NoTube Networks and ontologies for the transformation and unification of broadcasting and the Internet FP7 – 231761

D2.1 Requirements analysis

Coordinators: Gerhard Stoll, Peter Altendorf, Rico Zimmermann, Martin Riethmayer

With contributions from: Alex Conconi, Raoul Monnier, Libby Miller, Gerhard Stoll, Peter Altendorf, Rico Zimmermann, Martin Riethmayer, Matthias Elser, Michael Probst.

Quality Assessor: Fulvio Negro

Quality Controller: Lora Aroyo

<table>
<thead>
<tr>
<th>Document Identifier</th>
<th>NoTube/2010/D2.1/v2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version</td>
<td>V2.0</td>
</tr>
<tr>
<td>Date</td>
<td>23. February 2010</td>
</tr>
<tr>
<td>State</td>
<td>Final</td>
</tr>
<tr>
<td>Distribution</td>
<td>open</td>
</tr>
</tbody>
</table>
Executive Summary

This deliverable in its second version is dedicated to provide an overview of existing file formats and metadata models which are used in professional and non-professional domains and its application in the NoTube service scenarios. Furthermore the need of an internal metadata format for the NoTube environment is described. To be able to specify an appropriate metadata format the requirements for such a model are given in this version. The task of a metadata transformation between different metadata models, which are needed for the various use cases is outlined in this document as well.

This document is the revised version of the deliverable D2.1 which was finalised in M3. Due to the early stage of the project at that time, several points of the initial deliverable had to be substantiated.
# Document Information

<table>
<thead>
<tr>
<th>IST Project Number</th>
<th>FP7 - 231761</th>
<th>Acronym</th>
<th>NoTube</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full Title</strong></td>
<td>Requirements analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Project URL</strong></td>
<td><a href="http://www.notube.eu/">http://www.notube.eu/</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Document URL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EU Project Officer</strong></td>
<td>Leonard Maqua</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Deliverable Number</th>
<th>2.1</th>
<th><strong>Title</strong></th>
<th>Requirements analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Work Package Number</strong></td>
<td>2</td>
<td><strong>Title</strong></td>
<td>TV Metadata Interoperability</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date of Delivery</th>
<th>Contractual</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M 13</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Status</th>
<th>version 2.0</th>
<th>final □</th>
</tr>
</thead>
</table>

**Nature**
- prototype □
- report □
- dissemination □

**Dissemination level**
- public □
- consortium X

<table>
<thead>
<tr>
<th>Authors (Partner)</th>
<th>Institut für Rundfunkechnik GmbH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author</strong></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Alex Conconi</td>
</tr>
<tr>
<td>E-mail</td>
<td><a href="mailto:alex.conconi@txt.it">alex.conconi@txt.it</a></td>
</tr>
<tr>
<td><strong>Partner</strong></td>
<td>TXT Polymedia</td>
</tr>
<tr>
<td>Phone</td>
<td></td>
</tr>
</tbody>
</table>

| **Author**        |                                  |
| Name              | Raoul Monnier                      |
| E-mail            | raoul.monnier@thomson.n           |
| **Partner**       | Thomson Grass Valley France SA    |
| Phone             |                                  |

| **Author**        |                                  |
| Name              | Libby Miller                      |
| E-Mail            | libby.miller@bbc.co.uk            |
| **Partner**       | BBC                               |
| Phone             |                                  |

| **Author**        |                                  |
| Name              | Gerhard Stoll                     |
| E-mail            | stoll@irt.de                      |
| **Partner**       | Institut für Rundfunkechnik GmbH |
| Phone             |                                  |

| **Author**        |                                  |
| Name              | Peter Altendorf                   |
| E-mail            | altendorf@irt.de                  |
| **Partner**       | Institut für Rundfunkechnik GmbH |
| Phone             |                                  |

| **Author**        |                                  |
| Name              | Rico Zimmermann                   |
| E-mail            | zimmermann@irt.de                 |
| **Partner**       | Institut für Rundfunkechnik GmbH |
| Phone             |                                  |

| **Author**        |                                  |
| Name              | Martin Riethmayer                 |
| E-mail            | riethmayer@irt.de                 |
| **Partner**       | Institut für Rundfunkechnik GmbH |
| Phone             |                                  |

| **Author**        |                                  |
| Name              | Matthias Elser                    |
| E-mail            | elser@irt.de                      |
| **Partner**       | Institut für Rundfunkechnik GmbH |
| Phone             |                                  |

| **Author**        |                                  |
| Name              | Michael Probst                    |
| E-mail            | probst@irt.de                     |
| **Partner**       | Institut für Rundfunkechnik GmbH |
| Phone             |                                  |
Abstract
(for dissemination)

This deliverable is dedicated to provide an overview of existing file formats and metadata models which are used in professional and non-professional domains. Furthermore the need of an internal metadata format for the NoTube environment is described. To be able to specify an appropriate metadata format the requirements for such a model are given in this revised version. The task of a metadata transformation between different metadata models is outlined in this document as well.

Keywords
File Formats, Container Formats, Metadata Models, Transformations, Interoperability

Version Log

<table>
<thead>
<tr>
<th>Issue Date</th>
<th>Rev. No.</th>
<th>Author</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>23.04.09</td>
<td>1</td>
<td>M. Riethmayer</td>
<td></td>
</tr>
<tr>
<td>28.04.09</td>
<td>1.1</td>
<td>R. Zimmermann</td>
<td>Added introduction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Extended chapter 3.1.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Added Author</td>
</tr>
<tr>
<td>28.04.09</td>
<td>1.2</td>
<td>M. Riethmayer</td>
<td>Merged review informations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Extend 10 - 12 to reflect comments by RAI NoTubeTeam (Luca Vignaroli)</td>
</tr>
<tr>
<td>05.02.10</td>
<td>2.0</td>
<td>M. Riethmayer</td>
<td>Updating Deliverable for second version</td>
</tr>
<tr>
<td>Participants</td>
<td>Contact</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vrije Universiteit Amsterdam</td>
<td>Guus Schreiber&lt;br&gt;Phone: +31 20 598 7739/7718&lt;br&gt;Email: <a href="mailto:schreiber@cs.vu.nl">schreiber@cs.vu.nl</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>British Broadcasting Corporation</td>
<td>Libby Miller&lt;br&gt;Phone: +44 787 65 65 561&lt;br&gt;Email: <a href="mailto:Libby.Miller@bbc.co.uk">Libby.Miller@bbc.co.uk</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asemantics SRL Uninominale</td>
<td>Alberto Reggiori&lt;br&gt;Phone: +39 0639 7510 78&lt;br&gt;Email: <a href="mailto:alberto@asemantics.com">alberto@asemantics.com</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engin Medya Hizmetleri A.S.</td>
<td>Ron van der Heiden&lt;br&gt;Phone: +31 6 2003 2006&lt;br&gt;Email: <a href="mailto:ron@engin.tv">ron@engin.tv</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Institut fuer Rundfunktechnik GmbH</td>
<td>Christoph Dosch&lt;br&gt;Phone: +49 89 32399 349&lt;br&gt;Email: <a href="mailto:dosch@irt.de">dosch@irt.de</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ontotext AD</td>
<td>Atanas Kiryakov&lt;br&gt;Phone: +35 928 091 565&lt;br&gt;Email: <a href="mailto:naso@sirma.bg">naso@sirma.bg</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open University</td>
<td>John Domingue&lt;br&gt;Phone: +44 1908 655 014&lt;br&gt;Email: <a href="mailto:j.b.domingue@open.ac.uk">j.b.domingue@open.ac.uk</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAI Radiotelevisione Italiana SPA</td>
<td>Alberto Morello&lt;br&gt;Phone: +39 011 810 31 07&lt;br&gt;Email: <a href="mailto:a.morello@rai.it">a.morello@rai.it</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantic Technology Institute International</td>
<td>Lyndon Nixon&lt;br&gt;Phone: +43 1 23 64 002&lt;br&gt;Email: <a href="mailto:lyndon.nixon@sti2.org">lyndon.nixon@sti2.org</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stoneroos B.V.</td>
<td>Annelies Kaptein&lt;br&gt;Phone: +31 35 628 47 22&lt;br&gt;Email: annelies.kaptein@stoneroos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thomson Grass Valley France SA</td>
<td>Raoul Monnier&lt;br&gt;Phone: +33 2 99 27 30 57&lt;br&gt;Email: <a href="mailto:raoul.monnier@thomson.net">raoul.monnier@thomson.net</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TXT Polymedia SPA</td>
<td>Sergio Gusmeroli&lt;br&gt;Phone: +39 02 2577 1310&lt;br&gt;Email: <a href="mailto:sergio.gusmeroli@txtgroup.com">sergio.gusmeroli@txtgroup.com</a></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| KT Corporation                                                   | Myoung-Wan Koo<br>Phone: +82 2 526 6347<br>Email: mskim@kt.co.kr}
# Table of Contents

