for automatic update
19	PRO_PUBLISH (PPU)	
19.1	PPU_SOURCE_URL	Unique source system transmission identifier
19.2	PPU_CHANNEL	Foreign Key to Channel entity. The broadcast mother channel.
19.3	PPU_START_TIMESTAMP_1	Planned date and time for start of transmission DC field
19.4	PPU_STOP_ TIMESTAMP _1	Planned date and time for end of transmission DC field
19.5	PPU_START_TIMESTAMP_2	Date and time for actual start of transmission DC field
19.6	PPU_STOP_ TIMESTAMP _2	Date and time for actual end of transmission DC field
19.7	PPU_ISAPPROX_ TIMESTAMP	Flag indicates that the start time for transmission is approximately
19.8	PPU_ISRERUN	Rerun flag
19.9	PPU_ISQUICK_REPRISE	Quick reprise flag

	19.10	PPU_ISNOTPUBLISHED	Has not been transmitted flag
	19.11	PPU_ISLIVE	Is transmitted live flag
20		PUB_PRO_REL (PPR)	Relations between Publications and Programs are stored in the pub_pro_rel entity.
21		PUB_PUBLISH (PUU)	Publishing log that contains start-, stop timestamp and location.
	21.1	PUU_START_TIMESTAMP	Timestamp for start of publishing. DC field
	21.2	PUU_STOP_TIMESTAMP	Timestamp for end of publishing. DC field
	21.3	PUU_PORTAL	Foreign Key to Portal entity. Name of portal or high-level address for publication.
	21.4	PUU_SITE	Site name or low-level address for publication
22		PUB_TO_PUB (PTP)	Links between Publications extracted from Publications and mapping to PUB_ID's for internal navigation.
	22.1	PTP_HREF_URL	Original link extracted from Publication.
	22.2	PTP_TITLE_LINK	Name or description related to link.
23		PUBLICATION (PUB)	Material from productions packaged for publishing such as print, web, teletext etc.
	23.1	PUB_DRAMS_ID	Unique identifier for a DRAMS publication. DC field
	23.2	PUB_SOURCE_URL	Unique article source system identifier. DC field
	23.3	PUB_TITLE	Primary article title. DC field
	23.4	PUB_TITLE_ALT	Alternate article title. DC field
	23.5	PUB_DESCRIPTION	Description of the publication. DC field
	23.6	PUB_CATEGORY	Foreign Key to Category entity. Category description. DC field
	23.7	PUB_TARGETGROUP	Foreign Key to Target group entity. Target group (intended audience).
	23.8	PUB_INTENTION	Foreign Key to Intention entity. Intention for publication (teach, entertain...)
	23.9	PUB_DEPARTMENT	Foreign Key to Department entity. Department (descriptive name)
	23.10	PUB_ARCHIVE_COMMENT	Archive comment
	23.11	PUB_CREATE_TIMESTAMP	Creation timestamp.
	23.12	PUB_UPDATE_TIMESTAMP	Modification timestamp.
	23.13	PUB_ISLOCKED	Locked for automatic update
24		REGISTRATION_LEVEL (RLE)	The Registration Level contains a controlled list of registration remarks.
	24.1	RLE_CODE	Short name for registration level
	24.2	RLE_TEXT	Text describing the level of registration.
25		RIGHTSCODE (RIC)	The RightsCode entity contains a controlled list of DR defined codes and text for rights.
	25.1	RIC_CODE	Code for rights.
	25.2	RIC_TEXT	Should reflect rights restrictions specified in the DR Rights system.
26		ROLE (ROL)	The role entity is used to describe roles and role codes and relations to persons through Person Relations.
	26.1	ROL_ROLE_DC	Role identifier.

	26.2	ROL_ROLE_DK	Role name in Danish.
	26.3	ROL_ISCREATOR	Role only applicable for creator attributes.
27		RUNNING_ORDER (RUO)	A Running Order defining a schedule for a Program or a part of a Program and consist of an ordered sequence of Element Groups and/or Elements. Has metadata such as sequence number and title.
	27.1	RUO_DRAMS_ID	Unique identifier for a DRAMS running order. DC field
	27.2	RUO_SOURCE_URL	Unique source system identifier for a Running Order. DC field
	27.3	RUO_TITLE	Title for running order. DC field
	27.4	RUO_NEXT	Indicates situations where one Running Order is followed by another Running Order. Content is the unique RUO_DRAMS_ID for the next Running Order.
	27.5	RUO_PREVIOUS	Indicates situations where one Running Order follows another Running Order. Content is the unique RUO_DRAMS_ID for the previous Running Order.
28		SIMUL_CHANNEL (SIC)	The Simultaneously Channel entity describes child channels associated to one transmission of a Program.
	28.1	SIC_CHANNEL	Short name for transmission channel.
29		SUBJECT (SUB)	The Subject entity contains flat list of keywords that constitutes a controlled vocabulary.
	29.1	SUB_KEYWORD	Keywords
	29.2	SUB_TYPE	Valid keyword types D, T, U D: Keyword is a Danish librarian code (DK5) T: Keyword is part of the thesaurus U: Keyword is not a part of the thesaurus
30		SUBJECTREL (SRE)	Relations between Subject and Item, Program, Publication and Keyframe are stored in the Subject Relation entity.
31		TIMECODE_DESCRIPTION (TID)	Defines a description primarily related to a video Media Object. The entity description contains basic properties such as time code and maybe an image covering the actual keyframe related to timecode.
	31.1	TID_TIMECODE	Time code for when in the video the keyframe is captured.
	31.2	TID_DURATION	Duration for clip
	31.3	TID_DESCRIPTION	Text describing the clip DC field
	31.4	TID_KEYFRAME	Actual keyframe (image)

3.5.2 Conclusion

The DR Metadata model (consisting of entities, attributes and relations) is intended to serve as a framework for DR asset management (with a strong focus on core descriptive metadata.) The internal specifications are for managing material in both the production and archiving domain.

3.5.2.1 The main entities in DR-DM:

- On the "Archive"-side:
  - Main/Single Production
  - Item
  - Media Object Group(s)
  - Media Object (MOB)

- Media Object Instances (MIN)

- On the "Broadcasting/Publishing"-side:
  - Program + Publication
  - Running order
  - Element Group
  - Elements

The connection to the essence is maintained in a single entity, the Media Object Instance (MIN).

3.5.2.2 Standards and Exchange

During development relations to international standards for S2S, B2B and B2C exchange were taken into account (Therefore the mapping-process in chapter 6 showed up the advantages (mostly quick mapping solutions for the core elements) of this approach during development).

For the DR-DM a set of formal specifications and XML-based schemata exists for S2S[7] and B2B[8] exchange.

[7] http://www.dr.dk/omdr/pdf/metadata/appendix2ver12.pdf

[8] http://www.dr.dk/omdr/pdf/metadata/appendix3ver12.pdf

4 Data Models (Sample of Standards)

4.1 Dublin Core

4.1.1 Introduction

Dublin Core is presently the most “popular” and discussed metadata standard. The “Dublin Core Initiative” was founded by the Online Computer Library Center (OCLC)[9] and the National Center for Supercomputing (NCSA)[10] in March 1995. This Initiative always tries to optimize and upgrade the DC.

The Dublin Core Metadata Initiative (DCMI)[11] bases upon an international teamwork of members of the following groups: librarians, representatives of public administration and commercial enterprises, computer scientists. Every year workshops take place where the elements of DC are discussed.

First Dublin Core contained 13, later 15 elements. These elements describe DLOs (document-like objects); these can be specified by “qualifiers”. The qualifiers help the user to find DLOs more easily and quickly in the Internet.

According to *Thea Spiridonidou* only a few search engines are able to read and interpret metadata information at the same time. Those engines which manage to do that, take only certain elements of data into account. For example “Alta Vista” and “Infoseek” take only the elements “Keywords” and “Description” in HTML-documents into consideration.

Spiridonidou assumes that “if Dublin Core becomes a common standard, its elements will be recognized automatically by search engines”.

[9] http://www.oclc.org

[10] http://www.ncsa.uiuc.edu

[11] http://www.dublincore.org

4.1.2 Elements

FieldNr.	Fieldname	Description
1	Title (DC.TITLE)	A name given to the resource. Typically, Title will be a name by which the resource is formally known.
2	Creator (DC.CREATOR)	An entity primarily responsible for making the content of the resource. Examples of Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity.
3	Subject and Keywords (DC.SUBJECT)	A topic of the content of the resource. Typically, Subject will be expressed as keywords, key phrases or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.
4	Description (DC.DESCRIPTION)	An account of the content of the resource. Examples of Description include, but are not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.
5	Publisher (DC.PUBLISHER)	An entity responsible for making the resource available. Examples of Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity.
6	Contributor (DC.CONTRIBUTORS)	An entity responsible for making contributions to the content of the resource. Examples of Contributor include a person, an organization, or a service. Typically, the name of a Contributor should be used to indicate the entity.
7	Date (DC.DATE)	A date of an event in the lifecycle of the resource. Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and includes (among others) dates of the form YYYY-MM-DD.
8	Resource Type (DC.TYPE)	The nature or genre of the content of the resource. Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type Vocabulary [DCT1]). To describe the physical or digital manifestation of the resource, use the FORMAT element.
9	Format (DC.FORMAT)	The physical or digital manifestation of the resource. Typically, Format may include the media-type or dimensions of the resource. Format may be used to identify the software, hardware, or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [MIME] defining computer media formats).

10	Resource Identifier (DC.IDENTIFIER)	An unambiguous reference to the resource within a given context. Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Formal identification systems include but are not limited to the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).
11	Source (DC.SOURCE)	A Reference to a resource from which the present resource is derived. The present resource may be derived from the Source resource in whole or in part. Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.
12	Language (DC.LANGUAGE)	Language Recommended best practice is to use RFC 3066 [RFC3066] which, in conjunction with ISO639 [ISO639]), defines two- and three-letter primary language tags with optional subtags. Examples include "en" or "eng" for English, "akk" for Akkadian", and "en-GB" for English used in the United Kingdom.
13	Relation Resources (DC.RELATION)	A reference to a related resource. Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.
14	Coverage (DC.COVERAGE)	The extent or scope of the content of the resource. Typically, Coverage will include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [TGN]) and to use, where appropriate, named places or time periods in preference to numeric identifiers such as sets of coordinates or date ranges.
15	Rights Management (DC.RIGHTS)	Information about rights held in and over the resource. Typically, Rights will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the Rights element is absent, no assumptions may be made about any rights held in or over the resource.

4.1.3 Conclusion

At a first glance Dublin Core seems to contain all most important elements for the needs importing the essential parts of legacy-metadata into PrestoSpace to be managed and treated there. As a “least common denominator” a “simple” schema like DC may be suited to act as a starting point.

But DC is made of a plain list of attributes, it has no structure. The relations between entities can’t be expressed and it is e.g. not sufficient to produce analytical documentation, even if each item can contain the DC elements.

The simple “list”-nature of DC and the small number of elements were the main arguments for using this schema as the “red thread” in the mapping process in chapter 6.

4.2 SMEF (BBC)

4.2.1 Introduction

SMEF is recently used by the BBC[12] as „Standard Media Exchange Framework Data Model (SMEF-DM[13]). SMEF was developed by the BBC1 Media Data Group (led by Carol Owens) and is BBCs company Data Model. OpenSMEF, a subset of SMEF, was the part released to the public initially. In the beginning of 2001 OpenSMEF was withdrawn in favour of the SMEF-DM v 1.5

"It is a semantic, logical data model defining the meanings of items of data (attributes), of logical clusters of these items (entities), and of the relationships between the clusters." [14]

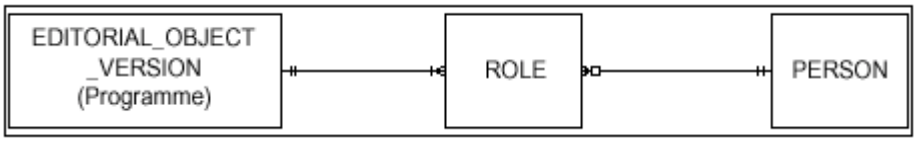
4.2.1.1 SMEF-DM Parts (Products)

The SMEF Data model itself consists (abrasive spoken) of two parts:

- A set of definitions (a Data-Dictionary in which data attributes, and the entities into which they are grouped, are defined (see an example below)

ENTITY: **ROLE_TYPE**
ENTITY DESCRIPTION: Abbreviation – ROT
A description of a standardised type of role, responsibility or task undertaken by an organisation or an individual in the development, management or control of material.
ATTRIBUTE: **ROT_Title**
ATTRIBUTE DESCRIPTION: The title of the type of role. Examples Include: Adaptation By, Book Author, Presenter, Producer, Director, Key Contributor, Production Contact, Material Creator, Cameraman, Vision Mixer, etc

- A set of Entity Relationship Diagrams (ERD) showing the rules for the structural integrity of the data as represented by relationships between the entities (see an example below)



An editorial object is SMEF's name for something that is a complete programme or item (it can also be a promotion or interstitial). Other names for it could be 'a work', episode, etc.

A Media Object is the description of a component of an Editorial Object. It can represent, for example, the audio, video and subtitles. Each Media Object is of a single type, so the video is represented by one Media Object and the subtitles are represented by another.

4.2.1.2 Elements (entities and attributes)

This list shows only the elements and attributes used in OpenSMEF and "FullSMEF" as well. For a full documentation and list please refer to In the original SMEF Architecture it is not foreseen to link (concrete) "Numbers" to the Definitions. But for the mapping-process we assigned fictive numbers ("Nr.") using the following numeration-rule: Entity.Attribute

[12] http://www.bbc.co.uk

[13] http://www.bbc.co.uk/guidelines/smef/

[14] OpenSMEF_v151_Definition_v1.0.pdf by the BBC Media Data Group 28.Feb.2000

Nr.	Entity	Attribute	Description
1	AUDIO_CLIP		Abbreviation - AUD The entity represents an editorial description of a discrete sound or a section of continuous sound that is editorially distinct from any other Audio Clip in a way that is defined in the Media Object description. The sound may be in planning to be captured, edited, or transmitted. All individual audio clip objects are monaural. Stereophonic sound, for example, is created by the association of two sound clips. Examples of Audio Clip include sound track, audio description (An audio description is a verbal description of what is happening in a section of a programme which has no dialogue, and is aimed at viewers with challenged vision.), jingle, etc.
1.1		AUD_Duration	The duration of play of the audio object. Measured in hours, minutes, seconds, and decimal fractions of a second. This is an optional attribute which is only populated where the duration is not derivable. Duration may be derivable, for example, from the difference between the capture start date/time and the capture end date/time.
2	AUDIO_CLIP_ INSTANCE		Abbreviation - AUI This entity holds the data which is specific to a given instance of an audio clip media object. The material instance need not have persistence.
2.1		AUI_Sample_Rate	Specifies the sample rate of essence data.
3	BRAND		Abbreviation - BRA The name applied to a collection of assets with a recognisable collective identity - this could include a series of programmes. The assets could cover programmes, books, videos, characters, magazines, toys, etc. A brand can be defined at a high level as BBC Sport or as a sub-Brand such as Grandstand.
3.1		BRA_Title	The title of the Brand. This could be the same as a programme title, or series, or could be a separately created 'umbrella' title.
4	CONTRACT		Abbreviation - CON A legally binding agreement with a person and/or organisation typically to provide a schedule of deliverables and / or services within a defined time period. A contract will comprise a number of line items.
4.1		CON_Number	The unique identifier allocated to a Contract with a person and/or organisation.

5	EDITORIAL_OBJECT_ VERSION		Abbreviation - EOV This entity identifies specific versions of a particular Editorial Object Concept. These may be, for example, transmittable versions intended for a particular publication event or market (Examples: "The Human Body", Episode 1 (UK); "The Human Body", Episode 1 (USA); the 28' and the 30' minute version of the same programme; pre and post watershed versions of the same programme) The entity has 2 subtypes, Programme Version and Item Version. One or other will be associated with it depending upon whether the editorial concept is seen as a programme in itself or as an item within a programme.
5.1		EOV_Colour_ Indicator	An indicator showing whether an Editorial Object Version is editorially deemed to be in Colour. This is, in the main, required for Programme Guide purposes.
5.2		EOV_Creation_Dat e	The date on which the Editorial Object Version was created.
5.3		EOV_Duration	The duration of an editorial object version measured in hours, minutes, seconds, and decimal fractions of a second.
5.4		EOV_Synopsis_ Description	A description of the content of a version of an Editorial Object. Where only a single version of the Editorial Object Concept exists this should be the same as that entity's equivalent description. The attribute includes comments on what distinguishes this version.
5.5		EOV_Title	The title of a particular version of an Editorial Object. Different versions of an Editorial Object Concept may have different titles.
6	EMAIL_ADDRESS		Abbreviation - EAD Data identifying an electronic mail address for a person, location, or organisation.
6.1		EAD_Name	An electronic mail address.
7	GRAPHIC_TYPE		GRAPHIC_TYPE Abbreviation - GRT Reference data defining the allowable types of graphic media objects. For example: Aston; Graphic Image; Modified Still.
7.1		GRT_Description	Description of the graphic type - e.g.: Aston; modified still; graphic image.
8	IMAGE_CODING_ STANDARD		Abbreviation - ICS This is the coding information applied to the particular instance of the image media object (shot, still, and graphic).
8.1		ICS_Frame_Rate	The rate that images are captured, expressed in frames per second. This, obviously, does not apply to single frame media objects, for example, stills.
8.2		ICS_Name	The colloquial name identifying a particular standardised coding scheme. For example: "PAL"; "NTSC"; "SDV", "HDTV".
8.3		ICS_Total_Lines_ Per_Frame_ Quantity	Specifies the number of lines in a total frame in the video scanning system. (SMPTE Data Dictionary).