1. SUMMARY.................................................................................................................................10

2. UNIFIED CONTAINER FORMATS FOR ESSENCE AND METADATA.................................13
   2.1. MATERIAL EXCHANGE FORMAT (MXF)..............................................................................13
      2.1.1. MXF introduction.............................................................................................................13
      2.1.2. MXF encoding..................................................................................................................13
      2.1.3. MXF file structure............................................................................................................14
   2.2. OTHER CONTAINER FORMATS..........................................................................................14
      2.2.1. Open Media Framework Interchange (OMFI or OMF) and Advanced Authoring Format (AAF)..........................................................................................................................15
      2.2.2. General Exchange Format (GXF).....................................................................................15
      2.2.3. 3GPP file format (3GP)....................................................................................................15
      2.2.4. DVB File Format (DVB-FF)............................................................................................15

3. DATA MODELS SURVEY .............................................................................................................16
   3.1. PROFESSIONAL DATA MODELS.........................................................................................16
      3.1.1. Broadcast Metadata Exchange Format (BMF).................................................................16
         3.1.1.1. BMF introduction.........................................................................................................16
         3.1.1.2. BMF class model.........................................................................................................16
         3.1.1.3. BMF ProductionElements .........................................................................................17
         3.1.1.4. BMF metadata exchange ............................................................................................18
         3.1.1.5. BMF Standardisation ..................................................................................................18
         3.1.1.6. BMF related projects ..................................................................................................18
         3.1.1.7. BMF in practice ...........................................................................................................19
         3.1.1.8. BMF in practice ...........................................................................................................19
      3.1.2. MPEG-7 ..........................................................................................................................19
         3.1.2.1. Introduction ..................................................................................................................19
         3.1.2.2. Part 3: Visual ...............................................................................................................20
         3.1.2.3. Audio ..........................................................................................................................21
         3.1.2.4. Multimedia Description Schemes ...............................................................................22
         3.1.2.5. Systems .......................................................................................................................22
         3.1.2.6. Other Projects using MPEG-7 ....................................................................................23
      3.1.3. MPEG-21 .......................................................................................................................23
         3.1.4. P META .........................................................................................................................24
      3.1.5. SMEF .............................................................................................................................25
         3.1.6. Dublin Core & EBU Core ..............................................................................................25
         3.1.6.1. Dublin Core ...............................................................................................................25
         3.1.6.2. EBU Core ....................................................................................................................26
      3.1.7. DMSI (SMPTEDoubtless SMPTE 380M) ........................................................................26
      3.1.8. PrestoSpace Format ........................................................................................................28
      3.1.9. NewsML-G2 ....................................................................................................................29
      3.1.10. RDF Introduction ..........................................................................................................31
         3.1.10.1. RDF components ......................................................................................................31
         3.1.10.2. Basic RDF model ......................................................................................................32
         3.1.10.3. Basic RDF syntax and Query ....................................................................................33
         3.1.10.4. Schemas and Namespaces ......................................................................................33
         3.1.10.5. Qualified Property values .........................................................................................34
         3.1.10.6. Containers ...............................................................................................................34
         3.1.10.7. RDF Schema ...........................................................................................................35
      3.1.11. EBU/ega Schema for Advertisement ..........................................................................35
         3.1.11.1. Association of Television and Radio Sales Houses (ega) ............................................35
      3.1.12. Liaison between EBU and ega .....................................................................................36
      3.1.13. Status of the current schema .........................................................................................36

3.2. CONSUMER DATA MODELS ...............................................................................................38
   3.2.1. TV-Anytime .......................................................................................................................38
      3.2.1.1. General description .......................................................................................................38
      3.2.1.2. Keys of TV Anytime system: the attractors .................................................................38
      3.2.1.3. Structure ......................................................................................................................39
      3.2.1.4. Program description ...................................................................................................40
      3.2.1.5. Program Location ......................................................................................................40
      3.2.1.6. Program Segmentation ...............................................................................................40
4. TAXONOMY MANAGEMENT ..................................................................................................................48
  4.1. EBU ESCORT 2007..............................................................................................................................48
  4.2. SMPTE UL (UNIVERSAL LABEL).....................................................................................................49
  4.3. ISO COUNTRY CODES.........................................................................................................................50
  4.4. ISO LANGUAGE CODES.......................................................................................................................51
5. ASSESSMENT (VALIDATION OF DATA MODELS).....................................................................................52
  5.1. ASSESSMENT FOR USE CASE “PERSONALISED SEMANTIC NEWS” (7a) ..................................52
  5.2. ASSESSMENT FOR USE CASE “PERSONALISED TV GUIDE WITH ADAPTIVE ADVERTISING” (7b) .54
  5.3. ASSESSMENT FOR USE CASE “INTERNET TV IN THE SOCIAL WEB” (7c) .................................55
    5.3.1. General Principles ...........................................................................................................................55
    5.3.1.1. Distinguish Brand, Series, Programme, Version and Broadcast ...................................................55
    5.3.1.2. Brand, Series, Programme, Version and Broadcast must have unique dereferenceable Web identifiers ...............................................................................................................................................................................................55
    5.3.1.3. Channels must have unique dereferenceable Web identifiers .......................................................56
    5.3.1.4. Good timekeeping ..........................................................................................................................56
    5.3.1.5. Be able to identify points and time-based segments in video ........................................................56
    5.3.1.6. Be extensible ..................................................................................................................................56
    5.3.2. Fields ...............................................................................................................................................56
6. METADATA HANDLING AND FORMAT TRANSFORMATION .......................................................................58
  6.1. BASIC PRINCIPLES OF MODEL TRANSFORMATIONS ..................................................................58
  6.2. METADATA MODEL TRANSFORMATIONS IN NO Tube.................................................................58
    6.2.1. Transformation of the selected NoTube metadata models ..............................................................59
    6.2.1.1. Transformation within Service Provider Ambient .........................................................................59
    6.2.1.2. Transformation within Home Ambient ........................................................................................59
    6.2.2. Transformation of external sources ................................................................................................59
  6.3. METADATA MODELS SELECTED BY NO Tube..................................................................................60
    6.3.1. Models for Personalized Semantic News (7a) ................................................................................60
    6.3.2. Models for Personalized TV Guide with adaptive Advertising (7b) .............................................61
    6.3.3. Models for Internet TV in the Social Web (7c) .............................................................................61
  6.4. SERVICE IMPLEMENTATION ............................................................................................................62
    6.4.1. Services Description .........................................................................................................................62
    6.4.2. Metadata Injection ............................................................................................................................63
    6.4.3. Metadata Conversion .......................................................................................................................63
  6.5. CONCLUSION ......................................................................................................................................64
7. CONCLUSION .............................................................................................................................................65
8. REFERENCES ...............................................................................................................................................66
9. ANNEX A ...................................................................................................................................................68
List of figures

Figure 1: Physical and logical view of an MXF File ([2])..........................................................13
Figure 2: A simple MXF File ([2])..........................................................................................14
Figure 3: Overview of BMF PEs and their hierarchical relationships........................................17
Figure 4: Relation between different tools and elaboration processes of MPEG-7.....................19
Figure 5: Overview of the MPEG-7 Multimedia DSs...............................................................22
Figure 6: Different streaming strategies of same XML file.......................................................22
Figure 7: The context of P_META ([16])...............................................................................24
Figure 8: The content structure around which P_META has been built ([16])...............................24
Figure 9: Descriptive metadata frameworks and their relationship to the content of an MXF file body..........................................................................................................................27
Figure 10: Production chain of PrestoSpace([22])...................................................................28
Figure 11: High-level structure of the PrestoSpace metadata format MAD ([23])......................29
Figure 12: Assembling of the PrestoSpace metadata format MAD ([23])..................................30
Figure 13: Resource Description Framework (RDF) model.....................................................32
Figure 14: Official egta logo.......................................................................................................35
Figure 15: The TV Anytime schema..........................................................................................39
Figure 16: Aggregation of program description by the unique document structure of TV Anytime...40
Figure 17: Simplified model of Programmes Ontology.............................................................43
Figure 18: The flexible structure of the EPG Data Model..........................................................44
Figure 19: Block diagram of the ESG data model, specified in TS 102 471...............................45
Figure 20: Generic transformations of broadcast sources provided by WP2..............................61
List of Abbreviations

AAF: Advanced Authoring Format
B2B: Business-to-Business
B2C: Business-to-Customer
BMF: Broadcast Metadata Format
C2B: Customer-to-Business
C2C: Customer-to-Customer
DVB: Digital Video Broadcast
EBU: European Broadcasting Union
egta: European Group of Television Advertisers
EPG: Electronic Program Guide
HDTV: High-Definition Television
KLV: Key-Length-Value
MXF: Material Exchange Format
NIC: News Item Container
NMHA: NoTube Metadata model for Home Ambient
NMSP: NoTube Metadata model for Service Provider
OMFI: Open Media Framework Interchange
RDF: Resource Description Framework
SMEF: Standard Media Exchange Framework
SMPTE: Society of Motion Picture and Television Engineers
TVA: TV-Anytime
1. Summary

Metadata is the key to the exchange of information, control of system behaviour and access to content in the digital society, both in the professional as well as in the consumer domain. It fits into all the different environments, such as archiving, delivery and communication, and the challenge is improving seamless interoperability across platforms in a competitive market.

This second version of deliverable 2.1 “Metadata Requirements Analysis” provides an introduction to the potential metadata schemes and models, which could be applied to the NoTube system with its typical use cases and its interoperability and transformation, which might be necessary between the professional and consumer domain and even between B2B (Business to business), C2C (Consumer to Consumer) and of course between B2C (Business to Consumer) and C2B (Consumer to Business) domains. In addition to the first version, this version includes the description of a metadata schema, especially defined by EBU (European Broadcasting Union) and egta (European Group of Television Advertising) for ads in the B2B domain and the assessment, i.e. validation of the appropriate metadata schema for the different use cases, selected by NoTube.

The metadata rationale is presented from an audio-visual and multimedia sector’s perspective, mainly of TV-broadcasters as the entity in the professional domain and typical consumers of TV programs in the consumer domain, not ignoring the underlying semantics information. NoTube considers both types of typical consumer of TV programs. This includes the old-fashioned typical leaned-back viewer, who wants a minimum of interaction with the TV-program and a maximum of comfort to provide the program, what he wants at the exact time, he chooses according to his personal profile. However, this includes as well the highly interactive viewer, who wants to retrieve all the information about and interact with the program he is interested in. Following the discussions and developments of the various use cases in the NoTube project, an analysis of potential metadata models and the required translation between models has been made.

Metadata is often presented as "data about data" or "data about content" or "information about content". Above all, it consists of information that needs to be shared within a domain of application, but can also be used to communicate with the external world. Several groups of interest have invested into a range of harmonisation efforts bearing in mind the potential technical and commercial advantages of a common solution. Both, standardisation of certain metadata schemes, as well as translation or mapping between the different schemes, where it is necessary, are the natural continuations of these activities for an easier exchange of information and for the creation of new services in a digital paradigm of TV, developing interoperable products in a competitive horizontal market environment.

With the introduction of digital TV, TV via Internet and Internet TV, Metadata has now reached a higher level of maturity and should be defined as "content about content" but the notion of information needs to remain, in particular for new, highly interactive and enriched multimedia services. On one hand, the NoTube system has to consider, that "Content is consumed" like for example when a user accesses the descriptive information contained in an EPG. On the other hand, the whole NoTube architecture is based on the fact that "information is processed". This is an important distinction, as metadata solutions will support electronic data processing, indexing, searching and retrieval of any electronic resource appropriately documented. Such information is an invaluable content and is to that respect considered as an asset that needs to be delivered from the content producer to the broadcaster (B2B) and then of course from the broadcaster to the consumer (B2C). In the case of consumer generated content, even between consumers (C2C) and from the consumer back to the broadcaster (C2B). In most of these conditions, the transmitted metadata need to be translated in order to match the formats used at the various instances. This has to be taken into account by the NoTube architecture in order to guarantee a
continuous flow of assets and metadata between all the entities involved in the generation, processing, distributing and consuming of interactive, rich multimedia services.

The awareness to the strategic importance of standardising and harmonising metadata has become obvious since the metadata added value has been clearly identified during the last years and the technology barriers have been removed at least partially. The use of standardised and harmonised metadata will allow:

- Optimising the content creation process for the production, storage and delivery of interoperable content and its associated metadata at lower cost for a better description of the resources and a higher quality of service (e.g. by recording technical metadata and mastering the processing chain for controlled content manipulation or by developing specialised protocols and application programming interfaces).

- Maximising the management and use of existing resources (content availability, content accessibility, improved interactivity of richer applications, management of associated data for searching and filtering, etc.) for re-purposing content and metadata, which are both assets generating revenues.

- Ensuring that usage rules and rights are respected and that content owners benefit from the revenues generated by their content and metadata.

- Supporting the development of competitive, interactive rich media services. Digital distribution of content over digital broadcasting (traditional and interactive TV) and the Internet (Semantic Web) is providing easier access to more content. In addition, the use of digital mass storage and media servers is changing the character of content consumption. New applications and services are taking benefit of richer content (e.g. virtual and augmented reality), particularly if facilitated by the use of associated metadata. EPGs will use content description and give a simple and direct access to programmes of primary user interest. Personalisation, electronic filtering and other media personalisation techniques will be highly appreciated by some of the consumers together with intelligent agents using user profiles and usage history information. Content referencing techniques allows to access content from different sources at different time of availability. Non-linear viewing and navigation will be possible using e.g. bookmarks or highlights within segmented content.

Metadata will be a key enabler for this new, interactive rich media audio-visual content and its associated information, applications and services. The usage of standardized and harmonized Metadata and their translation between the various instances must therefore strongly be considered by the NoTube architecture. It has to be considered, that interoperability is the warrant of the development of competitive horizontal markets.

This revised version of deliverable 2.1 is dedicated to provide an overview of existing file formats and metadata models, which are used in professional and non-professional domains (chapter 2.; 3.). In addition to its first version, this deliverable includes a description of the new schema for advertisement (see chapters 3.1.11 – 3.1.13), presently under definition by EBU and egta. Within chapter 4 information are given about taxonomy and thus in which way several vocabularies can be classified. Chapter 5., which has been completely rewritten in this version of the deliverable, provides the assessment of the three use cases which have been established in NoTube with respect to their requirements towards metadata. Furthermore the need of an internal metadata format for the NoTube environment is described. In this context the task of metadata transformation or better “translation” between different metadata models, which are going to be used, or being used already in the various domains of production, archiving, in-house distribution, transmission to the consumer, or in the consumer domain by itself, is outlined in chapter 6.. In order to be able to specify an
appropriate metadata format for the typical NoTube applications, the requirements from the three use cases are specified in chapter 6, as well. This chapter has been extended significantly and documented in this version of this deliverable. Finally, a conclusion of this deliverable is given.

Since Metadata standardisation, harmonisation and translation are vital for the development, storage and distribution of interactive rich media services, the NoTube project will be instrumental to contribute to the production and translation of Metadata between the various instances, as well as to contribute to existing and new Metadata standards and produce guidelines or recommendations of general interest as a result of their implementation trials and respective innovative developments of the Metadata technology.

Several aspects, which have been raised in the first version of this deliverable, which was due in M3, have been substantiated with more details in this second version of the deliverable.
2. Unified container formats for essence and metadata

2.1. Material Exchange Format (MXF)

2.1.1. MXF introduction

The Material Exchange Format (MXF) has been developed by several organizations including SMPTE, EBU and the ProMPEG-Forum addressing the user requirements for a mainstream IT-based TV-production. Thus, MXF defines a file format for the transportation of essence and metadata in the professional domain and it is optimized for the interchange of material in content creation industries. It is a container for several essence types and supports a flexible mechanism to transport metadata, based on different metadata schemes. The MXF file format is a mature SMPTE standard and can thus be implemented with interoperability with other systems in mind. (see [1])

In the broadcast environments MXF is already widespread and software modules that support the MXF technology can be often easily integrated in the technical workflows. The overhead of MXF is very low and ranges typically at about 100 kB per File.

2.1.2. MXF encoding

MXF is a container that holds the essence together with metadata synchronously, which describes the essence technically as well as in regards of the content. All data (essence and metadata) is identified with a 16 byte Key and the length (encoded in 4 bytes) of the data, followed by the value of the data. This triplet is called Key-Length-Value or KLV. The overhead of KLV is a neglectable 0.00008% with respect to video essence, ranging only slightly higher for audio essence. The overall overhead of MXF is typically in the range of 100-250 Kilobyte, even for files as large as several Gigabytes. The KLV approach ensures the extensibility of MXF, decoders can skip the length of any unknown value (identified by the key) and continue processing the encapsulated data at the start of the next KLV triplet package.

As shown in Figure 1, an MXF file can be viewed in two ways (see [2]):

- The physical view of an MXF file is the byte stream on a disk or a wire. It is the simplest view of a file showing physical properties e. g. partitions, KLV coding etc.
- The logical (metadata) view shows the description of the file contents obtained by decoding the data model. It is defined by the contents of metadata e. g. number of picture, sound and data tracks.