9	IMAGE_FORMAT_ TYPE		Abbreviation - IFT Source material for publication can be supplied in a variety of different picture and raster formats. In order to specify this formatting completely, three parameters must be defined: 1 Active Image Aspect Ratio 2 Raster Aspect Ratio 3 Protected Aspect Ratio A labelling system has been devised by BBC Post Production which holds the above information as a six character code as follows: Publication Format Code = aabccd where: aa = Active Image Aspect Ratio. b = Display Format. This can be deduced from the other parameters but it is included explicitly in the code for operational convenience. cc = Raster aspect ratio. d = Protected aspect ratio. Examples: 12F12C 4:3 image full frame on a 4:3 raster
9.1		IFT_Action_ Horizontal_Safe_ Percentage	When a programme is commissioned in widescreen the aspect ratio in which the programme will be transmitted on an analogue service will also be stipulated. In most cases this will be 14:9 letterbox but in some cases (e.g. sport) 4:3 centre cut-out will be requested. In exceptional cases only is a 16:9 deep letterbox image transmitted on the analogue service. Defining the analogue transmission format allows programme makers shooting in 16:9 to limit key elements of the scene to the section of the frame that will be seen by the viewer. This attribute specifies the percentage of the image which will lie outside the vertical safe area for action (cf. graphics). The percentage is of the total raster and not merely of the shoot to protect area and the figure is for both sides of the screen. Thus, a safe area of 3.5% would mean that 7% of the total horizontal portion of the picture is regarded as unsafe.

9.2		IFT_Action_ Vertical_Safe_ Percentage	When a programme is commissioned in widescreen the aspect ratio in which the programme will be transmitted on an analogue service will also be stipulated. In most cases this will be 14:9 letterbox but in some cases (e.g. sport) 4:3 centre cut-out will be requested. In exceptional cases only is a 16:9 deep letterbox image transmitted on the analogue service. Defining the analogue transmission format allows programme makers shooting in 16:9 to limit key elements of the scene to the section of the frame that will be seen by the viewer. This attribute specifies the percentage of the image which will lie outside the vertical safe area for action (cf. graphics). The percentage is of the total raster and not merely of the shoot to protect area and the figure is for both top and bottom of the screen. Thus, a safe area of 3.5% would mean that 7% of the total vertical portion of the picture is regarded as unsafe.
9.3		IFT_Active_Image_ Aspect_Ratio	Defines the Aspect Ratio of the actual picture (excluding any black edges). Expressed as a quantitative relation between two integers, the values being separated by a colon. Examples are: 4:3; 14:9; 15:9; 16:9.
9.4		IFT_Graphics_ Horizontal_Safe_ Percentage	When a programme is commissioned in widescreen the aspect ratio in which the programme will be transmitted on an analogue service will also be stipulated. In most cases this will be 14:9 letterbox but in some cases (e.g. sport) 4:3 centre cut-out will be requested. In exceptional cases only is a 16:9 deep letterbox image transmitted on the analogue service. Defining the analogue transmission format allows programme makers shooting in 16:9 to limit key elements of the scene to the section of the frame that will be seen by the viewer. This attribute specifies the percentage of the image which will lie outside the vertical safe area for graphics (cf. action). The percentage is of the total raster and not merely of the shoot to protect area and the figure is for both sides of the screen. Thus, a safe area of 3.5% would mean that 7% of the total horizontal portion of the picture is regarded as unsafe.

9.5		IFT_Graphics_ Vertical_Safe_ Percentage	When a programme is commissioned in widescreen the aspect ratio in which the programme will be transmitted on an analogue service will also be stipulated. In most cases this will be 14:9 letterbox but in some cases (e.g. sport) 4:3 centre cut-out will be requested. In exceptional cases only is a 16:9 deep letterbox image transmitted on the analogue service. Defining the analogue transmission format allows programme makers shooting in 16:9 to limit key elements of the scene to the section of the frame that will be seen by the viewer. This attribute specifies the percentage of the image which will lie outside the vertical safe area for graphics (cf. action). The percentage is of the total raster and not merely of the shoot to protect area and the figure is for both top and bottom of the screen. Thus, a safe area of 5% would mean that 10% of the total vertical portion of the picture is regarded as unsafe.
9.6		IFT_Raster_ Aspect_Ratio	Defines the Raster Aspect Ratio required to display the picture undistorted. Defines how the picture fits onto the raster, e.g. whether a 16:9 picture is recorded using an entire 4:3 active frame (sometimes called anamorphic), or as a letterbox. Expressed as a quantitative relation between two integers, the values being separated by a colon. Examples are: 4:3; 14:9; 15:9; 16:9.

10	MEDIA_OBJECT		Abbreviation - MOB The editorial information about a single content element. The content element may already have been captured and/or transmitted or it may be in planning to be captured and/or transmitted. A Media Object has content in only one medium - audio, video, text, graphic or still. The differing media have been modelled as subtypes containing their own attributes and relationships. A media object may be created as part of a concept, editorial version, or may be created as a stock item (which may relate to a brief). Thus, the media object may be, for example: the edited audio for a programme; the video for an item; audio for a single clip; a single still from a photograph library; a graphic prepared for a programme; text for transmission with a programme (e.g. subtitles). Distinctions are drawn between: A - the media object itself - because this can exist in the planning stages this can be thought of as the idea behind the shot, still, audio clip, &c. For example, a media object could be the idea of a long shot following an actor as he climbs up a hill. B - the action or event which realises the media object - in the case mentioned above, this would actually be the actor climbing the hill C - the capture of the action or event realising the media object - in the above, this is the sensor (here a camera) used to capture the first instance of the actor climbing the hill.
10.1		MOB_Description	Description of the Media Object. For an Audio Clip object the description will reflect heard contents. This can include details of interview subject matter and background sound effects. For a Graphic Object the description will reflect pictorial contents. This may include subject and composition. For a Still Object the description will reflect pictorial contents. This may include subject and composition. For a Shot Object the description will reflect a synopsis of pictorial contents and may include details of subject and action. It incorporates standard terminology such as wide shot (WS), long shot (LS), two-shot (2-S), etc. For a Text Object the description will reflect contents of the text and could be a summary or synopsis.
10.2		MOB_Title	The title of the media object.

11	ORGANISATION		Abbreviation - ORG The entity represents any recognised grouping of individuals, company or company structure with whom there is an association to carry out tasks or responsibilities for the development, usage or management of media assets. The scope encompasses: an internal organisational unit (e.g. the BBC production department or directorate); a copyright collection body; an independent production company; a grouping of individuals and/or organisations who/which may act together and who should be tracked as such (e.g. one or more named soloists singing or playing with a particular orchestra).
11.1		ORG_Name	The name of an organisation.
12	PERSON		Abbreviation - PER An individual involved in some manner in the creation, use or management of media assets.
12.1		PER_First_Name	The first name of the Person.
12.2		PER_Last_Name	The last name (or surname) of the Person.
12.3		PER_Middle_Name	The middle name, or initial, of the person.
12.4		PER_Salutation_ Short_Form	A shortened salutation as applied to a person's name. Examples are Mr.; Miss; Mrs.; Ms; Lord.
12.5		PER_Suffix_Name	Qualifications, awards, and/or honours normally appended to the end of a person's name. for example: B.Sc.; M.B.; Ch.B.; M.C.; V.C.; O.M.
13	POSTAL_ADDRESS		Abbreviation - PAD Data identifying a postal, address for a person, location, or organisation.
13.1		PAD_Country_ Name	The country name in the postal address.
13.2		PAD_County_ State_Name	The county (or, where appropriate, state or district) name in the postal address for the Location. Examples include Berkshire, Idaho, Manitoba, and Queensland
13.3		PAD_Line_1_Name	First undefined line in the postal address. May be used, for example, to specify a room number or a box number.
13.4		PAD_Line_2_Name	PAD_Line_2_Name Second undefined line in the postal address. May be used, for example, to specify a house number or a building name.
13.5		PAD_Line_3_Name	PAD_Line_3_Name Third undefined line in the postal address. May be used, for example, to specify a road name (or number) or an estate name.
13.6		PAD_Postal_Code	Delivery location code used by the relevant national postal service(s). For example Post Code (UK), ZIP-Code (USA).
13.7		PAD_Town_City_ Name	The town or city name in the postal address.

14	PRG_GROUP_LINK_ PRG_VER		Abbreviation - PGL This entity provides the link to show for any given programme group which programme versions have been included in it, and their episode numbers.
14.1		PGL_Sequence_ Count	The sequence number to denote that a programme version must be transmitted or viewed in a particular order within a series or serial. For example, an episode number.
15	PROGRAMME_ GROUP		Abbreviation - PGR A grouping of programmes with shared identification and branding linked by common characters, subject matter, style or story. Could be a series, serial or themed grouping. A fiction series (drama or comedy) will have common characters, themes and/or style between episodes, but individual stories. A fictional serial will have a common story running across all episodes, with part being told in each. A factual series may have either individual or shared stories/arguments, such as a history series. A series may be occasional or regular in its transmission pattern - a serial will always have a prescribed transmission pattern and order. A themed group may draw together programme versions based around a campaign or anniversary. To BBC Broadcast this may represent a collection of assets for which transmission rights have been acquired.
15.1		PGR_Count	Applies to series and serials only, an incremental number to identify which series this is, where more than one series of programmes has been created.
15.2		PGR_Title	The title for this group of programmes, it may be the same as one or all of the programme versions contained within it. In Broadcast terms this will be the name of the asset package.
16	PROGRAMME_TYPE		PROGRAMME_TYPE Abbreviation - PRT Programme Type is the category of programme type taken from a standardised list for transmission to the consumer. Commonly used in RDS delivery, DAB delivery and MPEG-2 delivery. Programme types include News, Sport, Traffic Information, Pop, Classical, with further sub-categorisation. Also used for EPGs. An example of this within the BBC is the application of the valid combinations of Editorial Category Levels 1, 2, and 3.
16.1		PRT_Title	The title of a specific programme type.

17	PROGRAMME_ VERSION		Abbreviation - PVE A version of a programme is, or may become a unit of transmission with a title and credits and as such may be individually scheduled for transmission. It may be billed in programme guides with a start time and has a pre-determined (possibly approximate) duration. A general definition is as something a viewer or listener would perceive as a "programme". It may be made up of one, or more, items. It may have a budget attached to it, or the budget may apply to a group of programmes. Programmes can contain other programmes - for example: Thought for the Day can be regarded as a programme within the Today programme as it is editorially independent and has a separate budget.
17.1		PVE_Episode_Title	A name given to a particular programme in a group of programmes. For example: a programme in the series 'Panorama' may have an Editorial Object Version Title 'Panorama' and a Programme Version Episode Title 'The BSE Crisis'; a programme in the series 'Dr Who: The Aztecs' may have the Programme Version Episode Title 'Warrior of Death'.
18	RIGHT		Abbreviation - RIG An interest, or permission, which is recognised and protected by law. This entity records the detail of each right which has been acquired for exploitation purposes. The details recorded here are not Outlet specific. The right may relate to an outlet e.g. the rights to transmit a specific programme on a specific outlet, though these details are recorded elsewhere.
18.1		RIG_Condition_ Desc	A description of the conditions which govern or restrict the Right. E.g. -cannot publish a specific media asset during Xmas period.
18.2		RIG_End_Date	The end date of the period covered by this Right, i.e. the date to which the rights can be exploited.
18.3		RIG_Jurisdiction_ Desc	Jurisdiction of law that the Right is covered by.
18.4		RIG_Start_Date	The start date of the period covered by this Right. i.e. the date from which the rights can be exploited.
19	RIGHT_TYPE		Abbreviation - RTY Describes the type of right, for example the right to broadcast or the right to publish.
19.1		RTY_Title	The title of the type of right.
20	ROLE_TYPE		Abbreviation – ROT A description of a standardised type of role, responsibility or task undertaken by an organisation or an individual in the development, management or control of material.
20.1		ROT_Title	The title of the type of role. Examples Include: Adaptation By, Book Author, Presenter, Producer, Director, Key Contributor, Production Contact, Material Creator, Cameraman, Vision Mixer, etc

21	SHOT		Abbreviation - SHO The entity provides the editorial description for a continuous section of an image or images that is editorially distinct from any other Shot in a way that is defined in its description. The section may be planned, captured, created from other recorded images or transmitted. It provides the editorial description of how the image or images are composed - e.g. Wide Screen of..., or Close Up of..., or Starts wide on... and tracks/crabs/zooms/to Medium Close Up of....,etc.
21.1		SHO_Colour_ Indicator	An indicator showing whether the shot is in colour. It is an Editorial Decision whether the shot is deemed to be in colour.
21.2		SHO_Duration	The duration of play of a shot. Measured in hours, minutes, seconds, and frames (where appropriate). This is an optional attribute which is only populated where the duration is not derivable. A duration may be derivable, for example, from start and end timecodes and the frame rate.
22	SHOT_INSTANCE		SHOT_INSTANCE Abbreviation - SHI This entity holds the data which is specific to a given instance of a shot media object. The material instance need not have persistence.
22.1		SHI_File_End_ Timecode	Time code of the last frame in the shot instance.
22.2		SHI_File_Start_ Timecode	Time code of the first frame in the shot instance.
22.3		SHI_Sample_Rate	Specifies the sample rate of essence data - this may be implicit in the shot instance storage format.
23	SOUND_FORMAT_ TYPE		Abbreviation - SFT Defines the type of sound format for an Editorial Object Version and in which it (the Editorial Object Version) may have been commissioned. Examples are: mono; 2 channel stereo; 5 channel surround sound.
23.1		SFT_Name	The name of the sound format type e.g. mono, 2 channel stereo, 5 channel surround sound, etc.
24	STILL		Abbreviation - STI An editorial description of a natural image (c.f. Graphic) with persistence, but no duration e.g. a photo, or single frame extracted from a shot. The description may apply to a still image that is planned to be taken, captured, edited or transmitted. For a Still Object the description will reflect pictorial contents. This may include subject and composition. The description must carefully distinguish this Still from any similar ones.
24.1		STI_Colour_ Indicator	An indicator showing whether the still is in colour. It is an Editorial Decision whether the still is deemed to be in colour.

25	STORAGE_TYPE		Abbreviation - STT This entity defines the types of storage medium. Examples of storage medium include Betacam SP, D3, DVC Pro, VHS, Server, Compact Disc, HDD Server.
25.1		STT_Title	The name by which the storage medium type is known. For example - VHS; ZIP Drive.
26	SUBJ_REF_CONCEPT _TERM		Abbreviation - RCT This entity contains the names which are synonymous with a given concept. It is used to provide both a natural language translation for a classification number and to allow searches on classification numbers in language format.
26.1		RCT_Concept_ Term_Desc	This is a textual value synonymous with a given concept. For example, the concept 4:302.003 (in a particular classification scheme) may be synonymous with both 'MUSEUMS (NATURAL HISTORY)' and with 'NATURAL HISTORY MUSEUMS'.
27	TELEPHONE_ ADDRESS		TELEPHONE_ADDRESS Abbreviation - TAD A telephone number for a person, location, or organisation. The number may be for a telephone or a facsimile machine - any device which can be addressed across a telecoms network in this fashion.
27.1		TAD_Number	A telephone number for a person, location, or organisation.
28	TERRITORY		Abbreviation - TER A defined geographical area. Could be a continent, a group of countries or a region within a country.
28.1		TER_Title	The title of the territory.
29	TEXT_TYPE		Abbreviation - TET Reference data defining the allowable types of text media objects. Examples are Subtitles and Web Text.
29.1		TET_Description	Description of the text type. For example - Sub-titles; WEB Text.

4.2.2 Conclusion

SMEF was a very committed and complex attempt towards the “perfect schema”, going far beyond core TV-business but taken into account future goals. It started in an era when the distress-calls for metadata-standards “fulfilling all needs, situations and business” swept over the Media communities worldwide.

The complex structure, relations and composition of SMEF seems to be the main obstacle against a worldwide success and usage of SMEF. Introducing Open-SMEF and its successors was a good attempt to boost the achievement of SMEF.

SMEF-DM is technology agnostic, which ensures wide applicability. It makes use of standard attributes and definitions where possible, e.g. SMPTE, ISO and EBU, and is not in competition with these organisations but seeks co-operation.

4.2.2.1 The Main Entities of SMEF-DM

Here the following main entities should be mentioned, which have been in use in OpenSMEF as well:

- Brand
- Programme Group
- Programme (Editorial Object Programme)
- Programme Item (Editorial Object Item)

Other entities exist e.g. for Distribution-Channels and related issues, Mediatypes, Contracts and other right-issues, and many, many more. The material/essence is linked to the entity “MEDIA_OBJECT_INSTANCE”(MOI).

4.2.2.2 Standards and Exchange

BBC used SMEF for inputs in other international Taskforces like SMPTE, MPEG7 and MPEG21, but also the EBU-Projects P/FTA, P/FRA,P/CHAIN, etc.

Due to the fact that SMEF was not only meant for “BBC-inhouse”-use, the exchange with other parties/systems is part of the system. The big number of entities and relations may be tricky to be handled when data is exchanged with other systems and the extensive use of manual intervention is to be taken into account.

4.3 MXF – Metadata-Schema

4.3.1 Introduction and Overview

The Material Exchange Format, MXF, is a File Format optimized for the interchange of material for the content creation industries. MXF is a wrapper format intended to encapsulate and accurately describe one or more "clips" of Essence. These Essence "clips" may be Pictures, Sound, Data or some combination of all of these. The core requirement for the design and development of MXF was to be able to bundle the essence and an "EDL" in an unambiguous way which was both essence agnostic and metadata aware. In order for an application to do anything, the file must contain data about the essence i.e. Metadata. This particular sort of metadata is called “structural metadata” and allow many applications and devices to process content without knowing a-priori what the content is. The accurate description of the underlying content is one of the key strengths of the MXF format.

Descriptive Metadata standards define optional editorial metadata that enhance the usability of the Essence content of an MXF file. (The MXF Descriptive Metadata plug-in is very simple. Guidelines on its use can be found in the MXF Engineering Guideline SMPTE EG41. General Guidelines for adding Descriptive metadata can be found in the Engineering Guideline for MXF Descriptive Metadata SMPTE EG42.) The MXF-Descriptive-Metadata-sets are based on a Generic universal label; There may be non-MXF schemes which use the MXF Descriptive Metadata plug-in mechanism. These external schemes may have different label and key values to the ones given here. All MXF defined schemes shall be identified with the Universal Label defined below.