![Figure 1: Physical and logical view of an MXF File ([2])](image)
The MXF container supports the possibility to transport various essence (compression) formats. The mapping of specific essence formats into the MXF structure is defined in according SMPTE documents. An excerpt of supported essence formats by MXF and the appropriate SMPTE documents are shown in Table 1. (see [1], [2])

<table>
<thead>
<tr>
<th>SMPTE Document</th>
<th>Essence Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMPTE 381M</td>
<td>MPEG Streams</td>
</tr>
<tr>
<td>SMPTE 382M</td>
<td>AES3 and Broadcast Wave</td>
</tr>
<tr>
<td>SMPTE 383M</td>
<td>DV-DFI</td>
</tr>
<tr>
<td>SMPTE 422M</td>
<td>JPEG 2000</td>
</tr>
<tr>
<td>RP 2008</td>
<td>AVC Streams</td>
</tr>
<tr>
<td>SMPTE 2019-4</td>
<td>VC-3 Coding (DNxHD)</td>
</tr>
</tbody>
</table>

Table 1: Excerpt of supported essence formats by MXF

2.1.3. **MXF file structure**

The MXF specification for the exchange of multimedia information as files in non-real time transfers with essence and metadata packaged together eases packetised transport of material of different compression formats in broadcast production environments using different networks (e.g. connecting NLEs, servers and storages). It is suitable for captured, ingested, finished and nearly finished material, designed for efficient storage on various types of media (incl. archiving e.g. storage on data tape) and transportation using communication links. The basic structure of an MXF file shows the following elements (see Figure 2) (see [2]):

- A file header which provides information about the file as whole,
- A file body which comprises picture, sound and data essence stored in essence containers,
- A file footer which terminates the file.

![Figure 2: A simple MXF File ([2])](image)

An optional but recommended index table provides rapid conversion from sample-based indexes (e.g. time code) into byte-offsets within an essence container. Optional body partitions can be inserted at intervals within the file body to provide features like robustness of metadata information by repeating the header metadata and more.

2.2. **Other Container Formats**

In this section, further container formats are being introduced. Their suitability for the NoTube project has to be researched and tested against the requirements that will evolve during the project.
2.2.1. Open Media Framework Interchange (OMFI or OMF) and Advanced Authoring Format (AAF)

The container formats Open Media Framework Interchange and Advanced Authoring Format both mainly were developed for interchange in a post-production environment. Both formats offer the ability to carry audio and video data as well as technical metadata like edit decision lists (EDL), colour corrections, transitions, and (depending on the format) other data relevant for editing and other post-production tasks. While neither OMFI nor AAF have been standardised at the Society of Motion Picture and Television Engineers (SMPTE), the AAF specification is available for the public while OMFI is the property of APPLE. Both formats do not have a defined way to carry other metadata then the technical metadata required in the post-production process and currently seem not fit to be used in NoTube. (see [3], [4], [5])

2.2.2. General Exchange Format (GXF)

The General Exchange Format has been primarily developed by Grass Valley Group. While at first only JPEG sequences were supported as video coding, more and more codecs were introduced to be carried inside a GXF file. GXF has been developed to enable quick access to material while still recording ("play-while-record") and for archiving. GXF can carry metadata and has been published in the standard SMPTE 360M. However, GXF is not quite as flexible as MXF as it does not use KLV encoding (see chapter 2.1.2.), which ensures extensibility for MXF. (see [3], [6])

2.2.3. 3GPP file format (3GP)

The 3rd Generation Partnership Project (3GPP) has developed the 3GPP file format, which itself is abbreviated 3GP. 3GP is an instance of the ISO base media file format and mainly to be used within Multimedia Messaging Services (MMS) and Packet-switched Streaming Service (PSS). The former is the well-known successor to Short Message Service (SMS) on mobile phones which can transfer higher data volumes than SMS. 3GP is intended “to be used for continuous media along the entire delivery chain envisaged by the MMS” (see [7]). 3GP files can hold metadata in a “user-data box”, which has only 10 entries. This format will most likely be one of the “external sources” (see Deliverable 6) which will hold user generated content, sent from the users mobile phone and sent via MMS or other file transfer mechanisms. The format has six “profiles” which differ in complexity and are designed for different uses (e.g. better streaming support or simplicity). (see [7])

2.2.4. DVB File Format (DVB-FF)

The DVB project is a consortium of broadcasters as well as manufacturers of different parts of hard- and software in the digital distribution of video material. The file format has been developed to be used for the storage and exchange of protected content and can carry audio, video and data. Like 3GP and MP4, the DVB-FF is derived from the ISO base media file format. The metadata to be used in DVB-FF is based on TV-Anytime structures (see chapter 3.2.1.). The essence can be encapsulated in MPEG-2 Transport Stream or in RTP packets. This format is still under development, where interoperability and developing guidelines are supposed to follow the initial release. (see [8])
3. Data Models survey

3.1. Professional Data Models

3.1.1. Broadcast Metadata Exchange Format (BMF)

3.1.1.1. BMF introduction

The Broadcast Metadata exchange Format BMF was developed by the Institut für Rundfunktechnik (IRT) in collaboration with broadcasters and relevant industrial partners. The purpose of BMF is to ensure the uniformly regulated exchange of metadata in a professional broadcast environment, especially for professional television production. It was possible to achieve this aim by analysing the relevant processes in a television production in depth, beginning with the idea for a programme, through its planning to creative arrangement and transmission. (see [9], [10])

Therefore BMF covers almost all areas that a television programme can ever deal with or touch upon: editorial work, programme planning, rights administration, production organisation, production of programmes and items, acquisition and recording of clips, handling programmes and items, compiling programme schedules, dealing with run-downs through to archiving and reuse. BMF also fully supports the description of the material exchange between broadcasting organisations and production companies.

A number of companies were and remain included in the development of BMF, such as S4M, Blue Order, Dalet, T-Systems, Dimetis, DAVID, AVID, IBM, SGI, Pinnacle, Thomson Grass Valley, Sony, Panasonic, VCS, Flying Eye, MCI, to name just a few. This wide industrial support ensures that system integrators can find their way with their applications within the data model. (see [10])

3.1.1.2. BMF class model

The BMF model is defined as a class model in UML (Unified Modelling Language) and it is available from the IRT web page (http://www.irt.de/bmf) (see [9]). This class model serves as a semantic exchange scheme between the various components of a file-based production platform. In the current-affairs area, this affects, among other things, the exchange between editorial management system and production platform, editorial management system and play-out control, editorial management system and archive, and the video-file transfer among the broadcasting stations, the regional studios and the production houses. BMF does not, however, describe the way information is stored in the individual systems. (see [9])

The BMF class model specifies data structures, which represent the conceptual BMF model. To conform to BMF, products must implement this class model. Complete implementation of the model is not necessary; it is sufficient to implement the class model packages that are required for the respective application, in order to be able to describe the product as conforming to BMF.
The introduction of file-based platforms in television production requires the management and the exchange of Metadata between the various systems and such a platform. Nowadays this exchange of Metadata between systems, and thus the integration of these systems in the operating process, is carried out exclusively on an individual and therefore proprietary basis. This means high costs for the planning and operational integration. With the BMF class model a basis has been created to reduce these costs. BMF provides the prerequisites for uniform description of the information. This format can be reused in the operating process at any interface between individual systems. Then a renewed definition for further interfaces is no longer necessary. The BMF class model describes the information to be exchanged, its meaning, and the relationships that are relevant and are exchanged in the course of a television production process. (see [9])

3.1.1.3. BMF ProductionElements

The core of the BMF class model is described by the business objects Shot (raw material), Scene, Item and Programme. These business objects are called ProductionElement (PE) in context of the BMF class model. (see [9])

The PE Programme represents a business object that is necessary to describe a television programme and the associated material. These programmes are transmitted within a framework of a channel. In Programmes, Items, Scenes and Shots (raw material) are used. Objects of class Programme therefore have references that point to objects of the type Item, Scene or Shot.

The PE Item is necessary as part of the preparation of a programme to describe the item and its material, from planning to transmission. Items are not transmitted as independent programmes, but always as part of a programme. Both Scenes and Shots are used in Items. An object of the class Item therefore has references to Shot or Scene typed objects.

A television production is often scene-based within feature production. A programme is planned and produced in scenes. A scene is described as PE Scene in context of the BMF class model. The Scenes are made up of Shots (raw material) and thus represent an intermediate product within such a production. An object in the BMF class Scene has references which refer objects of the type Shot.

The BMF class Shot represents a PE which is necessary as part of the preparation of an item or a programme, i.e. for its planning, description and production. It thus represents the raw material that is used in scenes, items and programmes. Raw material can be newly generated material from acquisition, but it can also be from the archive (excerpt) or from other broadcasting stations. A shot can also be a pre-processed material (field editing) from acquisition.
The different BMF PEs, described above, and their relationships together are shown in Figure 3. It is evident that these relationships build a hierarchy, initiating with the Programme, between the different PEs.

![Diagram of BMF PEs and their hierarchical relationships]

The structure of BMF PEs, shown above, completely incorporates the mechanisms of the data model FESADneu and the possibilities for describing the contents. Without the complete incorporation of FESADneu into the data model of BMF, there would be a break in the “life cycle” of content, and consistent reuse would therefore not be guaranteed. (see [10])

### 3.1.1.4. BMF metadata exchange

BMF supports the two common methods of metadata exchange: metadata being exchanged together with the essence within a given container format such as MXF, or an exchange separate from the essence using predominantly XML-files.

For the transportation within the MXF file format the BMF metadata are encoded as KLV (key length value) data (see chapter 2.1.2.). The base for an xml based metadata exchange is the BMF XML Scheme based on the class model developed by the IRT.

### 3.1.1.5. BMF Standardisation

The BMF class model is currently in the process of being registered in the registries of the Society of Motion Picture and Television Engineers (SMPTE). This will ease the exchange of content by adding KLV-coded metadata to the essence using MXF file transfer.

### 3.1.1.6. BMF related projects

#### 3.1.1.6.a. WMP

The project “Technologies for Media Production” (German: “Werkzeuge der MedienProduktion, WMP) developed a distributed system for film and broadcast work which offers the exchange of data between film and television production. The project started in 2006 and end in 2008. The developed system is able to replace the tape-based delivery of Content (video, audio and metadata) with a tapeless approach. The distributed production system has been developed based on Web-Service technology, which follows the SOA (“Service Oriented Architecture”) principles. The project allows productions to prepare content in an integrated production-, post-production-, and archiving environment for a variety of output formats, such as Digital Cinema, television, HDTV and mobile display devices. For the xml based metadata exchange between the cinema and the broadcast...
domain a BMF interoperability interface was developed, that enables an broadcast editor to search in an cinema production archive. The technologies for the project were developed by the Fraunhofer Institute for Integrated Circuits IIS, the Institut für Rundfunktechnik (IRT), CinePostproduction Bavaria Bild & Ton, AVT Audio Video Technologies GmbH, IRIDAS and associate partner Bayerischer Rundfunk.

3.1.1.6.b. CONTENTUS

This project is a compilation of technologies that facilitate the preservation of our cultural heritage. The core element of this project is an efficient processing pipeline “from information asset to knowledge network”, that describes the workflow for processing our cultural heritage from traditional storage media to community-enriched content accessible through semantic multimedia searches. IRT develops software modules for relevant data model conversions, such as a BMF converter, and contributes to the Requirements Engineering, System Design and test methods among other things.

3.1.1.6.c. VITALAS

This project is an innovative project designed to provide advanced solutions for indexing, searching and accessing large scale digital audio-visual content as it can be found in broadcasters’ archives. The focus of IRT’s contributions is to bridge the divide between content creators and broadcasters on one side and the other research partners in the project on the other side. This comprises the establishment of functional and technical requirements with respect to the steps of the media production chain, but also the development of services for the exchange of media essence and metadata in file-based production environments. One format for the metadata exchange in this project is the BMF metadata format.

3.1.1.7. BMF in practice

The adoption of the BMF class model in practice, especially in broadcast environments, is currently at the beginning. One of the first broadcast stations that implemented a metadata exchange based on BMF is the public-broadcasting organisation NDR (Norddeutsche Rundfunk) in Germany. Within the NDR news production BMF is used to exchange the description of a news programme after transmission to an archive system. Furthermore these BMF metadata are transmitted to a central repository to enable the access for other public-broadcasting organisation for editorial work.

At the beginning of this year, BMF has been chosen as a candidate which can be used to establish a uniform metadata exchange of TV programmes between the nine German Public Service Broadcasters within the ARD (Arbeitsgemeinschaft der öffentlich-rechtlichen Rundfunkanstalten der Bundesrepublik Deutschland, “Association of Public Broadcasting Corporations in the Federal Republic of Germany”). For this purpose, BMF will be updated and extended to version 2.0. This will happen according to a number of new requirements that arose from the fact that the use of BMF for the programme exchange also includes the exchange of programme planning data, radio broadcasting and the delivery of content to video on demand (VoD) platforms.

3.1.2. MPEG-7

3.1.2.1. Introduction

MPEG-7 is an ISO/IEC standard developed by ISO’s Moving Pictures Experts Group (MPEG), formally called Multimedia Content Description Interface. While former MPEG standards focus on the compression and transportation of audiovisual content, the objective of MPEG-7 is not to replace these standards but to provide additional functionality for efficient search, browsing and retrieval of multimedia content. As such, MPEG-7 focuses on the description part of the process chain including feature extraction, description and use of
the description (i.e. search, browse and retrieve). It can be used independently of other MPEG standards (e. g. attached to an analogue video file).

XML-based MPEG-7 defines Descriptors (D) for audio, video and graphics using the Description Definition Language (DDL) which is supported by several Description Schemes (DS) without specifying the way the metadata is extracted or used by an application. The MPEG-7 descriptions may either be located with the associated essence (audiovisual material) or elsewhere in the network, the latter needing a mechanism to link essence and metadata. A binary representation (BiM) for efficient transportation of the descriptions and a reference software implementation are also defined in the standard.

The Descriptors (D) and the Description Schemes (DS) are defined by the Description Definition Language (DDL) so that the Ds are structured into DSs (see Figure 4). Both are instantiated as descriptions in XML format and can then be encoded in binary format for efficient delivery and storage. For specific applications, the DDL allows the extension of particular DSs.

Description Tools (Descriptors and Description Schemes) cover basic visual features (colour, texture, shape, motion, localization, and face recognition) and for audio utilize low-level Descriptors (e. g., spectral, parametric, and temporal features of a signal) as well as high-level Description Tools that are more specific to a set of applications (including general sound recognition, instrumental timbre, spoken content, audio signature). Multimedia Description Schemes (MDS) comprises the set of Description Tools dealing with generic entities (used in audio and visual descriptions) as well as multimedia entities (used whenever more than one medium needs to be described).

1 Source of Figure 4: [http://en.wikipedia.org/wiki/Image:Mpeg7image2.svg](http://en.wikipedia.org/wiki/Image:Mpeg7image2.svg)
The DDL allows not only the creation of new DSs and Ds but also the modification (e.g. extension) of existing DS. Thus, as DDL schema defines the constraints for a valid MPEG-7 description. Encoded on XML, it defines the document structure expressed in a Document Type Definition (DTD). The MPEG-7 DTD uses W3C’s XML Schema Language extended by MPEG-7-specific mechanisms like array and matrix data types. The validity of the description with these rules is checked with a DDL parser which is first initialised with a DDL file.

The MPEG-7 specification is composed of 12 parts (2008), the following subsections introduce the most relevant parts of MPEG-7 for this metadata model survey. The information is collected from [11] and [12].

3.1.2.2. Part 3: Visual

The MPEG-7 Visual Description Tools cover basic visual features of media like colour, texture, shape, motion, localization, and face recognition. As an example the Motion Trajectory descriptor part of the Motion Descriptors in combination with the Spatio-Temporal Locator part of the Localization Tools can be used to identify the position of objects in video streams.

Other application examples for visual descriptions include digital libraries (image and video catalogue), broadcast media selection (TV channels), and multimedia editing (personalised electronic news service, media authoring). Among this diversity of possible applications, the MPEG-7 Visual feature descriptors allow users or agents to perform the following tasks taken as examples.

- **Graphics**: Draw a few lines on a screen and get, in return, a set of images containing similar graphics or logos.
- **Images**: Define objects, including colour patches or textures, and get, in return, examples among which you select the ones of interest.
- **Video**: On a given set of video objects, describe object movements, camera motion, or relations between objects and get, in return, a list of videos with similar or dissimilar temporal and spatial relations.
- **Video Activity**: On a given video content, describe actions and get a list of videos where similar actions happen.

3.1.2.3. Audio

MPEG-7 Audio provides structures for describing audio content. This includes a set of low-level Descriptors, for audio features that cut across many applications (e.g. spectral, parametric, and temporal features of a signal), and high-level Description Tools that are more specific to a set of applications. Those high-level tools include general sound recognition and indexing Description Tools, instrumental timbre Description Tools, spoken content Description Tools, an audio signature Description Scheme, and melodic Description Tools.

Example applications for the MPEG-7 Audio Description Tools

- **Query by humming**: Audio content is indexed by the melodic Description Tools of MPEG-7 and stored in a database. A user of that service connected via a hand-held hums the melody and the server extracts the melody from the humming and searches for matches in the database.
- **Query for spoken content**: A telephone message service converts incoming voice mails into text or a phonemic representation using MPEG-7 Spoken Content Description Tools where a translation to text fails. The user calls up the message server and asks for messages by a spoken query. The query is transformed into
a MPEG-7 description and this is used to look for matches in the message queue.

• Assisted Consumer Level Audio Editing: An audio signal has been processed to extract metadata to facilitate interactive applications. For example, many low-level audio features such as power, loudness, spectral centroid, and spectral envelope could be used in a hypothetical consumer-level audio editor. One could use these Descriptors for display, such as loudness, in the place of the waveform. With some knowledge of standard editing rules within a genre, a sophisticated program could use spectral features to equalize channels and thereby automatically enhance the clarity of a mix.
3.1.2.4. **Multimedia Description Schemes**

MPEG-7 Multimedia Description Schemes (MDS) are metadata structures for describing and annotating audio-visual content. Figure 5 provides an overview of the various tools of MDS.