Byte No.	**Description**	**Value (hex)**	**Meaning**
1	Object Identifier	06h	
2	Label size	0Eh	
3	Designator	2Bh	ISO, ORG
4	Designator	34h	SMPTE
5	Registry Category Designator	04h	Labels
6	Registry Designator:	01h	Labels
7	Structure Designator	01h	Labels
8	Version Number	01h	Registry Version
9	Item Designator	0Dh	Organizationally Registered
10	Organization	01h	AAF Association
11	Application	04h	MXF / AAF compatible Descriptive Metadata Labels
12	Label Version	01h	Version 1 of the MXF / AAF DM labels
13	Scheme Kind	xxh	Defined by the scheme specification
14~16	Reserved	00h	Reserved for use by each scheme

An MXF file may contain zero or more Descriptive Metadata (DM) Schemes. These schemes may or may not be MXF DM schemes. A Descriptive Metadata Scheme comprises one or more Descriptive Metadata Frameworks in which each DM Framework has several metadata items and sets grouped together (generally for semantic reasons).

For more information about MXF and the Metadata handling therein please refer to SMPTE S377M[15].

4.3.1.1 DMS1 (official and standardized MXF-DM)

The official SMPTE-standardized Metadata-schema for MXF is DMS1 (S380M[16]) (Descriptive Metadata Scheme), published in fall 2004.

In DMS1 three frameworks are defined:

- Production Framework
  - descriptive Metadata containing information about content and identification
- Scene Framework
  - descriptive Metadata containing information about actions, events and places occurring within the content
- Clip Framework
  - descriptive Metadata containing information about the creation of the content (shooting information)

4.3.1.2 An example for a non-standardized MXF-DM

The following DM-schema for MXF (according to the State University of Florida[17]) is based upon Dublin Core, but as well derived from the AAF-Data-Model, which is even more complex and targeting multimedia-authoring and post-production. In this paper this DM-schema (elements see below) has been used for the mapping process.

15 http://www.smpte.org/smpte_store/standards/

16 http://www.smpte.org/smpte_store/standards/index.cfm?scope=0&CurrentPage=18&stdtype=smpte

17 http://palmm.fcla.edu/strucmeta/mxf2mods.htm

- **Elements**

FieldNr.	Fieldname	Description
1	dc.Title	A name given to the resource. Typically, Title will be a name by which the resource is formally known.
2	dc.Creator	An entity primarily responsible for making the content of the resource. Examples of Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity.
3	dc.Subject	A topic of the content of the resource. Typically, Subject will be expressed as keywords, key phrases or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.
4	dc.Description	An account of the content of the resource. Examples of Description include, but are not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.
5	dc.Publisher	An entity responsible for making the resource available. Examples of Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity.
6	dc.Contributor	An entity responsible for making contributions to the content of the resource. Examples of Contributor include a person, an organization, or a service. Typically, the name of a Contributor should be used to indicate the entity.
7	dc.Date	A date of an event in the lifecycle of the resource. Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and includes (among others) dates of the form YYYY-MM-DD.
8	dc.Type	The nature or genre of the content of the resource. Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type Vocabulary [DCT1]). To describe the physical or digital manifestation of the resource, use the FORMAT element.
9	dc.Format	The physical or digital manifestation of the resource. Typically, Format may include the media-type or dimensions of the resource. Format may be used to identify the software, hardware, or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [MIME] defining computer media formats).
10	dc.Identifier	An unambiguous reference to the resource within a given context. Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Formal identification systems include but are not limited to the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).
11	dc.Source	A Reference to a resource from which the present resource is derived. The present resource may be derived from the Source resource in whole or in part. Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.

12	dc.Language	A language of the intellectual content of the resource. Recommended best practice is to use RFC 3066 [RFC3066] which, in conjunction with ISO639 [ISO639]), defines two- and three-letter primary language tags with optional subtags. Examples include "en" or "eng" for English, "akk" for Akkadian", and "en-GB" for English used in the United Kingdom.
13	dc.Relation	A reference to a related resource. Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.
14	dc.Coverage	The extent or scope of the content of the resource. Typically, Coverage will include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [TGN]) and to use, where appropriate, named places or time periods in preference to numeric identifiers such as sets of coordinates or date ranges.
15	dc.Rights	Information about rights held in and over the resource. Typically, Rights will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the Rights element is absent, no assumptions may be made about any rights held in or over the resource.
Qualified DC		
16	dc.title.Alternative	Any form of the title used as a substitute or alternative to the formal title of the resource. This qualifier can include Title abbreviations as well as translations.
17	dc.description.TableOfContents	A list of subunits of the content of the resource.
18	dc.contributor.role	
19	dc.format.extent	The size or duration of the resource.
20	dc.format.medium	The material or physical carrier of the resource.
21	dc.coverage.temporal	Temporal characteristics of the intellectual content of the resource.
22	dc.date.created	Date of creation of the resource.
23	dc.date.issued	Date of formal issuance (e.g., publication) of the resource.
24	dc.date.modified	Date on which the resource was changed.
25	dc.relation.hasPartOf	The described resource includes the referenced resource either physically or logically.
26	dc.relation.isVersionOf	The described resource is a version, edition, or adaptation of the referenced resource. Changes in version imply substantive changes in content rather than differences in format.
27	dc.relation.hasVersion	The described resource has a version, edition, or adaptation, namely, the referenced resource.
28	dc.relation.isFormatOf	The described resource is the same intellectual content of the referenced resource, but presented in another format.
29	dc.relation.hasFormat	The described resource pre-existed the referenced resource, which is essentially the same intellectual content presented in another format.
30	dc.subject	
31	dc.description.note	Additional information regarding the thesis or dissertation. Example: acceptance note of the department

32	dc.description.release	Description of the version of the work.

4.3.2 Conclusion

The official Data-model for MXF (DMS1) is relatively young, so when the work on this paper started only a short overview was planned to be given as it wasn't official published then. The MXF-format itself could be an important candidate for the transportation-bin for AV-items within PrestoSpace or from/to the customers, so the data-model within this wrapper-format should be taken into consideration as an important input for the PrestoSpace-Metadata-schema. Therefore we started already to analyze DMS1 more deeply and may be ready to provide an annex to this paper, containing more precise information on DMS1 later this year.

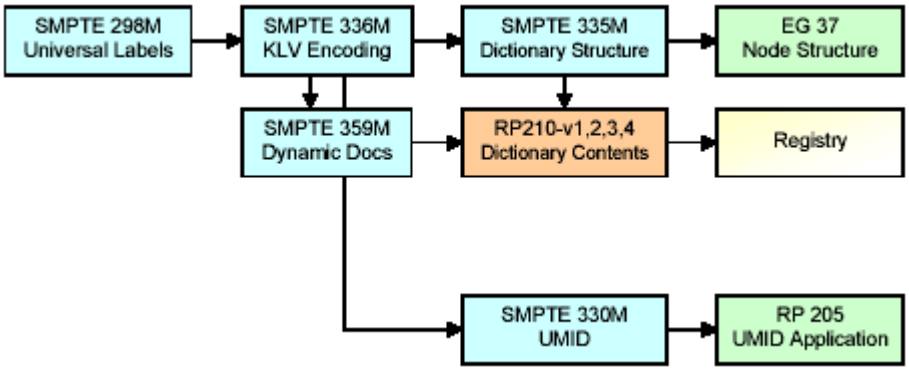
The non-official schema shortly introduced suffers from the same disadvantages as it's source (plain list of attributes, no structure, relations between entities can't be expressed, etc.). It's lack of being standardized is an important disadvantage too. (We used it for the mapping process in chapter 6 just due to the fact that is could be done quick and without special efforts.)

4.4 SMPTE Metadata Dictionary and structure

4.4.1 Introduction

SMPTE has been developing standards for Metadata since before the term was in common use (e.g. SMPTE 12M Timecode). The need for a registry of Metadata was perceived before the inception of the EBU/SMPTE Taskforce, but was put on hold pending the publication of the influential task force report in 1998. Since then, progress towards delivery of the full SMPTE Registry has been deliberate and in 2001 SMPTE Metadata Dictionary structure was officially published as standard. The Metadata Dictionary itself has been updated and revised several times, the latest version is RP210.8[18] from the 10th Aug. 2004.

The following table shows the relations between the different SMPTE-standards when the Metadata Dictionary was published:



"The SMPTE Metadata Dictionary picks up on the work of the EBU-SMPTE Taskforce for Harmonized Standards for the Exchange of Programme Material as Bit streams that completed its Final Report in 1998. The Dictionary provides flexibility in capturing metadata and exchanging it among applications through a standardized hierarchy of Universal Labels for the metadata-elements, grouped in classes. "Metadata Classes" are collections of metadata elements with common characteristics or attributes. Additional Classes are provided for user-defined metadata. The Dictionary also contains information on the required format of metadata values and the allowable range of values. The Dictionary consist of Structure (SMPTE XXXX) and Content (RPXXXX) which must be used together as a pair-neither must be used in isolation. The SMPTE Dictionary was recently awarded during NAB 2000 in Las Vegas."[19]

We do not provide the full information (list of elements) about this metadata-dictionary (SMPTE RP210) because it would boost the volume of this paper way too much. It contains about 1800 elements and therefore would take over 50 pages in this paper.

For more information on the elements please see:

http://www.smpte-ra.org/mdd/RP210v8-final-040810MC.xls

The metadata dictionary structure defined in SMPTE 335M[20] covers the use of metadata for all types of essence (video, audio, and data in their various forms). The dictionary itself must not be used alone as it's meant to be the practical part to SMPTE 335M.

---

[18] http://www.smpte-ra.org/mdd/rp210-8.pdf

[19] http://www.schemas-forum.org/registry/desire/activityreports.php3?field=filename&value=AV-smpte.rdf

[20] http://www.smpte.org/smpte_store/standards/index.cfm?scope=0&stdtype=smpte&CurrentPage=15

4.4.2 Conclusion

As the name states, this is a dictionary, not a schema or standard. So its usefulness can be that of providing standard keys and definitions for attributes to be implemented in planned to be developed schemata and standards. The dictionary is structured in 7 distinct classes:

- Identification
- Administration
- Interpretation
- Parametric
- Process
- Relational
- Spatio-Temporal

Introducing 457 nodes and 1363 leafs the dictionary provides elements for nearly every occasion and need in the A/V-world, acting perfectly as a rich source (“repository”) for every DM-developing taskforce.

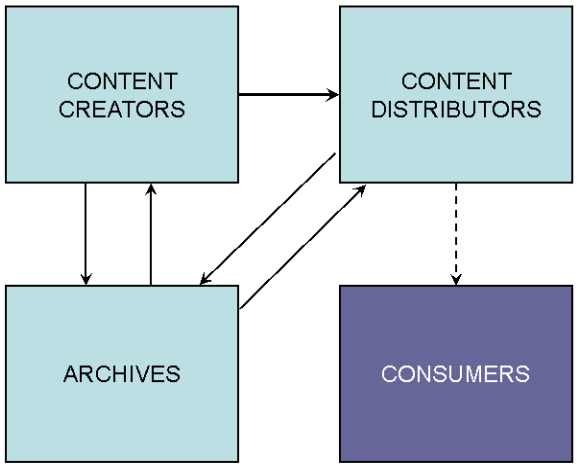
For the mapping-process we used the numeration provided in the SMPTE-list we gave the link to. Due to the fact that the mapping of the SMPTE-elements was the most controversial and time-consuming, going far beyond the timeframe provided for this paper, we decided to stop the process for the time being (only 1 element, UID, enclosed in mapping, acting as an example) and will proceed in case of exigency later in the project.

The SMPTE Metadata Dictionary is to be taken into account for the process of designation elements for the PS-DM-schema, acting potentially as a tool and “repository” in the definition-process.

4.5 P_META Metadata Exchange Scheme(EBU)

4.5.1 Introduction and Overview

P_Meta (v1.0) was originated 1999 by the EBU[21] Project Group to create a standardised metadata exchange scheme which offers a way of sharing the meaning of electronic information necessary or useful for the business-to-business exchange of content.. The P_META Scheme is intended for use in a business-to-business scenario where the participating organisations may retain their internal data structures, workflows, and concepts. The P_META process-model:



The P_META Scheme is basically a set of definitions which provide a semantic framework for the information which is typically exchanged along with audio-visual material. It includes the identification of concepts (simple or complex) that are referenced by P_META names and P_META Identifiers.

“As a key principle, the data analysis has been set at the lowest level to identify concepts or subjects which cannot be further divided, thus giving maximum precision in meaning, with maximum flexibility in the use and re-use of basic elements.” The list of main entities of P_Meta, built on the question “What kind of content will be the object of an information exchange ?”, contains the following:

- Programme
- Programme Group
- Item / Programme Item
- Media Object (MOB)
- Brand

4.5.1.1 Attribute List

“As the list is in itself unstructured, P_Meta also contains a number of sets (70) that enable the attributes to pass specific types of information by putting them into a specific context. It is also envisaged that groups of P_Meta users will also be able to build additional sets for their

[21] http://www.ebu.ch/en/index.php

own purposes by using a defined notation to combine attributes from the flat list into meaningful sets."[22]

The following list does not contain all attributes in P_Meta because of the great amount of them. For the complete list and deeper information please refer to:

http://www.ebu.ch/en/technical/publications/tech3000_series/tech3295/index.php?display=EN

CODE	Attribute-Name	Description
A1	ADDRESS_DELIVERY_CODE	Delivery location code used by the relevant national postal service(s). The nature and format of the code varies considerably between different countries. Note that some countries now include a character ahead of their internal postal code to indicate the country. This may be seen in the examples from Germany, Sweden and Norway below. Known aliases: Post Code, ZIP Code, Value Type: External Source of referenced data: National Postal Service of each country. Examples: (country indicated within parenthesis) W1A 1AA (United Kingdom); SE23 3NL (United Kingdom); 10135 (Italy); SE-105 (Sweden); D-80939 (Germany); N-0340 (Norway); 55101 (USA); R3N 1S8 (Canada)
A2	ADDRESS_LINE_NAME	An undefined line in the postal address. This might be a street name, a house name, a house number, a district of a city, etc. Value Type: Uncontrolled Text Examples: No. 45; Crowthorn Road; Ocean Boulevard; Rue de Montreal
A3	ADDRESS_TELEPHONE_NUMBER	A telephone number, the precise nature and/or context of which will be defined by the set containing the number. Value Type: Uncontrolled Text Examples: 020 8534 1234
A4	ADDRESS_TOWN_CITY_NAME	The town or city in a postal address Value Type: Uncontrolled Text Examples: Paris; Glasgow; Minneapolis
A9	AWARD_NAME	The name of an award presented to a person, programme, series, etc. and signifying public approbation, professional achievement, etc. Value Type: Uncontrolled Text Examples: Silver Rose of Montreaux.
A10	BRAND_TITLE	Title of a Brand, which is a collection of assets with a recognisable collective identity. Value Type: Uncontrolled Text Examples: BBC Drama; BBC Sport.

[22] http://www.ebu.ch/CMSimages/en/tec_t3295_tcm6-11463.pdf

A11	SERVICE_NAME	The name of a broadcast service. A broadcast service is a distinctive packaging of published (broadcast) material intended to provide an identifiable continuity to the programming. Identity is usually achieved through a style of presentation, visually by logos, graphics and images and/or aurally by 'jingles' pervading the packaging. Value Type: Uncontrolled Text Examples: BBC ONE; BBC Radio Four; Radio Essex; Jazz FM; RAI 1; Österreich-1;
A12	DAI_COMPRESSION_CODE	A controlled code for the type of processing applied to a data object in order to achieve data compression, but also file grouping and data encoding. The abbreviation DAI stands for Data object Instance. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.6 DAI_COMPRESSION_CODE
A13	DAI_COMPRESSION_NAME	A textual name for the type of processing applied to a data object in order to achieve data compression, but also file grouping and data encoding. The abbreviation DAI stands for Data object Instance. Value Type: Uncontrolled Text Examples: gzip
A14	CONTRACT_CLAUSE_DESCRIPTION	A description given to a clause within a contract Value Type: Uncontrolled Text
A15	CONTRACT_DATE	The date upon which a contract is made. Value Type: Formatted Code according to ISO 8601 complete date (YYYY-MM-DD) Examples: 2002-06-30
A18	CONTRACT_NUMBER	Number assigned to the contract for the transaction by the issuer of the contract for the purposes of identifying the contract. Value Type: Uncontrolled Text Examples: LB00313201/1; LB00313236/01
A19	CONTRACT_TYPE_CODE	A controlled code identifying the type of contract. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.5 CONTRACT_TYPE_CODE
A20	CONTRACT_TYPE_NAME	Describes a contract type recognized between people and/or organizations involved in the lifecycle of media-assets Value Type: Uncontrolled Text
A21	COUNTRY_CODE	Internationally agreed code for code for a specified country. Value Type: External Source of referenced data: ISO 3166-1: Examples: fr; gb; hu
A22	COUNTRY_NAME	Name of a country. Value Type: Uncontrolled Text Examples: France, Germany, Greece, Italy, Spain, United Kingdom.