![Figure 5: Overview of the MPEG-7 Multimedia DSs](image)

- **Basic Elements**
- **Content Management**: These tools describe the following information: creation and production, media coding, storage and file formats, and content usage.
- **Content Description**: The structural tools describe the structure of the audiovisual content in terms of video segments, frames, still and moving regions and audio segments. The semantic tools describe the objects, events, and notions from the real world that are captured by the audiovisual content.
- **Navigation and Access**: Provides tools for facilitating browsing and retrieval of audio-visual content by defining summaries, partitions and decompositions, and variations of the audio-visual material.
- **Content Organization**: These tools are used for organizing and modelling collections of audio-visual content and of descriptions.
- **User Interaction**: These tools describe user preferences and usage history pertaining to the consumption of the multimedia material.

3.1.2.5. **Systems**

MPEG-7 description schemes are represented as XML documents. Due to the nature of XML, these documents are hierarchical organized and not ideally designed for a real-time, constrained and streamed environment. To overcome the lack of efficiency of textual XML, MPEG-7 Systems defines a generic framework to facilitate the carriage and processing of MPEG-7 descriptions: BiM (Binary Format for MPEG-7). It enables the streaming and the compression of any XML documents. Figure 6 shows three different strategies using BiM to stream a XML file in one or more segments.

![Figure 6: Different streaming strategies of same XML file](image)
A textual format TeM has been added to the systems part for uses cases where compression is not an important requirement. While BiM uses the knowledge of the specific XML schema description (XSD) to allow for high compression rates, TeM can be used for any XML formatted document.

The MPEG-2 systems specification contains an amendment [13] for the transmission of metadata over MPEG-2 transport streams. This amendment allows for different delivery methods, including a synchronous and a cyclic transmission of metadata or fragments of metadata.

3.1.2.6. Other Projects using MPEG-7

3.1.2.6.a. GMF4iTV and porTiVity

Both projects focussed on object interactivity in broadcast services using MPEG-7. Objects in the program stream have been highlighted by means of motion trajectory descriptions and identified by the Semantics tools. An application made objects selectable and granted access to additional content. In GMF MPEG-7 metadata was transmitted to the terminal using the methods described in the previous section. The terminal used the metadata to render the application. porTiVity targeted mobile hand-held devices; therefore the MPEG-7 metadata was used to generate a MPEG-4 LASER[14] scene description containing the broadcast video and the object interaction.

3.1.2.6.b. PrestoSpace

The PrestoSpace project uses MPEG-7 Descriptors and Description Schemes in its metadata model. Please refer to chapter 3.1.8. for further details.

3.1.3. MPEG-21

The MPEG-21 standard (ISO/IEC 21000) is a metadata framework which aims to enable the transparent access to multimedia resources across a wide range of networks and devices. The intention of MPEG-21 is the adaptation of Digital Items defined as structural digital objects with standard representations, identifications and descriptions. It includes a variety of dimensions, including users, terminals, networks and delivery for the description. To describe the User preferences MPEG-21 addresses the areas of display, accessibility and mobility characteristics. Terminal capabilities include hardware, software and display properties. Moreover the device profiles indicate the supported media formats (e.g. MPEG-2). The delay, error and bandwidth characteristics display accessibility are described by the mobility characteristics. Finally, the delivery descriptors specify the types of supported transport protocols (e.g. TCP/IP) and the types of connections (e.g. multicast).

The MPEG-21 links previous coding and metadata standards (e.g. MPEG-1, -2, -4 and -7) together to build a protectable universal package of Digital Items for the consumption by users. The MPEG-21 standard is implemented as a XML Schema, that contains the entire grammar to describe Digital Items. For that reason the representation language is xml. (see [15])
3.1.4. **P_META**

The P_META format “is a library of common terms, data types and data structures representative of what are being used in production, to editorially identify, or technically describe content and associated rights.” It comprises a set of tools developed by the EBU project P/Meta and represents the semantic layer, i.e. the descriptive metadata layer which defines the meaning of each element of description relevant to production practices. P_META 2.0 defines and aggregates a basic catalogue of data items that can be used as the building blocks for common information exchange between content creators, distributors, archives and systems involved in production processes (see Figure 7). Thus, P_META focuses on business-to-business (B2B) processes, but it features also interoperability with business-to-consumer (B2C) metadata.

Broadcasters using internal metadata solutions can map their models to the P_META metadata schema as a basis for interface specifications. For this purpose, P_META was designed independent from a specific technology, so that it can, for example, be implemented as XML documents or embedded in file formats. (see [16])

There is explicitly no data model behind P_META, only a list of exchange sets. These exchange sets bundle information that is put together for a particular purpose (application). To this extent, P_META concerts an exchange process that is strictly application-orientated and must therefore be newly implemented upon every change and new application.
The P_META model is currently used by VRT (Flemish Belgian Broadcasters) to transmit metadata between different production islands within their environment. Furthermore the TVP (Poland Broadcast station) is starting the implementation of P_META to describe sport events. Parts of the P_META model has been integrated in the PrestoSpace metadata model as well. PrestoSpace is an European project on archive restoration (see chapter 3.1.8.).

### 3.1.5. SMEF

The Standard Media Exchange Framework (SMEF) Data Model provides a set of data definitions for the range of information involved in the production, development, use and management of media assets. Its purpose is to help ensure that systems will be able to operate together and exchange information by establishing a common understanding of shared data. It was developed by Siemens for the BBC.

Some aspects of the model are very specific to the BBC, and some are much more generally applicable. It includes terms for very specific descriptive pieces of information about the times of individual shots or audio clips, and also metadata associated with the broader description of programmes in the context of an EPG. It is designed to support businesses processes such as commissioning and archival as well as broadcast. It uses the ISO identifiers for unique and permanent identification of audio visual items: ISAN (International Standard Audiovisual Number), V-ISAN (Version-ISAN - ISAN with a version suffix) and ISRC (International Standard Recording Code) and also SMPTE UMID (Unique Material Identifier). (see [17])

The model uses entities, attributes, relationships, cardinality constraints, optionality indicators, primary keys, Java-style classes with inheritance of attributes, and exclusive groups of entities (groups where for X there can be either A, B or C).

Due to its broad scope SMEF is a model of huge complexity. The model is described as a series of eight large entity-relationship diagrams.

### 3.1.6. Dublin Core & EBU Core

#### 3.1.6.1. Dublin Core

The development of the Dublin Core metadata model has been done by the Dublin Core Metadata Initiative (DCMI), which is an open, non-profit Company incorporated in Singapore. The DCMI includes the work on architecture and modelling, discussions and collaborative work in DCMI communities, annual meetings and workshops and standards liaison.

The DCMI defines 15 core elements in the Dublin Core Metadata Element Set (DCEMS). The DCEMS deals with cross-domain information resource description. For Dublin Core applications, a resource typically is an electronic document that is described online using the 15 elements described in the standard ISO 1583. Implementations are typically based on the Resource Description Format (RDF) using XML. The DCEMS is part of a larger set of metadata vocabularies defined in DCMI Metadata Terms (DCMI-TERMS). These DCMI-TERMS define additional metadata elements to enable a more detailed description of the 15 core elements.
Dublin Core Metadata Element Set

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Title</td>
</tr>
<tr>
<td>2</td>
<td>Creator</td>
</tr>
<tr>
<td>3</td>
<td>Subject</td>
</tr>
<tr>
<td>4</td>
<td>Description</td>
</tr>
<tr>
<td>5</td>
<td>Publisher</td>
</tr>
<tr>
<td>6</td>
<td>Contributor</td>
</tr>
<tr>
<td>7</td>
<td>Date</td>
</tr>
<tr>
<td>8</td>
<td>Type</td>
</tr>
<tr>
<td>9</td>
<td>Format</td>
</tr>
<tr>
<td>10</td>
<td>Identifier</td>
</tr>
<tr>
<td>11</td>
<td>Source</td>
</tr>
<tr>
<td>12</td>
<td>Language</td>
</tr>
<tr>
<td>13</td>
<td>Relation</td>
</tr>
<tr>
<td>14</td>
<td>Coverage</td>
</tr>
<tr>
<td>15</td>
<td>Rights</td>
</tr>
</tbody>
</table>

Table 2: Dublin Core Metadata Element Set

Dublin Core neither forms a full set of metadata capable of fulfilling all purposes for broadcasting, neither does it suffice for broadcast archiving, where Dublin Core is most suitable: Several information relevant in broadcasting (e.g. timecode) are not covered by Dublin Core. (see [18])

Dublin Core has been implemented and used within many projects since the Dublin Core Metadata Initiative has started. A detailed overview of these projects is given on the official web page (http://dublincore.org/projects).

3.1.6.2. **EBU Core**

The EBU Core Metadata Set is based on the EBU Digital Strategy Group results. It defines a minimum list of metadata (attributes) to describe radio and television content – it is particularly intended to describe archived content. The EBU Core specification addresses the programme exchange between broadcast organisations or between production facilities in a distributed environment. (see [19])

The metadata defined in EBU Core are an extension to the Dublin Core Metadata Element Set. It is a minimum list of attributes that describe a media resource and these attributes should be used in Dublin Core centric environments. (see [19])

The EBU Core Metadata Set has a limited inventory of metadata as described above. Thus it doesn't support technical descriptions, timeline related descriptions and it haven't information regarding to the whole broadcast production process of the content. It is primarily intend to describe the content from the archiving point of view.

The main proposed implementation of the EBU Core Metadata Set is to develop an EBU Archive Portal. The search of material inside this archive should be supported across all repositories of EBU participants. Furthermore the EBU Core Metadata Set should allow EBU Members to contribute to the European Digital Project in collaboration with National Libraries and Museums. (see [19])

3.1.7. **DMS1 (SMPTE-Standard SMPTE 380M)**

The descriptive metadata scheme-1 (DMS-1) was developed to use it in the context of the MXF file format to transport descriptive information about with the audio-visual content within a MXF file (see also chapter 2.1.).

There are several frameworks for descriptive metadata (DM) in that are summarised in the MXF descriptive metadata scheme-1 (DMS-1). This collection of DM frameworks and their
associated metadata sets are defined in the SMPTE standard 380M and can be applied to any operational pattern specification of MXF (see [20]). The defined DM frameworks are all based on the same data model as a set structure. They may be attached to the header metadata according to a mechanism defined in the MXF file format specification. The metadata elements (attributes), used in the defined DMS-1 metadata sets, are registered in SMPTE RP210 (see [21]).

The following DM frameworks which group related metadata properties are defined (see [20]):

- Production framework: provision of identification and ownership details of the audio-visual content of the file body
- Clip framework: provision of capture and creation information regarding individual audio-visual clips in the file body
- Scene framework: description of actions and events within individual scenes of the audio-visual content of the file body.

The DM frameworks allow giving contextual meaning to metadata. Figure 9 illustrates the relation of the structure of the DM frameworks contained in the header metadata to the content of an MXF file body. In conjunction with MXF, DMS-1 is comprehensive for describing material to be exchanged, but it is not sufficient to support the entire production processes. Also, DMS-1 can exist only in conjunction with MXF. Without MXF, all relationships to the associated essences and structures are lost. (see [21])
3.1.8. **PrestoSpace Format**

The PrestoSpace metadata format was developed during the PrestoSpace project activities. This project has the main objective to realise an integrated system for audio-visual preservation and access. The main issue to preserve the audio-visual material is the migration from the analogue to digital material. Beside the digitisation process the restoration, the storage and the production of content information (metadata), to enable the access and the delivery, are included in the PrestoSpace production chain. Figure 10 illustrates the PrestoSpace production chain. (see [22])

![Figure 10: Production chain of PrestoSpace([22])]()
The assembling of the MAD metadata format is illustrated in a more detailed manner in Figure 12 below, where the different parts are associated to their functional tasks. The global root element is directly connected to the editorial object business entity and belongs to the defined Ad hoc structures of the MAD format. The Ad hoc structures are also used to describe the material realisation context and to hold the Ancillary Data. The P_META sets are used for the top level identification information, including production, publication and genre information. For the description of the editorial parts and views of an editorial object the MPEG7 profile nodes are used. Moreover the MPEG7 profile nodes express all content related information. (see [23])

The PrestoSpace metadata format is free of royalty and free of license available for users to enable archive organisations and providers of PrestoSpace services to share a common defined framework. Further information can be found on the web page (see [23]).

3.1.9. NewsML-G2

The NewsML-G2 developed by the IPTC (International Press Telecommunication Council) defines a generic model for exchanging all kinds of newsworthy information. As a member of IPTC family of G2-Standards, NewsML-G2 shares many of the used components with other standards of this family, such as EventML-G2, SportsML-G2. These shared components form the base for the IPTC News Architecture for G2. (see [24])

Inside of NewsML-G2 different layers are specified to provide the description of the content, information about the handling of the news, information about the packaging of news, and also information about the technical transfer. The format is implemented as a XML Schema and the latest version 2.2 is available from the web page including the specification documents. (see [24])
The news exchange format NewsML-G2 provides the exchange of general news which can be expressed by different types of media (textual news, articles, photos, graphics, audio and video). The description of news includes the technical representation of the content and information (e.g. abstract, category) about the content in a semantic manner. The combination of the content representation and metadata about the content builds the construct called News Item in the scope of NewsML-G2. (see [24])

To categorize and annotate the content in a factual way the NewsML-G2 format provides different concepts (Concept Item). The concepts representing any kind of object, such as persons, organisations, companies, themes, emotions etc., and can be seen as a reference to the news. Furthermore the exchange of controlled vocabularies is supported. (see [24], [25])

For the exchange of a number of News Items the conception of NewsML-G2 provides a flexible mechanism for packaging the news in a structural way. This is facilitated by the construct called Package Item that references to existing News Items or other kind of Items (also web resources) and structure them like a tree. (see [24], [25])

The administrative information (e.g. creation date, revision, embargo dates, provider) related to the News Items or other kind of Items are provided by the management layer of the NewsML-G2 format. The main feature of the management layer is supporting revisions of an Item. This is done by preserving the identification number over the revisions of an Item and incrementing a version number. (see [24], [25])

Finally, the specification for the NewsML G2 Architecture defines a construct called News Message. This construct is specified as a wrapper resp. an envelope for transmitting on or more Items via any electronic channel. (see [24], [25])
3.1.10. RDF Introduction

The W3C standard RDF (Resource Description Framework) is a framework to describe resources on the web and one of the key technologies for the Semantic Web. With RDF it is possible to interrelate distributed information describing the same resource, e.g. video, audio or document. The information about a web resource is expressed as a list of statements in form of subject-predicate-object triples. A RDF Statement consists of a resource (subject) expressed by a unique URI, a property (predicate) of the resource and the value of the property (object). The value of a property can be another resource. The most popular collection of properties is the Dublin Core Element Set. (see [26])

The significant benefit that RDF brings is that it will allow the resource description communities to primarily focus on the issues of semantics rather than the syntax and structure of metadata. The contents of the framework will be determined by the stakeholder communities - independently developed and maintained. RDF also allows for the re-use, extendibility and refinement of established resource description standards since these will be available in machine-readable form.

RDF provides interoperability between applications that exchange machine-understandable information on the Web. RDF emphasizes facilities to enable automated processing of Web resources. RDF can be used in a variety of application areas; for example: in resource discovery to provide better search engine capabilities, in cataloguing for describing the content and content relationships available at a particular Web site, page, or digital library, by intelligent software agents to facilitate knowledge sharing and exchange, in content rating, in describing collections of pages that represent a single logical "document", for describing intellectual property rights of Web pages, and for expressing the privacy preferences of a user as well as the privacy policies of a Web site. RDF with digital signatures will be key to building the "Web of Trust" for electronic commerce, collaboration, and other applications.

3.1.10.1. RDF components

The RDF specifications provide a lightweight ontology system to support the exchange of knowledge on the Web. RDF provides interoperability between applications that exchange metadata and is targeted for many application areas including; resource description, site-maps, content rating, electronic commerce, collaborative services, and privacy preferences. RDF is the result of members of these communities reaching consensus on their syntactical needs and deployment efforts.

RDF also allows statements to be made about other statements. This is useful when there is a need to express the properties of metadata. The first set of properties describes a report. Note that there is a bagID on the <Description> tag to give that collection of properties a unique identifier. The second description, which points to the first bagID, not another resource, describes the properties of the first report description. In this case, it indicates the time period that the metadata about the report should be valid for. The Valid From/To dates applies to the whole Description block.

The RDF also includes preliminary work on the definition of metadata schemas. A schema will define the meaning, characteristics, and relationships of a set of properties. This may include constraints on potential values and inheritance of properties from other schemas. The RDF schema is based on the same model as the RDF syntax specification. In the following example, for instance, an RDF schema defines two properties (Title and Creator). Each is defined as type "property" with a human-readable label and a short description (see Figure 13).
3.1.10.2. Basic RDF model

The foundation of RDF is a model for representing named properties and property values. The RDF model (see [26]) draws on well-established principles from various data representation communities. RDF properties may be thought of as attributes of resources and in this sense correspond to traditional attribute-value pairs. RDF properties also represent relationships between resources and an RDF model can therefore resemble an entity-relationship diagram. (More precisely, RDF Schemas — which are themselves instances of RDF data models — are ER diagrams.) In object-oriented design terminology, resources correspond to objects and properties correspond to instance variables.