A23	CURRENCY_CODE	An internationally agreed code for a currency. Value Type: External Source of referenced data: ISO 4217: Examples: frf; gbp;huf
A24	CURRENCY_NAME	The name of an internationally recognised currency Value Type: Uncontrolled Text Examples: Euro, Pound Sterling, US Dollar
A25	DEVICE_NUMBER	The identifier allocated by an organisation to a device. This is additional to the manufacturer's serial number. Value Type: Uncontrolled Text
A26	DEVICE_TYPE_CODE	A controlled code for a device type. Device types are defined to the level of identifying the technology used by the device without going down to the level of a complete technical specification of an individual device. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.7 DEVICE_TYPE_CODE
A27	DEVICE_TYPE_NAME	The name of the type of device e.g. camera, audio tape machine, digital linear tape, random access memory (RAM), etc. Device types are defined to the level of identifying the technology used by the device without going down to the level of a complete technical specification of an individual device. Value Type: Uncontrolled Text Examples: Audio Mixer, Microphone, Router, Tape Recorder, Minidisk Recorder
A28	ADDRESS_ELECTRONIC_NAME	An electronic mail address. External reference: OpenSMEF Attribute Value Type: Uncontrolled Text Examples: john.jordan@bbc.co.uk
A29	EOV_DURATION	The duration of a particular version of a programme or item, at the frame rate it is intended to be shown at, measured in hours, minutes, seconds, and decimal fractions of a second. This may be derivable from the duration of its constituent media objects. The abbreviation EOV stands for Editorial Object Version. External reference: OpenSMEF Attribute Value Type: Formatted Code according to ISO 8601, with three digits for representing the decimal fractions of a second (hh:mm:ss.sss) Examples: 01:15:30.500
A30	EVENT_END_DATE	The date (GMT/UTC) at the end of an event specified by the set containing the attribute. Value Type: Formatted Code according to ISO 8601 complete date (YYYY-MM-DD) Examples: 2002-06-30

A31	EVENT_END_ELAPSED_TIME	The elapsed time from the beginning of the programme/item/media object to the end of an event specified by the set containing the attribute. Value Type: Formatted Code according to ISO 8601, with three digits for representing the decimal fractions of a second (hh:mm:ss.sss) Examples: 01:15:30.500
A32	EVENT_END_TIME	The time (GMT/UTC) at the end of an event specified by the set containing the attribute. Value Type: Formatted Code according to ISO 8601, with three digits for representing the decimal fractions of a second (hh:mm:ss.sss) Examples: 23:40:01.500
A33	EVENT_END_TIMECODE	The timecode at the end of an event specified by the set containing the attribute. External reference: Format defined by SMPTE Value Type: Formatted Code
A34	EVENT_START_DATE	The date (GMT/UTC) at the beginning of an event specified by the set containing the attribute. Value Type: Formatted Code according to ISO 8601 complete date (YYYY-MM-DD) Examples: 2002-06-30
A35	EVENT_START_ELAPSED_TIME	The elapsed time from the beginning of the programme/item/media object to the beginning of an event specified by the set containing the attribute. Value Type: Formatted Code according to ISO 8601, with three digits for representing the decimal fractions of a second (hh:mm:ss.sss) Examples: 01:15:30.500
A36	EVENT_START_TIME	The time (GMT/UTC) at the beginning of an event specified by the set containing the attribute. Value Type: Formatted Code according to ISO 8601, with three digits for representing the decimal fractions of a second (hh:mm:ss.sss) Examples: 23:40:01.500
A37	EVENT_START_TIMECODE	The timecode at the beginning of an event specified by the set containing the attribute. External reference: Format defined by SMPTE Value Type: Formatted Code
A38	FESTIVAL_NAME	The name of a festival or award ceremony Value Type: Uncontrolled Text Examples: Festival de Cannes, The Academy Awards, Mostra del Cinema di Venezia.
A40	GRAPHIC_USAGE_TYPE_CODE	A controlled code for the type of usage to which a graphic is put. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.11 GRAPHIC_USAGE_TYPE_CODE

A41	PCS_ACTIVE_LINES_PER_FRAME_QTY	The number of lines within the television frame available to carry image information. The abbreviation PCS stands for Picture Coding Scheme. Value Type: Integer
A42	PCS_ACTIVE_SAMPLES_PER_LINE_QTY	The number of samples within the television line available to carry image information. The abbreviation PCS stands for Picture Coding Scheme. Value Type: Integer Examples: 720
A43	PCS_BITS_PER_PIXEL_QUANTITY	Number of bits in each sample describing each pixel component. The abbreviation PCS stands for Picture Coding Scheme. Value Type: Integer
A44	PCS_FRAME_RATE_QUANTITY	The rate at which the material should be shown in order to achieve the intended editorial effect - expressed in frames per second. This does not apply to single frame media objects, for example, stills. The abbreviation PCS stands for Picture Coding Scheme. Value Type: Floating Point Examples: 25
A46	PCS_SAMPLING_HIERARCHY_CODE	A code that specifies the component sampling hierarchy for the video pixel matrix (SMPTE Data Dictionary). The abbreviation PCS stands for Picture Coding Scheme. External reference: SMPTE MDD Known aliases: chroma_format Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.22 PCS_SAMPLING_HIERARCHY_CODE
A47	PCS_SAMPLING_STRUCTURE_CODE	A code that specifies the analogue or digital sampling structure for the video scanning system. For example: Interlaced (SMPTE Data Dictionary). The abbreviation PCS stands for Picture Coding Scheme. External reference: SMPTE MDD Known aliases: interlaced/progressive flag Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.23 PCS_SAMPLING_STRUCTURE_CODE
A48	PCS_TOTAL_LINES_PER_FRAME_QTY	Specifies the number of lines in a total frame in the video scanning system. (SMPTE Data Dictionary). The abbreviation PCS stands for Picture Coding Scheme. External reference: SMPTE MDD Value Type: Uncontrolled Text Examples: 625

A49	PCS_TOTAL_SAMPLES_PER_LINE_ QTY	Specifies the number of samples in a total line in the video pixel matrix (SMPTE Data Dictionary). The abbreviation PCS stands for Picture Coding Scheme. Value Type: Uncontrolled Text
A50	PFT_ACTION_VERTICAL_ SAFE_PERCENTAGE	When a programme is commissioned in widescreen the aspect ratio in which the programme will be transmitted on an analogue service will also be stipulated. In most cases this will be 14:9 letterbox but in some cases (e.g. sport) 4:3 centre cut-out will be requested. In exceptional cases only is a 16:9 deep letterbox image transmitted on the analogue service. Defining the analogue transmission format allows programme makers shooting in 16:9 to limit key elements of the scene to the section of the frame that will be seen by the viewer. This attribute specifies the percentage of the image which will lie outside the vertical safe area for action (cf. graphics). The percentage is of the total raster and not merely of the shoot to protect area and the figure is for both top and bottom of the screen. Thus, a safe area of 3.5% would mean that 7% of the total vertical portion of the picture is regarded as unsafe. The abbreviation PFT stands for Picture Format Type. External reference: OpenSMEF Attribute Value Type: Floating Point Examples: 3.5
A51	PFT_ORIGINAL_FRAMING_ ASPECT_RATIO	The aspect ratio of the shot as it is framed for capture and, thus, the aspect ratio at which the image must be displayed to avoid distortion. Expressed as a quantitative relation between two integers, the values being separated by a colon. Examples are: 4:3; 14:9; 15:9; 16:9. The abbreviation PFT stands for Picture Format Type. Known aliases: Capture aspect ratio (SMPTE) Value Type: Formatted Code Examples: 4:3; 16:9
A53	PFT_GRAPHICS_HORIZONTAL_SAF E_ PERCENTAGE	When a programme is commissioned in widescreen the aspect ratio in which the programme will be transmitted on an analogue service will also be stipulated. In most cases this will be 14:9 letterbox but in some cases (e.g. sport) 4:3 centre cut-out will be requested. In exceptional cases only is a 16:9 deep letterbox image transmitted on the analogue service. Defining the analogue transmission format allows programme makers shooting in 16:9 to limit key elements of the scene to the section of the frame that will be seen by the viewer. This attribute specifies the percentage of the image which will lie outside the vertical safe area for graphics (cf. action). The percentage is of the total raster and not merely of the shoot to protect area and the figure is for both sides of the screen. Thus, a safe area of 3.5% would mean that 7% of the total horizontal portion of the picture is regarded as unsafe. The abbreviation PFT stands for Picture Format Type. External reference: OpenSMEF Attribute Value Type: Floating Point Examples: 10

A54	PFT_GRAPHICS_VERTICAL_ SAFE_PERCENTAGE	When a programme is commissioned in widescreen the aspect ratio in which the programme will be transmitted on an analogue service will also be stipulated. In most cases this will be 14:9 letterbox but in some cases (e.g. sport) 4:3 centre cut-out will be requested. In exceptional cases only is a 16:9 deep letterbox image transmitted on the analogue service. Defining the analogue transmission format allows programme makers shooting in 16:9 to limit key elements of the scene to the section of the frame that will be seen by the viewer. This attribute specifies the percentage of the image which will lie outside the vertical safe area for graphics (cf. action). The percentage is of the total raster and not merely of the shoot to protect area and the figure is for both top and bottom of the screen. Thus, a safe area of 5% would mean that 10% of the total vertical portion of the picture is regarded as unsafe. The abbreviation PFT stands for Picture Format Type. External reference: OpenSMEF Attribute Value Type: Floating Point Examples: 5
A56	INDEX_KEYWORD_NAME	Important word or phrase from a title, a document, the material description, or other used as an index to content. Value Type: Uncontrolled Text Examples: Louis XIIII; Economy; Brasil; Health Service; Defence.
A59	ITEM_SUB_TITLE	The sub-title given to an item (constituent editorial part of a programme) by its producer. Value Type: Uncontrolled Text
A60	ITEM_SYNOPSIS_DESCRIPTION	A synopsis of an item (constituent editorial part of a programme). Value Type: Uncontrolled Text
A61	ITEM_TITLE	The title given to an item (constituent editorial part of a programme) by its producer. Value Type: Uncontrolled Text Examples: Washing machine item; Steve McQueen interview; Verdi Aria.
A62	ITEM_SEQUENCE_NUMBER	An integer indicating where, in the sequence of items forming a programme, a particular item is to be found. Value Type: Integer
A65	LANGUAGE_CODE	Internationally agreed code for a specified language. Value Type: External Source of referenced data: ISO 639 (both version 1 - two letters codes and version 2 three letters codes): Examples: en; ha; gd; fr; eng; ita; deu; fra
A66	LANGUAGE_NAME	The name of a language. Value Type: Uncontrolled Text Examples: Anglais; American English; Hausa

A67	LOCATION_CAPTURE_NAME	The name of the location of the sensor (e.g. camera) used in the production of content. For example, if St Paul's Cathedral was filmed from the top of the BT Tower then the LOCATION_CAPTURE_NAME would be the BT Tower. Not to be confused with the name of the place that the real location is standing in for (see LOCATION_SETTING_NAME). Known aliases: PLACE OF SHOOTING Value Type: Uncontrolled Text Examples: BT Tower, Eiffel Tower, Northern Ireland border near Dundalk.
A68	EDITORIAL_CONTENT_CODE	Classification of a programme, or programme segment, or group of programmes, according to its content or subject. This may also apply to publication services - e.g. a 'sport' channel. External reference: EBU Escort 2.4 Value Type: External Source of referenced data: EBU/Escort: Examples: 3.1.1
A69	EDITORIAL_CONTENT_NAME	Classification of a programme, or programme segment, or group of programmes, according to its content or subject. This may also apply to publication services - e.g. a 'sport' channel. External reference: EBU Escort 2.4 Value Type: Uncontrolled Text Examples: Daily News; Religious Philosophies; Show Business; Medicine; Erotica.
A70	EDITORIAL_CONTROL_CODE	The extend to which the broadcaster has control of the programme content at the production stage. External reference: EBU Escort 2.4 Value Type: Controlled Code Source of referenced data: EBU/Escort/P_Meta: 9.1.8 EDITORIAL_CONTROL_CODE
A71	EDITORIAL_CONTROL_NAME	The extend to which the broadcaster has control of the programme content at the production stage. External reference: EBU Escort 2.4 Value Type: Uncontrolled Text Examples: Own Production; Commissioned Production; Co-production; Exchange; Purchase.
A72	EDITORIAL_FORMAT_CODE	Classification of a programme, or programme segment, or group of programmes, according to its formal structure. How is it structured, regardless of the subject with whom it is dealing. External reference: EBU Escort 2.4 Value Type: Controlled Code Source of referenced data: EBU/Escort/P_Meta: 9.1.9 EDITORIAL_FORMAT_CODE

A73	EDITORIAL_FORMAT_NAME	Classification of a programme, or programme segment, or group of programmes, according to its formal structure. How is it structured, regardless of the subject with whom it is dealing. External reference: EBU Escort 2.4 Value Type: Uncontrolled Text Examples: Magazine; Reading; Performed Drama.
A74	EDITORIAL_TARGET_GROUP_CODE	Classification of a programme, or programme segment, or group of programmes according to the specific audiences, defined by age, cultural/ethnic background profession, etc, for which it is intended. External reference: EBU Escort 2.4 Value Type: Controlled Code Source of referenced data: EBU/Escort/P_Meta: 9.1.10 EDITORIAL_TARGET_GROUP_CODE
A75	EDITORIAL_TARGET_GROUP_NAME	Classification of a programme, or programme segment, or group of programmes according to the specific audiences, defined by age, cultural/ethnic background profession, etc, for which it is intended. External reference: EBU Escort 2.4 Value Type: Uncontrolled Text Examples: Young Children; Senior Citizens; Ethnic Groups; Motorists; General Audience.
A77	MOB_SCRIPT_DESCRIPTION	A description of a media object, written before the instantiation of the object. It is therefore, descriptive of the intended object, rather than of an object which exists. The description may be used as directorial instructions during the future capture process. In post-scripted programmes this may not exist. This will use terms appropriate to the type of media object being described (e.g. Shot, Audio). The abbreviation MOB stands for Media Object. Value Type: Uncontrolled Text
A78	MOB_TITLE	The title of a media object which is a single content element (only one medium). The abbreviation MOB stands for Media Object. Value Type: Uncontrolled Text
A80	GRAPHIC_USAGE_TYPE_NAME	A type of usage to which a graphic is put. Value Type: Uncontrolled Text Examples: title sequence; graphic overlay
A81	ORG_ACCOUNT_NAME	The name associated with an account held by an organisation (or a person) with a financial institution. The abbreviation ORG stands for ORGanisation. Value Type: Uncontrolled Text Examples: BBC WORLDWIDE LTD STERLING ACCOUNT
A82	ORG_ACCOUNT_NUMBER	The identifying number provided by a financial institution for use by an organisation (or a person) in specifying a particular account held with the financial institution. The abbreviation ORG stands for ORGanisation. Value Type: Uncontrolled Text

A83	ORG_NAME	The name of an organisation, that is any grouping of individuals, company or company structure with whom there is an association to carry out tasks or responsibilities for the development, usage or management of media assets. The abbreviation ORG stands for ORGanisation. Value Type: Uncontrolled Text Examples: BBC Media Data Group; RAI; SMPTE; World Health Organisation
A86	ADDRESS_COUNTY_STATE_NAME	The county (or, where appropriate, state or district) name in a postal address. External reference: OpenSMEF Attribute - PAD_COUNTY_STATE_NAME Value Type: Uncontrolled Text Examples: Examples include Berkshire, Idaho, Manitoba.
A87	PERSON_SALUTATION_SHORT_ FORM	A shortened salutation as applied to a person's name. External reference: OpenSMEF Attribute Value Type: Uncontrolled Text Examples: Mr.; Miss; Mrs.; Ms; Lord.
A88	PERSON_FIRST_NAME	The first minor name of the person. In the western European tradition this would actually be the person's first name. For example, in the Western European name John Smith Jordan this would be John. You would not expect this name to be the primary determinant of the person's position on an alphabetically ordered list. Known aliases: Christian Name; Given Name Value Type: Uncontrolled Text Examples: Richard, Daniel, Laurent, Giorgio
A89	PERSON_LAST_NAME	The person's main name. You would expect this name to be the primary determinant of the person's position on an alphabetically ordered list. In the Western European tradition, and many others, this would be the person's surname or family name. In the name John Smith Jordan the main name is Jordan. This field, however, is also used for names which do not fall into the, typically, western model of first name, middle name(s), and last name. Known aliases: Surname; Family Name Value Type: Uncontrolled Text Examples: Jordan, Boch, Teruggi, Carter
A90	PERSON_MIDDLE_NAME	The secondary minor name(s) of the person. For example, in the Western European name John Smith Jordan this would be Smith. You would not expect this name to be the primary determinant of the person's position on an alphabetically ordered list. Value Type: Uncontrolled Text
A93	PERSON_SUFFIX_NAME	Qualifications, awards, and/or honours normally appended to a person's name Value Type: Uncontrolled Text Examples: O.M.

A94	PEV_INTENTION_CODE	The publication event (PEV) intention. That is, the classification of a programme, or programme segment, or group of programmes, according to the broadcaster's primary intention in publishing it/them. The abbreviation PEV stands for Publication Event. External reference: EBU Escort 2.4 Value Type: Controlled Code Source of referenced data: EBU/Escort/P_Meta: 9.1.25 PEV_INTENTION_CODE
A95	PEV_INTENTION_NAME	The publication event (PEV) intention. That is, the classification of a programme, or programme segment, or group of programmes, according to the broadcaster's primary intention in publishing it/them. The abbreviation PEV stands for Publication Event. External reference: EBU Escort 2.4 Value Type: Uncontrolled Text Examples: Pure Entertainment; Infotainment; Involvement.
A96	PGR_EPISODE_QUANTITY	The number of programme episodes in serial or series. The abbreviation PGR stands for Programme Group. Value Type: Uncontrolled Text
A97	PGR_SUB_TITLE	The sub-title for a group of programmes. The abbreviation PGR stands for Programme Group. Value Type: Uncontrolled Text
A98	PGR_SYNOPSIS_DESCRIPTION	A synopsis of a group of programmes. The abbreviation PGR stands for Programme Group. Value Type: Uncontrolled Text
A99	PGR_TITLE	The title for a group of programmes. The abbreviation PGR stands for Programme Group. Value Type: Uncontrolled Text Examples: Soul of Britain (BBC series); London's Burning (commercial serial); Wildlife on One (BBC strand)
A100	PGR_TX_CYCLE_CODE	The nature of the group of programmes for which this programme was created. This group may be composed solely of the programme itself or it may include other programmes. The abbreviation PGR stands for Programme Group. External reference: EBU Escort 2.4 Known aliases: Programme transmission cycle Value Type: Controlled Code Source of referenced data: EBU/Escort/P_Meta: 9.1.27 PGR_TX_CYCLE_CODE

A101	PGR_TX_CYCLE_NAME	The nature of the group of programmes for which this programme was created. This group may be composed solely of the programme itself or it may include other programmes. The abbreviation PGR stands for Programme Group. External reference: EBU Escort 2.4 Known aliases: Programme transmission cycle Value Type: Uncontrolled Text Examples: Anthology, Series, Serial, Single production
A103	COLOUR_CODE	A code specifying the editorial colour description for the content being exchanged. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.3 COLOUR_CODE
A104	PROGRAMME_EPISODE_COUNT	The sequence number for an individual programme within a serial or series Value Type: Uncontrolled Text
A105	IDENTIFIER_NUMBER	A unique identifier assigned to the programme group, programme, item, media object, or an instance of one of those by the producing organisation for identification and other purposes. The type of identifier used must also be exchanged. Known aliases: Production Number, Value Type: Uncontrolled Text Examples: LCD1230Nx; ABH521021A
A107	PROGRAMME_SUB_TITLE	The sub-title given to a programme by its producer Value Type: Uncontrolled Text
A109	PROGRAMME_SYNOPSIS_DESCRIPTION	A synopsis of a programme. Value Type: Uncontrolled Text
A110	PROGRAMME_TITLE	The title given to a programme by its producer External reference: EOV_TITLE Value Type: Uncontrolled Text Examples: Tony Bennett: Live at Leeds Castle
A114	PROGRAMME_EPISODE_TITLE	A name given to a particular programme in a group of programmes. For example: a programme in the series 'Panorama' may also have the Programme Title 'Panorama' and an Episode Title 'The BSE Crisis'; a programme in the series 'Dr Who and the Aztecs' may have a Programme Title 'Dr Who' and an Episode Title 'The Warrior of Death'. External reference: OpenSMEF Attribute Value Type: Uncontrolled Text
A116	RIGHT_CONDITION_DESC	A description of the conditions which govern or restrict the Right. For example, a specific media asset cannot be published over the Christmas period. External reference: OpenSMEF Attribute Value Type: Uncontrolled Text

A117	RIGHT_END_DATE	The end date of the period covered by this Right, i.e. the date to which the rights can be exploited. External reference: OpenSMEF Attribute Value Type: Formatted Code according to ISO 8601 complete date (YYYY-MM-DD) Examples: 2002-06-30
A118	RIGHT_START_DATE	The start date of the period covered by this Right, i.e. the date from which the rights can be exploited. External reference: OpenSMEF Attribute Value Type: Formatted Code according to ISO 8601 complete date (YYYY-MM-DD) Examples: 2002-06-30
A121	RIGHT_TYPE_CODE	The internationally agreed identifier for the type of right. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.30 RIGHT_TYPE_CODE
A122	RIGHT_TYPE_DESCRIPTION	Describes the type of right, for example the right to broadcast or the right to publish. Value Type: Uncontrolled Text Examples: 'hereby grants to the Licensee a non-exclusive licence to incorporate the Selected Material in the Licensee's production ('the Production') and to exploit the Production as it incorporates the Selected Material on the following conditions: '
A123	ROLE_NAME	The name given, where appropriate, to the role played by a contributor in a production. Value Type: Uncontrolled Text Examples: Lady Bracknell; First Policeman; Ulysses; Mother Courage.
A124	ROLE_TYPE_CODE	The title of the type of role, responsibility, or task undertaken by an organisation or an individual in the development, management or control of material Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.31 ROLE_TYPE_CODE
A125	ROLE_TYPE_NAME	The title of the type of role, responsibility, or task undertaken by an organisation or an individual in the development, management or control of material Value Type: Uncontrolled Text Examples: Cameraman; Producer; Chaperone; Animal Trainer; Publicist
A131	SOUND_FORMAT_CODE	A controlled code for a sound format type. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.32 SOUND_FORMAT_CODE
A132	SOUND_FORMAT_NAME	The name of the sound format type e.g. mono, 2 channel stereo, 5 channel surround sound, etc. Value Type: Uncontrolled Text Examples: Mono; Stereo; 5 Channel surround sound

A135	STORAGE_TYPE_CODE	The code given to identify the form of storage used for material. "Form of storage" is a bundled definition which includes all of the elements of specification required to support storage. The attribute value table points to the specification(s) defining everything needed to extract the information from the medium: the physical container (e.g. D3 cassette), the medium (e.g. tape), any significant characteristic of the medium (e.g. metal-particle tape), characteristics of the head-to-tape interface (e.g. head penetration), and track format. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.34 STORAGE_TYPE_CODE
A136	STORAGE_TYPE_NAME	The name given to a form of storage used for material. "Form of storage" is a bundled definition which includes all of the elements of specification required to support storage. The attribute value table points to the specification(s) defining everything needed to extract the information from the medium: the physical container (e.g. D3 cassette), the medium (e.g. tape), any significant characteristic of the medium (e.g. metal-particle tape), characteristics of the head-to-tape interface (e.g. head penetration), and track format. Value Type: Uncontrolled Text Examples: DAT; Minidisk; U-MATIC SP; Digital D3.
A137	STORAGE_IDENTIFIER	An identity used by an organisation for a storage medium occurrence, e.g. a tape number Value Type: Uncontrolled Text
A141	TERRITORY_CODE	A controlled code for a Territory. The code may be selected from any of the referenced sources of pre-defined values. ISO 3166-1 provides country codes, ISO 3166-2 country sub-division codes, and EBU/P_Meta provides codes only for a few particular cases. Value Type: External Source of referenced data: ISO 3166-1, ISO 3166-2, EBU/P_Meta: Examples: GB; IT-MI; DE; XW; XE; XX
A143	TET_USAGE_TYPE_NAME	Description of the type of usage for which a text data object text is intended. For example -Sub-titles; WEB Text. Value Type: Uncontrolled Text
A146	PROGRAMME_WORKING_TITLE	The title given to a programme by its producer before the final title is decided. This may be used for reasons of commercial security and resources may be booked to the working title as a way of assuring their availability. External reference: EOV_WORKING_TITLE Value Type: Uncontrolled Text Examples: Production x; Shakespeare 1.
A148	ITEM_SCRIPT_SCENE_NUMBER	The scene number which identifies the position of the scene, in relation to others in the production, within the shooting script. External reference: SMPTE key 01 05 06 00 00 00 00 00 Value Type: Uncontrolled Text Examples: 1; 34; 2a; 3c.

A149	ITEM_SCRIPT_SCENE_TAKE_COUNT	An incrementing number identifying the particular take of a specified scene within a production. External reference: SMPTE key 01 05 07 00 00 00 00 00 Value Type: Integer Examples: 1; 5; 10; 4.
A150	IDENTIFIER_TYPE_CODE	A code identifying the type of identifier being exchanged. Known aliases: Programme number, Production number Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.12 IDENTIFIER_TYPE_CODE
A151	IDENTIFIER_TYPE_NAME	A text name identifying the type of material identifier being exchanged. Known aliases: Programme number, Production number Value Type: Uncontrolled Text Examples: Organisation ID; ISAN; UPID; UPN; UMID.
A152	PRODUCTION_END_DATE	The date on which the production of the material referred to in the encompassing set or communication was completed. External reference: EBU ESCORT 2.4 Value Type: Formatted Code according to ISO 8601 complete date (YYYY-MM-DD) Examples: 2002-06-30
A154	STORAGE_IDENTIFIER_TYPE_CODE	A code identifying the type of storage identifier being used. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.33 STORAGE_IDENTIFIER_TYPE_CODE
A158	ORIGINATION_CODE	A controlled code identifying the (EBU defined) process sector from which the material was originated. External reference: EBU Escort 2.4 Value Type: Controlled Code Source of referenced data: EBU/Escort/P_Meta: 9.1.20 ORIGINATION_CODE
A159	PERSON_DESCRIPTION	A textual description of a person used, where necessary, to distinguish them from another person of the same name. Value Type: Uncontrolled Text Examples: Member of Scottish Parliament; Dentist; Poet; Historian; American farmer.
A160	ADDRESS_FACSIMILE_NUMBER	A facsimile machine number, the precise nature and/or context of which will be defined by the set containing the number. Value Type: Uncontrolled Text Examples: 020 8534 1234
A162	RIGHT_TERRITORY_DESCRIPTION	A description of the geographical territory over which the assigned right(s) apply. Value Type: Uncontrolled Text
A164	CURRENCY_AMOUNT	The amount of money in the currency defined by the associated code or name Value Type: Floating Point

A165	TET_USAGE_TYPE_CODE	A controlled code of the type of usage for which a text data object is intended. For example - Sub-titles; WEB Text. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.35 TET_USAGE_TYPE_CODE
A166	MOB_TYPE_CODE	This is the type of media object being exchanged: SHOT; GRAPHIC; DATA; STILL; AUDIO. The abbreviation MOB stands for Media Object. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.17 MOB_TYPE_CODE
A168	DEVICE_MODEL_NAME	The model name as applied to a device by the manufacturer and by which the device is known. Value Type: Uncontrolled Text
A169	DEVICE_SERIAL_NUMBER	Serial number applied to a particular instance of a device by its manufacturer. Value Type: Uncontrolled Text
A174	AUC_SAMPLE_RATE	Specifies the rate of sampling used to create the audio clip instance in samples per second. The abbreviation AUC stands for AUdio media object instanCe. Value Type: Uncontrolled Text Examples: 44000; 32000.
A175	AUC_BITS_PER_SAMPLE_QUANTITY	The maximum number of significant bits for the value without compression. The abbreviation AUC stands for AUdio media object instanCe. External reference: SMPTE MDD 04 02 03 03 01 00 00 00 Known aliases: Audio resolution; Bits per sample Value Type: Uncontrolled Text Examples: 16; 20; 24;
A176	SUBTITLE_FLAG	This indicates whether an item/programme/programme group includes subtitles. The nature of the subtitle (e.g. closed caption or teletext) is communicated as information related to the subtitle text object. Value Type: Boolean

A177	PFT_ACTION_HORIZONTAL_ SAFE_PERCENTAGE	When a programme is commissioned in widescreen the aspect ratio in which the programme will be transmitted on an analogue service will also be stipulated. In most cases this will be 14:9 letterbox but in some cases (e.g. sport) 4:3 centre cut-out will be requested. In exceptional cases only is a 16:9 deep letterbox image transmitted on the analogue service. Defining the analogue transmission format allows programme makers shooting in 16:9 to limit key elements of the scene to the section of the frame that will be seen by the viewer. This attribute specifies the percentage of the image which will lie outside the vertical safe area for action (cf. graphics). The percentage is of the total raster and not merely of the shoot to protect area and the figure is for both left and right sides of the screen. Thus, a safe area of 3.5% would mean that 7% of the total horizontal portion of the picture is regarded as unsafe. The abbreviation PFT stands for Picture Format Type. External reference: OpenSMEF Attribute Value Type: Floating Point Examples: 3.5
A178	AUC_REFERENCE_LEVEL_QUANTITY	The audio level in Dbm for 0VU. The abbreviation AUC stands for AUdio media object instanCe. External reference: SMPTE MDD RP210A 04 02 01 01 03 00 00 00 Value Type: Uncontrolled Text
A180	AUC_BITRATE_QUANTITY	Specifies the bitrate (in bits per second) of the used audio compression scheme. The abbreviation AUC stands for AUdio media object instanCe. External reference: SMPTE MDD 04 02 03 01 02 00 00 00 Known aliases: Audio Average Bitrate; Audio compression bitrate Value Type: Integer
A181	PCS_AVERAGE_BITS_PER_ SECOND_QUANTITY	Where required, the average data rate (in bits per second) of the coded shot media object. Used by playback/decoding software to estimate buffer size. The abbreviation PCS stands for Picture Coding Scheme. Value Type: Integer
A183	PCS_FIXED_BITRATE_INDICATOR	A flag indicating whether the coding standard delivers a fixed bitrate to the decoding/playback device. The abbreviation PCS stands for Picture Coding Scheme. Value Type: Boolean
A184	AUC_FIXED_BITRATE_INDICATOR	A flag indicating whether the coding standard delivers a fixed bitrate to the decoding/playback device. 'True' indicates a fixed bitrate. The abbreviation AUC stands for AUdio media object instanCe. Value Type: Boolean
A185	PCS_SAMPLE_RATE	Specifies the rate of luminance sampling in samples per second. The abbreviation PCS stands for Picture Coding Scheme. Value Type: Uncontrolled Text
A187	PROGRAMME_SCRIPT_DESCRIPTION	A text script for a complete programme. This may be a complete production script or it may only contain the spoken text. It differs, therefore, from a media object script which is a description of a single media object in terms applicable to that media object type. Value Type: Uncontrolled Text

A188	MSI_TITLE	This is used with music (live or recorded) and other recorded material (e.g. speech or sound) which is sourced commercially and the use of which must be reported to the appropriate rights bodies. This is the full title of the piece of material used as shown on the recording or sheet music (if used). Where only a part of a long work is used, the overall title of the work, followed by the details of the movement(s) used must be given. The abbreviation MSI stands for Music or Speech Item. Known aliases: track Title, Work title Value Type: Uncontrolled Text
A189	MSI_SIDE_COUNT	This is used with music (live or recorded) and other recorded material (e.g. speech or sound) which is sourced commercially and the use of which must be reported to the appropriate rights bodies. Where appropriate, this is the side number of the recording on which the track used appears. The abbreviation MSI stands for Music or Speech Item. Value Type: Integer
A190	MSI_TRACK_COUNT	This is used with music (live or recorded) and other recorded material (e.g. speech or sound) which is sourced commercially and the use of which must be reported to the appropriate rights bodies. Where appropriate, this is the track number of the item used. The abbreviation MSI stands for Music or Speech Item. Value Type: Integer Examples: 4, 3, 10.
A191	MSI_LABEL_NAME	This is used with music (live or recorded) and other recorded material (e.g. speech or sound) which is sourced commercially and the use of which must be reported to the appropriate rights bodies. This is the name of the publication label from which the material used was taken. The abbreviation MSI stands for Music or Speech Item. Value Type: Uncontrolled Text Examples: REPRISE; STRANGE FRUIT; TOPIC; UNITED ARTISTS.
A192	MSI_COLLECTION_NAME	This is used with music (live or recorded) and other recorded material (e.g. speech or sound) which is sourced commercially and the use of which must be reported to the appropriate rights bodies. Where appropriate, this is the name of the published collection (e. g.: CD name; LP name; set name) from which the material used was taken. The abbreviation MSI stands for Music or Speech Item. Value Type: Uncontrolled Text Examples: Wheels of Fire; Rainmaker; Sinatra Live at the Sands; Live at the Village Vanguard.
A193	MSI_NUMBER	This is used with music (live or recorded) and other recorded material (e.g. speech or sound) which is sourced commercially and the use of which must be reported to the appropriate rights bodies. This is the main catalogue number for the recording from which the material used was taken. The abbreviation MSI stands for Music or Speech Item. Value Type: Uncontrolled Text Examples: 298; 899; 0322; 060.

A194	MSI_PREFIX_NUMBER	This is used with music (live or recorded) and other recorded material (e.g. speech or sound) which is sourced commercially and the use of which must be reported to the appropriate rights bodies. This is the alphanumeric prefix to the catalogue number for the recording from which the material used was taken. The abbreviation MSI stands for Music or Speech Item. Value Type: Uncontrolled Text Examples: SFRSCD; LC; PY; TSCD.
A195	MSI_SIGNATURE_TUNE_INDICATOR	This is used with music (live or recorded) and other recorded material (e.g. speech or sound) which is sourced commercially and the use of which must be reported to the appropriate rights bodies. This indicates that the music was used as the signature tune of the item/programme in which it is used. The abbreviation MSI stands for Music or Speech Item. Value Type: Boolean
A196	MSI_TV_BACKGROUND_INDICATOR	This is used with music (live or recorded) and other recorded material (e.g. speech or sound) which is sourced commercially and the use of which must be reported to the appropriate rights bodies. Music used in a television programme is categorised as either BACKGROUND or VISUAL. Background music is music added solely for the benefit of the television audience and, thus, forms no intrinsic part of the programme. VISUAL music forms part of the programme in as much as it appears to be, or is, audible to those taking part in the programme. Examples of visual music include: music sung or played in vision; music which characters in a play can hear, even if it comes from a performer or player not in vision. All actuality music is visual. A value of TRUE would indicate that the music is BACKGROUND music, a value of FALSE indicates VISUAL music. The abbreviation MSI stands for Music or Speech Item. Value Type: Boolean
A198	PGR_WORKING_TITLE	Where appropriate, the working title for a group of programmes. The abbreviation PGR stands for Programme Group. Value Type: Uncontrolled Text Examples: Wildlife series 2002/3; Costume drama - Christmas 2003;
A200	CONTRACT_LINE_NAME	A free text description for the service or product covered by the contract line. Value Type: Uncontrolled Text
A201	CONTRACT_LINE_NUMBER	An identifier number for the contract line within a contract. Value Type: Uncontrolled Text
A202	CONTRACT_TERMS_OF_ BUSINESS_DESCRIPTION	This is a free text definition of the terms of business which apply to a specified contract or transaction, defined and/or specified by the organisation issuing the contract. These will vary from organisation to organisation and, perhaps, within organisations from transaction to transaction. Value Type: Uncontrolled Text
A203	RIGHT_TRANSMISSION_COUNT	The right to transmit specified material a set number of times (specified by the attribute value) as granted in an agreement or contract. Value Type: Formatted Code Examples: 1; 4; 10; Unlimited.

A204	CONTRACT_PAYMENT_ INSTALMENT_COUNT	Where contracts/agreements allow for payment by instalment this attribute identifies the particular instalment in the sequence of payments (e.g. instalment 1, instalment 5, etc.) Value Type: Integer Examples: 1, 4, 6, 23.
A205	CONTRACT_PAYMENT_ INSTALMENT_DUE_DATE	Where contracts/agreements allow for payment by instalment this attribute identifies the date by which the particular instalment in the sequence of payments (e.g. instalment 1, instalment 5, etc.) is due. The date is inclusive in that the payment may be made on the date quoted. Value Type: Formatted Code according to ISO 8601 complete date (YYYY-MM-DD) Examples: 2002-06-30
A206	CONTRACT_PAYMENT_ INSTALMENT_AMOUNT	Where contracts/agreements allow for payment by instalment this attribute identifies the amount of money due in a specified instalment. Value Type: Formatted Code
A207	LOCATION_SETTING_NAME	The dramatic setting of a piece of content. For example, a play set in Moscow might be shot in Dundee. In this case, although the material was filmed in Dundee the LOCATION_SETTING_NAME would be Moscow. The setting name might include more than just place information. It might, for example, define historical period. In the case above, the setting might have been Moscow, 1930s. The setting might also be completely fictional. An example of this would be a setting of 'Utopia' in a dramatisation of Sir Thomas Moore's book or 'An Outer Circle of Hell' in an adaptation of Dante's "Inferno". Value Type: Uncontrolled Text Examples: Paris 1789; Quantico; Naples, Beppe's Coffee shop
A209	PFT_INTENDED_DISPLAY_ ASPECT_RATIO	The aspect ratio of the complete image as intended for display (including any black edges) and, thus, the aspect ratio at which the complete image must be displayed to avoid distortion. Expressed as a quantitative relation between two integers, the values being separated by a colon. Examples are: 4:3 ; 16:9. The abbreviation PFT stands for Picture Format Type. Value Type: Formatted Code Examples: 4:3; 16:9
A210	PFT_DISPLAY_FORMAT_CODE	Identifies the visual effect of the illuminated portion of the image as displayed within the complete image at the relevant intended display aspect ratio. Effects include: Fullframe, Letterbox, Pillarbox The abbreviation PFT stands for Picture Format Type. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.26 PFT_DISPLAY_FORMAT_CODE
A212	RIGHT_GRANT_OF_RIGHTS_ MEDIA_NAME	A textual description of the type of media covered by the rights granted by a contract/agreement. Value Type: Uncontrolled Text Examples: Basic Cable Service; Pay Satellite Service; Theatric Distribution.

A213	RIGHT_GRANT_OF_RIGHTS_ MEDIA_CODE	A controlled code defining the type of media covered by the rights granted by a contract/agreement. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.29 RIGHT_GRANT_OF_RIGHTS_MEDIA_CODE
A214	LOCATION_ACTION_NAME	The name of the location of the captured (or to be captured) action. For example, if St Paul's Cathedral was filmed from the top of the BT Tower then the LOCATION_ACTION_NAME would be St. Paul's Cathedral. Not to be confused with the name of the place that the real location is standing in for (see LOCATION_SETTING_NAME). Value Type: Uncontrolled Text Examples: BT Tower, Eiffel Tower, Northern Ireland border near Dundalk.
A216	MOB_CAPTURED_DESCRIPTION	A description of a captured media object. This is the description of the shot which would be found on a 'shot list'. In post-scripted programmes this may be the only description of the media object. This will use terms appropriate to the type of media object being described (e.g. Shot, Audio). The abbreviation MOB stands for Media Object. Value Type: Uncontrolled Text
A217	EVENT_LOCAL_END_DATE	The local date at the end of an event specified by the set containing the attribute. Value Type: Formatted Code according to ISO 8601 complete date (YYYY-MM-DD) Examples: 2002-06-30
A218	EVENT_LOCAL_START_DATE	The local date at the beginning of an event specified by the set containing the attribute. Value Type: Formatted Code according to ISO 8601 complete date (YYYY-MM-DD) Examples: 2002-06-30
A219	EVENT_LOCAL_END_TIME	The local time at the end of an event specified by the set containing the attribute. Value Type: Formatted Code according to ISO 8601, with three digits for representing the decimal fractions of a second (hh:mm:ss.sss) Examples: 23:40:01.500
A220	EVENT_LOCAL_START_TIME	The local time at the beginning of an event specified by the set containing the attribute. Value Type: Formatted Code according to ISO 8601, with three digits for representing the decimal fractions of a second (hh:mm:ss.sss) Examples: 23:40:01.500
A221	TIME_GMT_OFFSET_COUNT	A signed integer defining the offset, in minutes, between local time and GMT. At 12:00 (noon) GMT an offset of +60 would mean a local time of 13:00 (1 p.m.) an offset of -60 would mean a local time of 11:00 (11 a.m.). Value Type: Integer
A222	FILE_FORMAT_TYPE_NAME	The name of the format of a computer file. Value Type: Uncontrolled Text Examples: MXF; Broadcast Wave; Mpeg Transport Stream; PDF

A223	FILE_BYTES_QUANTITY	The number of bytes of a computer file. Value Type: Unsigned integer Examples: 123456
A224	FILE_PATH_NAME	The name of the file path to a complete digital media, data, metadata etc file, that will be unique within a file server or a storage medium. External reference: SMPTE MDD Value Type: Uncontrolled Text Examples: /storage_A/PMeta/export/
A225	ISAN_PROGRAMME_TYPE_CODE	It specifies the programme category according to the International ISAN Agency Value Type: External Source of referenced data: ISAN
A226	PCS_COMPRESSION_NAME	The name of the analogue or digital compression scheme applied to the picture. The abbreviation PCS stands for Picture Coding Scheme Value Type: Uncontrolled Text Examples: Examples include: PAL; NTSC; SECAM; DV; DV-BASED; MPEG; M-JPEG; JPEG.
A227	MATERIAL_RELATIONSHIP_TYPE_CODE	This specifies the type of relationship existing between materials identified in association with this attribute in a superset. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.16 MATERIAL_RELATIONSHIP_TYPE_CODE
A245	SHI_GOOD_SHOT_INDICATOR	A flag, held against the original instance of any shot media object, indicating that the creator suggests this particular take to be that preferred for the purposes for which it was created. The abbreviation SHI stands for SHot media object Instance. Value Type: Boolean
A247	SHO_LOGO_INDICATOR	A flag which indicates that all instances of the specified shot contain an on-screen logo. This only applies to shot media objects. The abbreviation SHO stands for SHOt media object. Value Type: Boolean
A249	REFERENCE_SCHEME_NAME	This is the name of an external (non-P/Meta) reference scheme which may be used to enable communication of information. Value Type: Uncontrolled Text Examples: SMPTE Meta Data Dictionary
A250	REFERENCE_SCHEME_CODE	This is code identifying an external (non-P/Meta) reference scheme which may be used to enable communication of information. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.28 REFERENCE_SCHEME_CODE

A251	REFERENCE_ATTRIBUTE_NAME	This is the name of an attribute as defined by an external (non-P/Meta) reference scheme and which may be used to enable communication of information. The name should be sufficient to identify the attribute unambiguously within the scheme from which it comes. External reference: The scheme identified by the REFERENCE_SCHEME_CODE/NAME value communicated within the same sub-set. Value Type: Uncontrolled Text Examples: Genre; 04 02 01 01 03 00 00 00 (this being the ID for Audio Reference Level from the SMPTE Meta Data Dictionary)
A252	REFERENCE_ATTRIBUTE_CODE	This is a code identifying an attribute as defined by an external (non-P/Meta) reference scheme and which may be used to enable communication of information. The code value should be sufficient to identify the attribute unambiguously within the scheme from which it comes. Value Type: External Source of referenced data: The scheme identified by the REFERENCE_SCHEME_CODE/NAME value communicated within the same sub-set. Examples: 04 02 01 01 03 00 00 00 (this being the ID for Audio Reference Level from the SMPTE Meta Data Dictionary)
A253	REFERENCE_ATTRIBUTE_VALUE_NAME	This attribute contains the value of an attribute (as opposed to the attribute ID or name) from an external (non-P/Meta) description scheme. The information is exchanged as part of a P/Meta set (e.g. S31 - EXTERNAL_SCHEME_EXCHANGE_SET). It is incumbent on the user of this attribute that they should format the attribute value according to the rules required for that attribute within its definition scheme and, where the attribute within the scheme has values limited to a pre-defined list, ensure that the attribute value is valid. External reference: The scheme identified by the REFERENCE_SCHEME_CODE/NAME value communicated within the same sub-set. Value Type: Uncontrolled Text
A254	ORG_IDENTIFIER	An identifier associated with an organisation. Examples include, individual company registration numbers, individual VAT registration numbers, internal department or branch numbers and so on. The abbreviation ORG stands for ORGanisation. Value Type: Uncontrolled Text
A255	ORG_IDENTIFIER_TYPE_NAME	The name of a type of identifier applying to an organisation The abbreviation ORG stands for ORGanisation. Value Type: Uncontrolled Text Examples: British Bank Branch Sort Code
A256	ORG_IDENTIFIER_TYPE_CODE	A code for a type of identifier applying to an organisation. The abbreviation ORG stands for ORGanisation. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.18 ORG_IDENTIFIER_TYPE_CODE

A257	AUC_COMPRESSION_NAME	The name for any compression scheme, noise reduction scheme, or other non-linear processing applied to an audio signal. The abbreviation AUC stands for AUdio media object instanCe. Value Type: Uncontrolled Text Examples: Examples: MPEG Layer II; Dolby A; IEC pre-emphasis.
A258	AUC_COMPRESSION_CODE	The controlled code identifying any compression scheme, noise reduction scheme, or other non-linear processing applied to an audio signal. The abbreviation AUC stands for AUdio media object instance. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.1 AUC_COMPRESSION_CODE
A259	PCS_COMPRESSION_CODE	The name of the analogue or digital compression scheme applied to the picture. The abbreviation PCS stands for Picture Coding Scheme Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.21 PCS_COMPRESSION_CODE
A260	SET_EXPLANATORY_NOTE_NAME	This attribute is defined for use in EBU/PMETA sets to convey any required explanatory information. Value Type: Uncontrolled Text
A261	PEV_COUNT	This is a number of complete publications of a specified programme group, programme, item, or media object. The exact meaning of the attribute is clarified and circumscribed by the set within which it is used. The abbreviation PEV stands for Publication Event. Value Type: Integer
A361	FILE_FORMAT_TYPE_CODE	A controlled code communicating the format of a computer file. The value shall be a pair "content type"/"content sub-type" registered at IANA, as defined by IETF RFC 2045 and 2046. Value Type: External Source of referenced data: IETF (RFC 2045,2046 and other) and IANA Examples: text/plain; text/rtf; audio/basic; image/jpeg ; image/png; video/mpeg; video/quicktime; application/pdf; application/postscript
A362	INDEX_EXTERNAL_SCHEME_CODE	Where a subject description requires the use of an external (i.e. non-EBU) scheme, and where the descriptive scheme has been previously identified and registered in the EBU P/Meta metadata scheme (attribute values table), this code identifies the descriptive scheme. An example would be the Library of Congress Subject Headings. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.13 INDEX_EXTERNAL_SCHEME_CODE

A363	INDEX_EXTERNAL_SCHEME_NAME	A free-text field used where a subject description requires the use of an external (i.e. non-EBU) description scheme. This name identifies the description scheme being used. This is of use where the description scheme has not previously been identified and registered as a code in the EBU P/Meta metadata scheme. An example of this might be an internal scheme in use by one of the partners in the transaction. Value Type: Uncontrolled Text Examples: Library of Congress Subject Headings.
A364	INDEX_EXTERNAL_SCHEME_TERM_NAME	Where a subject description requires the use of an external (i.e. non-EBU) description scheme this free-text attribute contains the descriptive term taken from the given scheme. Value Type: Uncontrolled Text
A365	INDEX_EXTERNAL_SCHEME_TERM_CODE	Where a subject description requires the use of an external (i.e. non-EBU) description scheme this attribute contains a controlled code, taken from the scheme in use, which represents a descriptive term. Because this is not a P/Meta controlled code (that is, the codes and de-codes are not registered as part of the EBU P/Meta scheme but are controlled by the organisation responsible for the external description scheme) care must be taken by the party originating the information that the attribute value is encoded exactly as required by the external description scheme. Value Type: External Source of referenced data: The organisation responsible for the Index External Scheme as identified with either INDEX_EXTERNAL_SCHEME_CODE or INDEX_EXTERNAL_SCHEME_NAME.
A366	INDEX_EXTERNAL_SCHEME_TERM_SCOPE_DESCRIPTOR	When a descriptive term taken from an external (i.e. non-EBU) description scheme is used this free-text attribute may contain information further defining the scope of the term used. Value Type: Uncontrolled Text
A367	SET_DATA_VALIDITY_DATE	This attribute is defined for use in EBU/PMETA sets to convey the local date on which the information conveyed by the set is deemed to be valid. This is generally used in conjunction with the SET_DATA_VALIDITY_TIME and TIME_GMT_OFFSET_COUNT attributes. Value Type: Formatted Code according to ISO 8601 complete date (YYYY-MM-DD) Examples: 2002-06-30
A368	SET_DATA_VALIDITY_TIME	This attribute is defined for use in EBU/PMETA sets to convey the local time at which the information conveyed by the set is deemed to be valid. This is generally used in conjunction with the SET_DATA_VALIDITY_DATE and TIME_GMT_OFFSET_COUNT attributes. Value Type: Formatted Code according to ISO 8601, with three digits for representing the decimal fractions of a second (hh:mm:ss.sss) Examples: 14:33:59.999
A401	REVIEW_TITLE	The title given to a review by its author. Value Type: Uncontrolled Text
A402	REVIEW_SUB_TITLE	The sub-title given to a review by its author. Value Type: Uncontrolled Text

A403	REVIEW_EXCERPT_TEXT	A text selected and extracted from a review. Value Type: Uncontrolled Text
A404	REVIEW_COMPLETE_TEXT	The complete text of a review. Value Type: Uncontrolled Text
A405	REFERENCE_YEAR	The year with which a particular production is generally associated for cataloguing and reference purposes. This may either refer to the year of production or of first publication (or of first publication in a particular version) but is generally synonymous with the year given in the copyright or performance copyright. Note that although some instances of year of reference may be shown in Roman Numerals within copyright statements (e.g. MCMLXXXVII for 1987) that P/Meta will hold the value in Arabic Numerals. Value Type: Formatted Code Examples: 1987; 2002;
A406	LIVE_ACTION_CODE	Indicates whether a production features live action, animation or a mixture of both. This is required for ISAN and V-ISAN registration purposes. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.15 LIVE_ACTION_CODE
A407	LANGUAGE_USAGE_CODE	A code to indicate how a particular language is used in a particular context. Examples might include commentator language, original spoken dialogue, subtitle language, dubbed language and so on. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.14 LANGUAGE_USAGE_CODE
A408	LANGUAGE_USAGE_DESCRIPTION	A free form description of how a particular language is used in a particular context. Examples might include commentator language, original spoken dialogue, subtitle language, dubbed language and so on. Value Type: Uncontrolled Text Examples: Original Spoken Dialogue, Commentator Language, Original Narration.
A409	PEV_LIVE_TRANSMISSION_FLAG	A flag communicating whether a publication event (PEV) is a live transmission or not. Value Type: Boolean
A410	SERVICE_LINEUP_TIME	The time interval, in seconds, which a service, used for instance for a material exchange, is set up and ready before actual material exchange. Value Type: Unsigned integer
A411	SERVICE_HANDOFF_TIME	The time interval, in seconds, which a service, used for instance for a material exchange, is still set up and ready after the end of actual material exchange. Value Type: Unsigned integer

A412	ADDRESS_WEB_ADDRESS	An address of a resource on the world-wide web in the form of a URL or an http. External reference: Web addresses are formatted according to standards laid down by IETF/RFC Standard 1630 Value Type: Uncontrolled Text Examples: http://www.ebu.org.ch/ , http://www.bbc.co.uk/ , http://www.rai.it/ , http://www.ina.fr/
A413	ORG_TYPE_CODE	A codified value which indicates an organisation's type. This is stated in terms of the organisation's constitution. E.g. whether it is an internal unit of another organisation, a limited company, charity, co-operative and so on. The abbreviation ORG stands for ORGanisation. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.19 ORG_TYPE_CODE
A414	ORG_TYPE_NAME	The name of a particular type of organisation. The abbreviation ORG stands for ORGanisation. Value Type: Uncontrolled Text Examples: Limited Company, Partnership, Co-operative, Charity, Internal Division
A415	SIGN_LANGUAGE_CODE	A codified value for a system of sign language. Value Type: External Source of referenced data: SIL/Ethnologue: Examples: ASE; BHO; GSG; ISE
A416	SIGN_LANGUAGE_NAME	The name of a system of sign language. Value Type: Uncontrolled Text Examples: British Sign Language, American Sign Language.
A417	SIGN_LANGUAGE_PRIMARY_ LANGUAGE_INDICATOR	Indicates whether sign language is the primary language of the content, or whether it has been added later. Value Type: Boolean
A418	SIGN_LANGUAGE_ TRANSLATION_INDICATOR	Indicates whether the sign language used is a direct translation of the spoken dialogue of the content or is a commentary or summary of what is going on. Value Type: Boolean
A419	SERVICE_TRANSMISSION_ BITRATE_QUANTITY	Specifies the bitrate (in bits per second) allocated to a transmission service, used for instance for a material exchange. Value Type: Unsigned integer
A420	MOB_EXCHANGE_TRACK_ IDENTIFIER	An identifier of the track containing the Media Object instance during an exchange. This can be used, when applicable, for exchange by a physical medium (e.g. a Video Tape) as well as by a transmission link. The abbreviation MOB stands for Media Object. Value Type: Uncontrolled Text
A421	RIGHT_SUB_LICENCE_FLAG	Indicates whether the granted right can be granted by the licensee in sub-licence to third parties. Value Type: Boolean

A422	RIGHT_EXCLUSIVITY_FLAG	Indicates whether the right is exclusively granted to the licensee. Value Type: Boolean
A423	PMETA_VERSION_NUMBER	An identifier of the version for the P_Meta Definitions used in the communication. The value will be expressed by two integers separated by dot. Value Type: Formatted Code Examples: 1.0
A424	CONTENT_ALERT_DESCRIPTION	A description that alerts people about content they might not want to see/hear (or they might not want their children to see/hear). Value Type: Uncontrolled Text Examples: "Scenes of explicit sexual behaviour suitable for adults only"; "Occasional use of very strong language"; "Scenes with extreme horror effects"; "Strobing that could impact on those suffering from Photosensitive epilepsy"
A425	CONTENT_ALERT_CODE	A code identifying an alert to people about content they might not want to see/hear (or they might not want their children to see/hear), as defined by a content alert scheme communicated within the same sub-set. Value Type: External Source of referenced data: The scheme identified by the CONTENT_ALERT_SCHEME_CODE/NAME value communicated within the same sub-set.
A426	CONTENT_ALERT_SCHEME_CODE	A code identifying a content alert scheme which defines alerts to people about content they might not want to see/hear (or they might not want their children to see/hear). Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.4 CONTENT_ALERT_SCHEME_CODE
A427	CONTENT_ALERT_SCHEME_NAME	A name identifying a content alert scheme which defines alerts to people about content they might not want to see/hear (or they might not want their children to see/hear). Value Type: Uncontrolled Text Examples: "TVAnytime"
A428	MOB_DURATION	The duration, when applicable, of a Media Object instance, indicated in milliseconds, at either its canonical or intended presentation rate. The abbreviation MOB stands for Media Object. Value Type: Unsigned integer
A429	CLASSIFICATION_SCHEME_CODE	A code identifying a registered classification scheme for audio-visual works. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.2 CLASSIFICATION_SCHEME_CODE
A430	CLASSIFICATION_SCHEME_NAME	The name of a classification scheme for audio-visual works. Value Type: Uncontrolled Text Examples: TVAnytime; EBU/Escort;

A431	CLASSIFICATION_DIMENSION_CODE	A code identifying a classification dimension within a multidimensional classification scheme for audio-visual works, according to its specifications. Value Type: External Source of referenced data: The organisation responsible for the classification scheme and identified via either CLASSIFICATION_SCHEME_CODE or CLASSIFICATION_SCHEME_NAME:
A432	CLASSIFICATION_DIMENSION_NAME	The name of a classification dimension within a multidimensional classification scheme for audio-visual works. Value Type: Uncontrolled Text Examples: Intention; Content; Format; Intended Audience
A433	CLASSIFICATION_TERM_CODE	A code identifying a classification term within a classification scheme for audio-visual works, according to its specifications. Value Type: External Source of referenced data: The organisation responsible for the classification scheme and identified via either CLASSIFICATION_SCHEME_CODE or CLASSIFICATION_SCHEME_NAME:
A434	CLASSIFICATION_TERM_NAME	The name of a classification term within a classification scheme for audio-visual works. External reference: w : Value Type: Uncontrolled Text Examples: Infotainment; Western; Horror; Language Learning;
A435	PEV_AUDIENCE_RATING_PERCENTAGE	The percentage of the TV private household population or sub-population (as estimated from the sample) viewing the PEV across its duration, averaged across the minutes which comprise that period. The abbreviation PEV stands for Publication Event. External reference: EBU Escort 2.4 Value Type: Floating Point Examples: 46.3
A436	PEV_AUDIENCE_REACH_PERCENTAGE	The cumulative percentage of the TV private household population or sub-population (as estimated from the sample) viewing the PEV across its duration, cumulated across the minutes which comprise that period. The abbreviation PEV stands for Publication Event. External reference: EBU Escort 2.4 Value Type: Floating Point Examples: 61.8
A437	PEV_AUDIENCE_SHARE_PERCENTAGE	The amount of viewing obtained by the PEV expressed as a percentage of all TV viewing across the period of time of the PEV. The abbreviation PEV stands for Publication Event. External reference: EBU Escort 2.4 Value Type: Floating Point Examples: 70.0

A438	PEV_AUDIENCE_SCORE_ RECORDING_TECHNIQUE_CODE	A controlled code for the technique of recording the audience score for PEVs. The abbreviation PEV stands for Publication Event. Value Type: Controlled Code Source of referenced data: EBU/P_Meta: 9.1.24 PEV_AUDIENCE_SCORE_RECORDING_TECHNIQUE_CODE
A439	PEV_AUDIENCE_SCORE_ RECORDING_TECHNIQUE_NAME	The name of the technique of recording the audience score for PEVs. The abbreviation PEV stands for Publication Event. Value Type: Uncontrolled Text Examples: poll; people meter;
A440	ITEM_SCRIPT_DESCRIPTION	A text script for a programme item. This may be a complete production script or it may only contain the spoken text. It differs, therefore, from a media object script which is a description of a single media object in terms applicable to that media object type. Value Type: Uncontrolled Text
A441	PERSON_STAGE_NAME	The name by which a performer is known for professional purposes (e.g. Reginald Dwight performs under the name Elton John). This may be all, or part of, their real name or it could be another name entirely. A stage name may comprise of a number of parts but is represented here as a single string. It may also sometimes include a salutation. Value Type: Uncontrolled Text Examples: Sting; Madonna; Hans Moser; Bud Spencer
A442	SI_UNIT_OF_MEASURE_CODE	The symbol for a particular standard unit for the measurement of something (e.g. volume, size, weight, power, resistance, etc.) In P/Meta terms this will normally be for use in a contractual agreement. Usually this will apply to technical parameters. The prefix SI stands for "Système International d'Unités" but is generally referred to in English as the International System of Units. External reference: SI units are standardised and established by the 11th General Conference on Weights and Measures (CGPM, Conférence Générale des Poids et Mesures) in 1960. The CGPM is the international authority that ensures wide dissemination of the SI and modifies the SI as necessary to reflect the latest advances in science and technology. Value Type: External Source of referenced data: SI/BIPM (Bureau International de Poids et mesures) Examples: km; W; A; m; s; K;

A443	SI_UNIT_OF_MEASURE_NAME	The name of a particular standard unit for the measurement of something (e.g. volume, size, weight, power, resistance, etc.) In P/Meta terms this will normally be for use in a contractual agreement. Usually this will apply to technical parameters. The prefix SI stands for "Système International d'Unités" but is generally referred to in English as the International System of Units. External reference: SI units are standardised and established by the 11th General Conference on Weights and Measures (CGPM, Conférence Générale des Poids et Mesures) in 1960. The CGPM is the international authority that ensures wide dissemination of the SI and modifies the SI as necessary to reflect the latest advances in science and technology. Value Type: Uncontrolled Text Examples: metre; kilogram;
A444	SI_UNIT_OF_MEASURE_QUANTITY	A numeric value representing a particular amount of units specified by the accompanying unit of measure attribute. NB this attribute must be used in conjunction with either SI_UNIT_OF_MEASURE_CODE or SI_UNIT_OF_MEASURE_NAME. Value Type: Formatted Code
A445	ITEM_START_SCRIPT_CUE_TEXT	Text which provides the cue for the start of an item. Typically, this will be spoken dialogue by a performer or presenter, but it could refer to another action such as a stage direction or piece of music which marks the beginning of the item. Value Type: Uncontrolled Text
A446	ITEM_END_SCRIPT_CUE_TEXT	Text which provides the cue for the end of an item. Typically, this will be spoken dialogue by a performer or presenter within the item, but it could refer to another action such as a stage direction or piece of music which marks the end of the item. Value Type: Uncontrolled Text
A447	PGR_ORDERED_FLAG	This indicates whether within the programme group the individual programmes or episodes have to be ordered for giving the correct meaning to the programme group. The abbreviation PGR stands for Programme Group. Value Type: Boolean

4.5.2 Conclusion

P_Meta makes use of other standards and schemas as well as using successfully the practical experience of the members of the developing body. A large amount of elements had been added to provide a entire description-base for the essence/media used, targeting at the exchange of essence and metadata in a B2B scenario.

As the developing consortium points out, "P_META is designed to be as flexible as possible in implementation while retaining consistency of meaning. - It is not intended for use in databases (!) (although it could be used as a starting point), but is designed to be the "language on the wire" in implementation."[23]

Using numerical codes for attributes and standard values, allowing data exchanges to be translated into different languages without corruption, is another advantage of this standard

[23] http://www.ebu.ch/CMSimages/en/tec_t3295_tcm6-11463.pdf

4.5.2.1 Main entities in P_META

The main entities in P_META are:

- Programme
- Programme-Group
- Item (Programme-item)
- Media Object (MOB)
- Brand

4.5.2.2 Standards and Exchange

P_META makes reference to other standards from standardising bodies like ISO, SMPTE, EBU, IEFT/RFC.

As been pointed out by the group of developers P_META is not meant to be used mainly in Databases but has been designed to act as a “Metadata-Exchange-Standard” supporting the exchange between the different members of EBU and third parties as well.

5 Overall Conclusion

As already mentioned in the Overview-chapter of this document the main purpose of this document changed from a mere survey and short analysis of existing data-models, functioning as a discussion-base for the discussion over a PrestoSpace-Data-model, to a) providing an overview what PrestoSpace has to expect as legacy metadata and which of the existing and analysed standard formats can be acting as candidate for input; b) showing up the features PrestoSpace has to consider when working on the final PrestoSpace schema.

This little survey draw the bow from “simple” listing-models to complex relational models. At this point no appraisal or voting for or against a single approach will be given, as every model or schema has been developed and designed under different requirements and preconditions, sometimes to fulfil special needs and tasks.

Beneath others one thing showed up: the adaptation and matching of data coming from one model (may it be a simple “listing-model”, a fully developed relational model, an old legacy model, a newly developed one, etc.) to another model will be a time-consuming, rather expensive and even sometimes unsuccessful process, as explicit and/or satisfactory matching is not always possible. Even when succeeding, notable manual efforts during migration has to be taken into account.

6 MAPPING

6.1 Introduction

The mapping-process was in some points quite controversial and vivid. So the reader of this paper should feel free to add comments and proposals to the mapping done as well as new candidates (Models, schemata, etc) to be mapped too.

The main purposes of this chapter is to act as a discussion-contribution as well as a tool for the decision process for a PS-data-model.

6.2 Mapping of entities

ENo.	RAI-DM	INA-DM	FARAO	DR-DM	SMEF	P_META	…
1	Collection				Programme-Group	Programme-Group	
2	Program	Programme + Progr. EDOB	Program	Main/Single Production	Programme (EDOB_Programme)	Programme	
3	Item	Programme Item	Program Item (“Contribution”)	Item	Programme Item (EDOB_Item)	Item	
4				MOB (Media Object)	MOB (Media Object)	MOB (Media Object)	
5				MIN (Media Object Instance)	MOI (Media Object Instance)		
6					Brand	Brand	

Entity the essence/material is linked to

6.3 Mapping of attributes

Field-number	Attributes in Common	Dublin Core	FARAO	P_META (EBU)	(Open) SMEF (BBC)	MXF (non official)	MXF DMS1	SMPTE	DR Metadata standard
1	***Title***	**1**	**4**	A59,A61,A99, A107, A110, A114,A146,A198,A401	10.2, 15.2,17.1	**1**	TBC	TBD	5.3
2	***Creator***	**2**	**18**	A81,A82,A83,A87,A88,A89,A90,A125, A254, A255,A256,A413, A414	20,11,12,13	**2**	TBC	TBD	16.1
3	***Subject and Keywords***	**3**	**10**	Difficult to define, Coverage could be used here	/	**3**	TBC	TBD	/
4	***Description***	**4**	**8**	/	5.4	**4**	TBC	TBD	5.5
5	***Publisher***	**5**	**/**	A81,A82,A83,A87,A88,A89,A90,A125,A254,A255,A256,A413,A414	20,11,12,13	**5**	TBC	TBD	5.6
6	***Contributor***	**6**	**9**	A11,A81,A82,A83,A87,A88,A89,A90,A125,A254,A255, A256,A413, A414	20,11,12,13	**6**	TBC	TBD	/
7	***Date***	**7**	**24**	A152,A217,A218,A219,A22, A367,A368,A405	5.2	**7**	TBC	TBD	5.7
8	***Resource Type***	**8**	all fields in **TECHNICAL DATA**	A12,A13,A226	25	**8**	TBC	TBD	5.16
9	***Format***	**9**	all fields in **CARRIER DATA**	A72,A73,A222, A361	8,9,23	**9**	TBC	TBD	5.11
10	***Resource Identifier***	**10**	**28**	A105	/	**10**	TBC	01.01.01.00.00.00.00.00 01.01.02.00.00.00.00.00 01.01.03.00.00.00.00.00 01.01.04.00.00.00.00.00 …. 01.01.0F.00.00.00.00.00 01.01.15.00.00.00.00.00 01.01.15.40.00.00.00.00 01.01.15.40.01.00.00.00	5.1 (9.1 / 11.1 / 18.1 / 23.1)1

11	***Source***	**11**	all fields in **STORAGE DATA**	A223,A224	/	**11**	TBC	TBD	5.14
12	***Language***	**12**	/	A21,A22,A65,A66,A141, A407,A415, A416,A417, A418	/	**12**	TBC	TBD	6.1
13	***Relation Resources***	**13**	/	/	/	**13**	TBC	TBD	/
14	***Coverage***	**14**	12	A1,A9,A21, A22,A38,A67, A123,A141, A207,A214	21,28	**14**	TBC	TBD	5.19
15	***Rights Management***	**15**	all fields in RIGHTS	A14, A15, A18, A19, A20, A116, A117,A118, A119, A120, A121, A122, A162, A200, A201,A202, A203, A204, A205, A206, A212, A421, A422	4,18	**15**	TBC	TBD	5.13
16	***Storage History***	/	all fields in STORAGE DATA	A135, A136, A137, A154, A158	/	/	TBC	TBD	5.9
17	***Carrier Data***	/	all fields of CARRIER DATA	A25, A26, A27, A168, A169, A227	25	**20**	TBC	TBD	5.11
18	***Recording Data***	/	all fields of RECORDING DATA	A12, A29, A30, A31, A32, A33, A34, A35, A36, A37, A148, A149	21,22	**19**	TBC	TBD	5.11
19	***Technical Information***	/	all fields of TECHNICAL DATA	A41, A42, A43, A44, A46, A47, A48, A49, A50, A51, A53, A54, A103, A131, A132, A174, A175, A177, A178, A180, A181,A183, A184, A185, A209, A210, A257, A258, A259, A419	1,5,8,9,23,	/	TBC	TBD	5.10

6.4 Mapping Entities – Attributes

ENo	FARAO	DM	SMEF	P_META	
1		18(?), 27(?)	15	A198, A447	
2	1, 3-5,16, 20, 21, 28-30, 32-45, 47-60	10(?), 17	5(?), 16, 17	A96-A101, A104, A107, A109, A110, A114, A146, A225	
3	2, 6-15, 17-19, 22-27, 31, 46, 61, 62	5	21	A59-A62, A148, A149, A440, A445, A446	
4		9	1, 10	A77, A78, A166, A188-A196, A216, A245, A247, A420, A428	
5		11	2		
6			3	A10	

7 Links and Resources

http://www.oclc.org

http://www.ncsa.uiuc.edu

http://www.dublincore.org

http://www.rai.it

http://www.ina.fr

http://www.orf.at

http://www.dr.dk

http://www.bbc.co.uk

http://palmm.fcla.edu/strucmeta/mxf2mods.htm

http://www.smpte.org

http://www.smpte-ra.org/mdd/RP210v8-final-040810MC.xls

http://www.schemas-forum.org/registry/desire/activityreports.php3?field=filename&value=AV-smpte.rdf

http://www.ebu.ch/en/technical/publications/tech3000_series/tech3295/index.php?display=EN

http://www.ebu.ch/en/index.php

http://www.ebu.ch/CMSimages/en/tec_t3295_tcm6-11463.pdf

http://www.lookuptables.com

8 APPENDIX A

8.1 Detailed Description conceptual model IMMIX pilot system

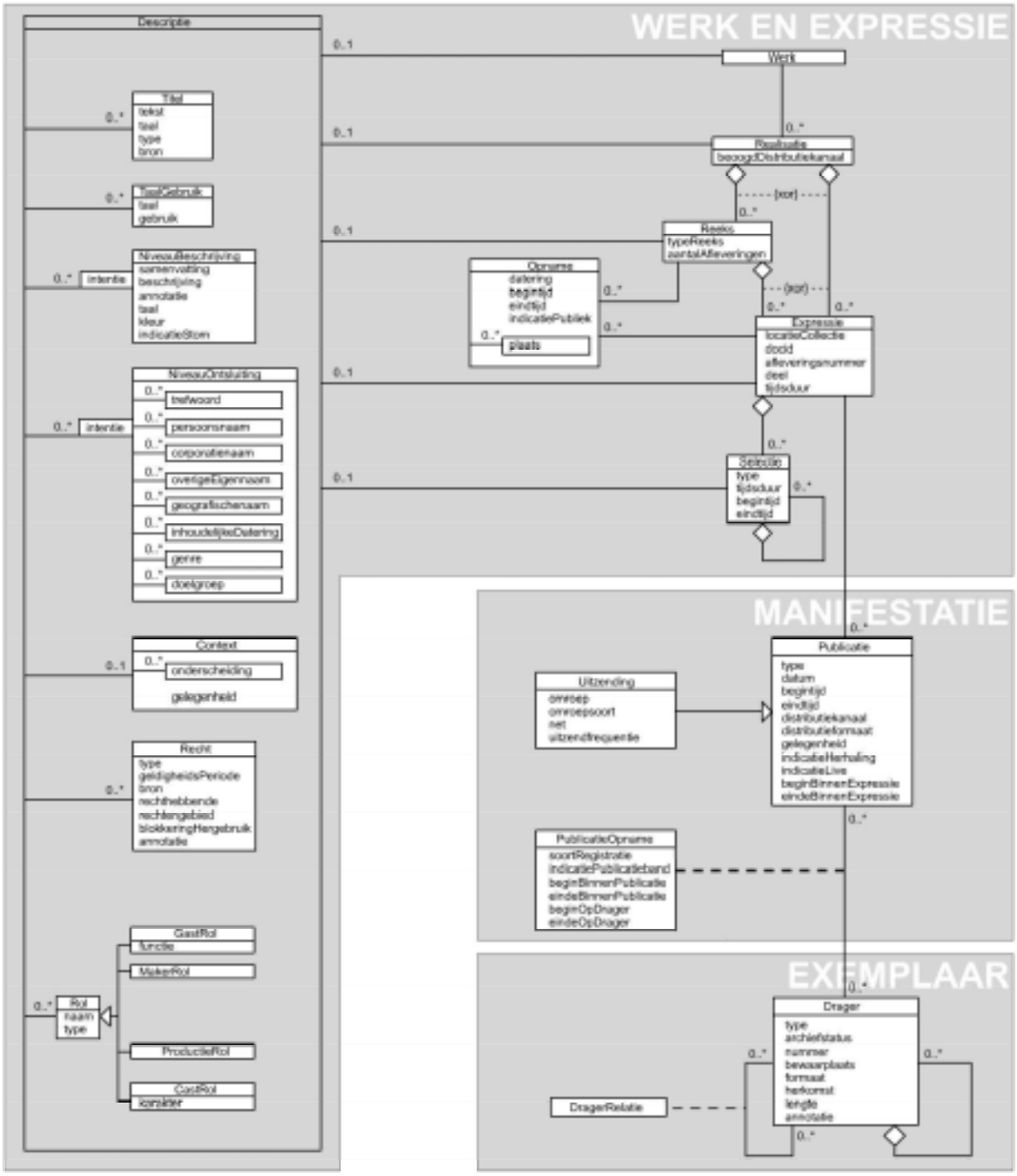
In this chapter a description is given of the conceptual metadata model that was the basis for the iMMix pilot system, the new multimedia catalogue of Beeld en Geluid. It represents the first step towards the model that is actually implemented at this point in time and of which a diagram is given above. The final metadata model (and data model) is still work in progress. For example in the near future the rights metadata and technical metadata will be extended and some minor other improvements are made.

The first paragraph shows the metadata model as a UML diagram; and makes explicit in which way the different dimensions Work, Expression, Manifestation and Exemplar are divided over the model.

In the next paragraphs the UML-diagram is textually explained. Each class has a textual description structured as follows:

- Description: description of the meaning of the class, eventually with examples.
- Relations: enumeration of the relations of this specific class with other classes.
- Attributes: a list of attributes per class. Each attribute is explained shortly and if possible examples are given. For some attributes a list of possible values is given. If there is an * placed in front of an attribute, it means the attribute can be repeated.

8.1.1 UML-diagram of pilot system



8.1.2 The dimensions Work and Expression

Work (Werk)

Description

A Work is the intellectual and artistic *concept* (or idea) that is the basis of zero or more Realisations. A Work is the highest level within the hierarchy of metadata. Please notice that it is not an individual programme, but the concept behind it. A work is not something concrete that we can grasp.

Examples are:

- the concept 'Soldaat of Oranje';
- the concept 'In de Vlaamse Pot';
- the concept 'Het NOS Journaal'.

Relations

Related to	0 or 1	Descriptions
Concept of	0 or more	Realisations

Attributes

-

Realisation

Description

A Realisation is a concrete elaboration of a Work. It describes the more general data or parts of which the Realisation consists.

Examples of Realisations of the Work concept 'Soldaat van Oranje';

- the series 'Soldaat van Oranje'
- the movie 'Soldaat van Oranje'

Relations

Related to	0 or 1	Description
Belongs to	exactly 1	Work
Exists of	0 or more	Series
Or of	0 or more	Expressions

Attributes	IntendedDistributionchannel *(BeoogdDistributiekanaal)*		Distribution channel for which the video, film, audio, or text is originally intended (television, radio, internet, and cinema).

Series (Reeks)

Description

A Series is a collection or grouping of Expressions. A season of a drama series is an example of such a Series. Also logically related Expressions can form a Series. 'All episodes of Netwerk, a current affairs programme, broadcasted by the AVRO form together a logical Series. Information that is valid for all Expressions within a Series, can be annotated at this level.

Relations	Related to	0 of 1	Description(s)
	Related to		Recording(s)
	Part of	0 or more	Realisation
	Consists of		Expressions
Attributes	*typeSeries (Reeks)*		Type of the Series (f.e. a series, group, Series, etc).
	AmountOfEpisodes (aantalEpisoden)		Amount of episodes belonging to one Series.

Expression (Expression)

Description

An Expression is the actual 'programme'.

Examples are:

- the movie ‘Soldaat van Oranje’;
- an episode of the television series ‘Soldaat van Oranje’;
- episode 10 of the series ‘In de Vlaamse Pot’;
- the NOS Daily News of Wednesday December 29, 2004;
- a documentary about sharks.

Relations	Related to	0 or 1	Description(s)
	Related to		Recording(s)
	Part of		Realisation

	OR Consists of	0 or more	Series (Reeks) Selections
	Related to		Publications
Attributes	*(LocationCollection)* *LocatieCollectie*		The original Name of the collection of which the original Description, i.c. the Description as in the legacy database, is derived
	DocID (docid)		The original Document ID from the legacy database
	Episodenumber (episodenummer)		Number of the episode, in case the Expression is part of a Series (Reeks).
	Part (deel)		Ranking number of the programme, in case the programme is described in several parts. Instead of an original Expression mostly the broadcast is described. It is possible for one programme to be broadcast in different parts. To facilitate migration from legacy databases it is possible to describe (using *part*) several Expressions originally meant to be one and the same Expression.
	Duration (tijdsduur)		Duration in units of time, if possible automatically derived.

Selection (Selectie)

Description

A Selection is part of an Expression. The Description at the Selection-level gives information about a specific part of the Expression.

Examples of a Selection are:

- the Item of the NOS Daily News of Wednesday 2 January 2002 in which is explained how much money the Dutch have spend on Christmas;
- a Shot within a documentary about sharks with a close-up of the skin.

Relations	Related to	0 of 1	Description
	Part of	0 or more	Expression
	Consists of		Selections

Attributes	*Type*	Examples of characterizations for a Selection are: • Item: a closed, edited topic in f.e. a current affairs programme on which something very specific can be annotated; (selection is in this case in fact made by the editor) • Shot: a continuous take of one object, event or person about which or whom something specific can be annotated; • Sequence: a by the documentalist selected longer fragment of images or sound on which something specific can be annotated.
	Duration (tijdsduur)	Duration in units of time, if possible automatically derived.
	Starting time (begintijd)	Starting point within the Expression on which the Selection starts, if possible Related to a keyframe.
	Endtime (eindtijd)	Point in time within the Expression at which the Selection ends, if possible Related to a keyframe.

8.1.3 The classes Recording (Opname) and Description

Recording (Opname)

Description	*Recording* (Opname) groupes information about the recording or a Series (Reeks) or an Expression. It gives information about the exact time and location a recording is made at.		
Relations	Related to or to	1	Series (Reeks) Expression
Attributes	*Date (Datering)*		Date or period in which the recording is made.
	Start time (Begintijd)		Point in time the recording started.
	End time (Eindtijd)		Point in time the recording ended.
	IndicationAudience (indicatiePubliek)		Indication: audience joining the recording or not
	** location (plaats)*		Location of the recording.

Description (Descriptie)

Description	Description clusters all descriptive parts that can be linked to all classes of the dimensions WORK and EXPRESSION (these are: Work, Realisation, Series (Reeks), Expression and Selection). This class has been created to make the model more comprehensive and has as class in itself no meaning.		
Relations	Related to	exactly 1	Work (Werk)
	or		Realisation (Realisatie)
	or	exactly 1	Series (Reeks)
	or	exactly 1	Expression (Expressie)
	or		Selection (Selectie)
		0 or more	
	Consists of	0 or more	Titles
	Consists of	0 or more	UsedLanguage (TaalGebruik)
	Consists of		LayerDescription (NiveauBeschrijving)
	Consists of	0 or 1	LayerIndexing (NiveauOntsluiting)
	CONSISTS OF	0 or more	Context
	Consists of	0 or more	Rights
	Consists of		Roles
Attributes	-		

Title (Titel)

Description	A Title is a given name at a certain level of description. The title consists of the name, the type of title en the origin of the title.	
Relations	Part of	Description
Attributes	*Text*	Titletext.
	Language (Taal)	Language of the title.

	type	Type of title. Possible values: • *Title* • *Subtitle* • *Working title* • *Given title* • *Other titles* • *Short Title*
	source	Source of the title

UsedLanguage (TaalGebruik)

Description	UsedLanguage (TaalGebruik) describes the used language.	
Relations	Part of	Description
Attributes	*Language (Taal)*	Name of the language.
	Use (gebruik)	The way the language is used. Possible values: • *Dubbing* • *Subtitling* • *Original language* • *Other languages/dialects* • *Adaptation*

LayerDescription (NiveauDescription)

Description	LayerDescription describes the subject of the audiovisual material (the content). NiveauDescription is being identified by means of an *Intention.* The *intention* indicates the target group of the Description, f.e. broadcast professional, museum, education, etc.	
Relations	Identification via	*intention*
	Part of	Description
Attributes	*Summary (samenvatting)*	Summary (short Description) of the content
	Description	OF THE CONTENT (TOPIC)
	annotation	Additional information on the theme/subject.

Language	The language this LayerDescription (NiveauBeschrijving) is written in. Default value is "Dutch".
colour	Use of colour, black/white, tinted, coloured by hand, etc.
indicationSilent	Indication if production is a so-called 'silent movie'.

LayerIndexing

Description

LayerIndexing describes the story of the audiovisual material (non-formal description). The attributes in this class are strongly related to the Beeld en Geluid thesaurus.

LayerIndexing is identified via an *Intention.* The *intention* indicates for whom and what the Description is meant, f.e. professionals, museum, education, etc.

Relations

Identification	*intention*
by means of	
	Description
Part of	
* *keyword (trefwoord)*	Subject.
* *personName (persoonsNaam)*	Name of a person.
**corporationName (corporatieNaam)*	Name of an organisation, corporation or company, etc.
* *other Names (overigeEigennamen)*	Other Names (not persons-, corporations-, or geographical names), f.e. names of fictional characters.
**geographicalName (geographischeNaam)*	Name of a location.
* *inhoudelijkeDatering*	Date or period referring to the story (content) within the program.
* *genre*	Classification of the production that can denote the form, the style, or the type of the program.
* *target group*	Target group, f.e. children, women, etc.

Context

Description

Context consists of data belonging to the dimensions WORK and EXPRESSION, but doesn't describe features related to the subject/content.

Relations	Part of	Description
Attributes	* *Award (onderscheiding)*	Name of an award or prize won by a production
	Event (gelegenheid)	Specific Event for which the production is made (Event for broadcasting is annotated at the class Publication).

Rights (Recht)

Description Rights (Recht) describes the rights for a specific part.

Relations	Part of	Description
Attributes	*type*	Type of rights, f.e. reuse, broadcast rights, etc
	Period (geldigheidsPeriode)	Period in which the rights are valid
	Source (bron)	Person, organisation, corporation or company the program is coming from
	Rights owner (rechthebbende)	Person, organisation, corporation or company that holds the rights
	RightsTerritory (rechtengebied)	Territory for which the rights apply, f.e. The Netherlands, Europe, etc.
	Blocking (blokkeringHergebruik)	Information about the reuse of the material
	annotation	Additional information on rights

Role (Rol)

Description A Role describes the role of a person, organisation, corporation or company. There is a distinction in GuestRole, MakerRole, ProductionRole and CastRole.

Relations	Part of	Description
	Generalisation of	GuestRole
	Generalisation of	MakerRole
	Generalisation of	ProductionRole
	Generalisation of	CastRole

Attributes	*Name*	Name of the person, organisation, corporation or company.
	type	Type of the role.

GuestRole (GastRol)

Description	A GuestRole is a specific form of a Role; a person, an organisation, corporation or company acts as a guest.	
Relations	Specialisation of	Role
Attributes	*(Name)*	
	(type)	Possible values: • Speaker • Performer
	Position	The position of the guest.

ContributorRole (MakerRol)

Description	ContributorRole (MakerRol) describes the responsibility for the conceptual, artistical and/or technical realisation of the material	
Relations	Specialisation of	Role
Attributes	*(Name)*	
	(type)	Possible values: • Dubbing • Composer • Textual editor • Textwriter • Camera • Director • Comments • Interviewer • Editing • Presentation • Production • Composition • Script • Etc

ProductionRole (ProductieRol)

Description	A Producer is the person, organisation, corporation or company responsible for the production.	
Relations	Specialisation of	Role
Attributes	*(Name)*	
	(type)	Possible values: • Commissioner • Producer • Sponsor • Subsidiser

CastRole (CastRol)

Description	CastRole describes the role of a member of the cast.	
Relations	Specialisation of	Role
Attributes	*(Name)*	*cast member*
	(type)	Standard f.e. actor/actrice or member of the cast.
	character (karakter)	Name of the character or role

8.1.4 The dimensions Manifestation and Exemplar

Publication (Publicatie)

Description	An Expression can be 'published', by means of publication. The class Publication contains information on the publication of an Expression. A way of Publication is the play-out of a television- or radio program. Other possibilities are publication via the Internet or publication on DVD or CD. An Expression can be published in several parts. Each episode is interpreted as a separate publication (broadcast). In Publication will be indicated which part of the Expression of the programme will be published.	
Relations	Belongs to	Expression
	Related to	Carriers
		PublicationRecording (PublicatieOpname)
	Generalisation of	Broadcast (Uitzending)

Attributes	*Type*	Type of publication, f.e. broadcast, release, review, etc.
	Date	Date of publication. (Or: number of the week, in the case of Polygoon/Profilti-material)
	Start time	Point in time the publication started.
	Endtime	Point in time the publication ended.
	distribution channel	The communication system used for the transfer of the material from transmitter to receiver, f.e. television, radio, cinema, internet, trade.
	distribution format	Format of the publication, f.e. PAL, DVB, DAB, widescreen, DVD, CD-ROM, etc.
	Event (gelegenheid)	Event in the scope of which the publication takes place.
	IndicationRepeat	Indication if the publication is a repeat.
	IndicationLive	Indication if the publication was a live event.
	beginWithinExpression	Point in time within the Expression at which the Publication starts (f.e. a television program can be broadcast in two parts).
	EndWithinExpression(eindeBinnenExpressie)	Point in time within the Expression at which the Publication ends.

Broadcast (Uitzending)

Description	A Broadcast is a special form of Publication, f.e. of a television or radio programme.	
Relations	Specialisation of	Publication
Attributes	*Broadcaster (omroep)*	The name of the broadcaster or the cooperation of broadcasters responsible for the broadcast.
	BroadcastType (omroepsoort)	Description of the type of broadcaster, f.e. national, regional, foreign, local, commercial, public.
	Channel (net)	Name of the distribution channel, f.e. Ned 1, Ned 2, Ned 3, Radio 1, Radio 2, etc.
	Broadcastfrequency (uitzendfrequentie)	Indication of the frequency with which the Expressions are broadcasted. This is annotated for each Broadcast.

PublicationRecording(PublicatieOpname)

Description	PublicationRecording (PublicatieOpname) describes the recording (registration) of a Publication on one or more Carriers (Dragers)	
Relations	Relation between	Publication
	and	Carrier
Attributes	*RegistrationType*	PUBLICATION IS RECORDED.
	(soortRegistratie)	
	indicationPublicationtape	Indicates if the Carrier (Drager) has been used for the Publication (Publicatie).
	(indicatiePublicatieband)	
	startWithinPublication (beginBinnenPublicatie)	Startingpoint within the Publication of the part that is on the Carrier (one publication can be divided over two or more carriers)
	endWithinPublication (eindeBinnenPublicatie)	Endtime within Publication of the part that is on the Carrier.
	startOnCarrier (beginOpDrager)	Point in time on the Carrier at which the specific part starts.
	endOnCarrier (eindeOpDrager)	Point in time on the Carrier at which the specific part starts.

Carrier (Drager)

Description	A Carrier describes a carrier on which one or more (parts of) publications are kept. A Carrier can consist out of different Carriers; a movie can be divided over several reels of film. It's also possible that there is a relation between two Carriers, which is described by a CarrierRelation .	
Relations	Related to	Publications
		PublicationRecording
	Related to	Carriers
		CarrierRelation
	Exists of	Carriers
Attributes	*Type (type)*	Type of carrier, f.e. Digibeta, Ampex, CD, DAT, server, etc.
	Archive status (archiefstatus)	Status of the carrier in the archive, f.e. viewing copy, master, deposit exemplar.
	Number (nummer)	Identification number of the carrier.
	Depository (bewaarplaats)	Physical location of the carrier.
	Format (formaat)	Format of the carrier, f.e. 8 mm, MPEG-1, etc.

Origin (herkomst)	ERSON, ORGANISATION, CORPORATION OR T HAS OFFERED, GIVEN, DONATED, OR E CARRIER.
Length(lengte)	The length or volume of the carrier.
Annotation (annotatie)	Additional information about the carrier.

CarrierRelation/DragerRelatie

Description	A CarrierRelation describes the relation between two Carriers.	
Relations	Relation between	Carrier
	and	Carrier
Attributes	*?*	Attributes for this class are under development.

8.2 INA-DM: Description fields

Fields	Types of fields	Type of index	Thesaurus	Index multifields associated	Format and content
Identification data					
Id notice	text	unique			number of the notice
Title of document	text	inclusive	Index titles		Individual Title of the document
Title of collection	text	inclusive	Index titles		Collection Title when each document is independant
Title of programming session	text	inclusive	Index titles		Title of the hourly session
Title of series	text	inclusive	Index titles		Collection Title when documents are following (serial, documentary with several parts...)
Date of transmission	text	exact			Date of transmission of the program Format dd.mm.yyyy
Duration	text				format hhmmss
Production type (code)	text	exact			Code giving legal feature according to producer or broadcaster
Producers	text	exact		Table of contents	Content : producer role, production location, production company name and year of production

Authors (patronymic)	text	inclusive	thesaurus PP PM	Index Keywords	Terms from lexicon (without indexing complement)
Genre	text	exact		List of authority	Typology of the domain covered (information, sport, history, music, technology…) from 1 to 3 terms
Form	text	exact		List of authority	Typology of audiovisual production genre (news, documentary, interview, serial, show…) from 1 to 5 terms
Technical data about contents and media					
Principal Descriptors or keywords (index)	text		thesaurus	Index keywords	Term from thesaurus Used to feature the principal content information. By default the request is done on neighbour terms (sons and synonyms)
Secondary Keywords	text		thesaurus	Index Keywords	Term from thesaurus Used to feature the secondary thematic content information. By default the request is done on neighbour terms (sons and synonyms)
Cast (patronymic)	text	inclusive	thesaurus PP PM	Index Keywords	Terms from the lexicon First and second names of presentators, guests, interpreters and technical staff
Person Keywords	text		thesaurus	Index Keywords	Term from thesaurus Used to feature the Person names information. By default the request is done on neighbour terms (sons and synonyms)

Location Keywords	text	inclusive	thesaurus	Index Keywords	Term from thesaurus Used to feature the Location where the event occurs. By default the request is done on neighbour terms (sons and synonyms)
Works	text	inclusive		Index Texts	In natural language. Interpreters and works.
Content list	text	inclusive			List of the program elements with title, start hour and duration.
Abstract	text	inclusive			In natural language. First level of summary synthetic and thematic
Synopsis	text	inclusive		Index Texts	In natural language. Summary of the document
Sequences	text	inclusive		Index Texts	In natural language. List of the sequences or shots which may be reused and shot format.
Date of shooting	text	exact			Date of shooting or recording
Date of retransmission	text	exact		Index Dates	Date(s) of retransmission of the program
Hour of transmission	text	exact			format hhmmss
Color	text	exact			Three values: color, B&W or mixed
Material	text				Description of the professional carriers
Version short / long	text	exact			Code indicating the type of version (VS, VL)
Version original / foreign	text	exact			Code indicating the type of version (VE, VO)

Notes	text	inclusive			In natural language. Various contents : legal information, awards...
Notes on legal issues	text				Notes on INA's property rights
Public attendance	text	exact			Typology about public attending
Public destinataire	text	exact			Typology about public target
Thesaurus	text	exact	Thesaurus		Termes des champs auteur, générique et descripteurs hors précisions
Type of notice (code)	text	exact			Coded value indicating which kind of notice (for a program, principal or abstract, or part of serial)
Librarian	text	exact			code on 3 caracters
Date of modification	text	exact		Index Dates	Date of last modification of the notice Form dd.mm.yyyy
Link to principal notice	text	exact			Number of the principal notice to make links with its related sub-documents