RDF is an application of XML (W3C, 1998b). It utilises the Namespace facility of XML (W3C, 1999b). The XML Namespace, which points to a URI, allows RDF to scope and uniquely identify a set of properties. This set of properties, called a schema, can be accessed at the URI identified by the namespace. Also, since RDF is based on XML, it inherits the language tag, thus enabling the support of multi-lingual metadata.

The RDF data model is a syntax-neutral way of representing RDF expressions. The data model representation is used to evaluate equivalence in meaning. Two RDF expressions are equivalent if and only if their data model representations are the same. This definition of equivalence permits some syntactic variation in expression without altering the meaning.

The basic data model consists of three object types:

Resources:

- All things being described by RDF expressions are called resources. A resource may be an entire Web page; such as the HTML document "http://www.w3.org/Overview.html" for example. A resource may be a part of a Web page; e.g. a specific HTML or XML element within the document source. A resource may also be a whole collection of pages; e.g. an entire Web site. A resource may also be an object that is not directly accessible via the Web; e.g. a printed book. Resources are always named by URIs plus optional anchor ids. Anything can have a URI; the extensibility of URIs allows the introduction of identifiers for any entity imaginable.

Properties:

- A property is a specific aspect, characteristic, attribute, or relation used to describe a resource. Each property has a specific meaning, defines its permitted values, the
types of resources it can describe, and its relationship with other properties. This
document does not address how the characteristics of properties are expressed; for
such information, refer to the RDF Schema specification.

Statements:

- A specific resource together with a named property plus the value of that property for
  that resource is an RDF statement. These three individual parts of a statement are
called, respectively, the subject, the predicate, and the object. The object of a
statement (i.e., the property value) can be another resource or it can be a literal; i.e.,
a resource (specified by a URI) or a simple string or other primitive datatype defined
by XML. In RDF terms, a literal may have content that is XML mark-up but is not
further evaluated by the RDF processor. There are some syntactic restrictions on
how mark-up in literals may be expressed.

3.1.10.3. **Basic RDF syntax and Query**

The RDF data model (see [26]) provides an abstract, conceptual framework for defining and
using metadata. A concrete syntax is also needed for the purposes of creating and
exchanging this metadata. This specification of RDF uses the Extensible Markup Language
[XML] encoding as its interchange syntax. RDF also requires the XML namespace facility to
precisely associate each property with the schema that defines the property.

The specification defines two XML syntaxes for encoding an RDF data model instance. The
serialization syntax expresses the full capabilities of the data model in a very regular fashion. The
abbreviated syntax includes additional constructs that provide a more compact form to
represent a subset of the data model. RDF interpreters are expected to implement both the
full serialization syntax and the abbreviated syntax. Consequently, metadata authors are
free to mix the two.

RDF/XML is the specified XML syntax used for the exchange of RDF between different
applications or operating systems. The other syntax for writing RDF is Notation 3 (N3) (see
[27]). This RDF Notation is mostly used for readability during RDF application development
or RDF Storage. RDF data stores are so called triple stores. Every RDF statement (triple) is
stored in the database. The resulting RDF repository can be queried by the standard query
language SPARQL (see [28]). Today's triple store vendors also support greater machine
interpretability languages like Web Ontology Language (OWL) or RDF Schema (RDFS) (see
[27], [28], [29]).

OWL and RDFS are based on RDF. Both extend the RDF model with richer vocabulary for
descibing properties and the hierarchical description of classes, attributes and associations
of these classes and their relationship to each other.

3.1.10.4. **Schemas and Namespaces**

When we write a sentence in natural language we use words that are meant to convey a
certain meaning. That meaning is crucial to understanding the statements and, in the case of
applications of RDF, is crucial to establishing that the correct processing occurs as intended.
It is crucial that both the writer and the reader of a statement understand the same meaning
for the terms used, such as Creator, approvedBy, Copyright, etc. or confusion will result. In a
medium of global scale such as the World Wide Web it is not sufficient to rely on shared
cultural understanding of concepts such as “creatorship”; it pays to be as precise as
possible.

Meaning in RDF is expressed through reference to a schema. You can think of a schema as
a kind of dictionary. A schema defines the terms that will be used in RDF statements and
gives specific meanings to them. A variety of schema forms can be used with RDF, including
a specific form defined in a separate document (see [26]) that has some specific characteristics to help with automating tasks using RDF.

A schema is the place where definitions and restrictions of usage for properties are documented. In order to avoid confusion between independent and possibly conflicting definitions of the same term, RDF uses the XML namespace facility. Namespaces are simply a way to tie a specific use of a word in context to the dictionary (schema) where the intended definition is to be found. In RDF, each predicate used in a statement must be identified with exactly one namespace, or schema. However, a Description element may contain statements with predicates from many schemas.

3.1.10.5. Qualified Property values

Often the value of a property is something that has additional contextual information that is considered "part of" that value. In other words, there is a need to qualify property values. Examples of such qualification include naming a unit of measure, a particular restricted vocabulary, or some other annotation. For some uses it is appropriate to use the property value without the qualifiers. For example, in the statement “the price of that pencil is 75 U.S. cents” it is often sufficient to say simply “the price of that pencil is 75”.

In the RDF model a qualified property value is simply another instance of a structured value. The object of the original statement is this structured value and the qualifiers are further properties of this common resource. The principal value being qualified is given as the value of the value property of this common resource.

3.1.10.6. Containers

Frequently it is necessary to refer to a collection of resources; for example, to say that a work was created by more than one person, or to list the students in a course, or the software modules in a package. RDF containers are used to hold such lists of resources or literals.

RDF defines three types of container objects:

Bag:

- An unordered list of resources or literals. Bags are used to declare that a property has multiple values and that there is no significance to the order in which the values are given. Bag might be used to give a list of part numbers where the order of processing the parts does not matter. Duplicate values are permitted.

Sequence:

- An ordered list of resources or literals. Sequence is used to declare that a property has multiple values and that the order of the values is significant. Sequence might be used, for example, to preserve an alphabetical ordering of values. Duplicate values are permitted.

Alternative:

- A list of resources or literals that represent alternatives for the (single) value of a property. Alternative might be used to provide alternative language translations for the title of a work, or to provide a list of Internet mirror sites at which a resource might be found. An application using a property whose value is an Alternative collection is aware that it can choose any one of the items in the list as appropriate.
3.1.10.7. RDF Schema

The RDF Schema (see [26]) introduces a special RDF vocabulary, with some fixed constraints on its interpretation; in particular, the notion of a ‘class’. We can define all of RDFS in terms of a single RDF property rdf:type which relates entities to the classes which contain them.

Descriptions used by these applications can be modelled as relationships among Web resources. The RDF data model, as specified in [27], defines a simple model for describing interrelationships among resources in terms of named properties and values. RDF properties may be thought of as attributes of resources and in this sense correspond to traditional attribute-value pairs. RDF properties also represent relationships between resources. As such, the RDF data model can therefore resemble an entity-relationship diagram. The RDF data model, however, provides no mechanisms for declaring these properties, nor does it provide any mechanisms for defining the relationships between these properties and other resources. That is the role of RDF Schema.

3.1.11. EBU/egta Schema for Advertisement

3.1.11.1. Association of Television and Radio Sales Houses (egta)

egta, the association of television and radio sales houses, is a non-profit trade association based in Brussels. It was founded in 1974 to address the needs of advertising sales houses. In 2002, egta opened to radio sales houses and in 2004 changed from the acronym European Group on Television Advertising to the brand name “egta”. In 1999, the first non-European partner joined egta and the number of members reached 100 in 2007.

Today, egta gathers 38 member radio advertising sales houses in Europe and 10 non-European members from five countries including Canada and Korea besides its 77 member television advertising sales houses in 30 European countries. According to the official homepage [1], 1000 people and more are attending egta’s meetings in the course of one year, in which egta holds 10 events, meetings or conferences. 2000 people are working in sales houses across Europe and other continents who are in regular contact with egta.

egta’s sovereign body is the General Assembly which adopted the egta statutes in 1990 (last modified in 2005). Since egta’s mission is to give sales houses a common voice, guidelines on the position of sales houses in the advertising industry have been formulated on certain topics of importance. Such guidelines are officially adopted by egta members at the association’s Annual General Meeting and can be used by egta members for their national activities. Some examples of these documents are the egta ethical charter or the egta blueprint on the future of television audience measurement which can be found on [30].

egta also cooperates with the European Broadcasting Union (EBU) [31]. Within this cooperation, a collaborative metadata exchange model for the unified exchange of metadata of television ads is currently under development. This “EBU-egta-scheme” will be considered in NoTube since it is currently the only available scheme for ads metadata. In fact, the scheme is under review by partners of the NoTube consortium and feedback will be given to both EBU and egta.

Figure 14: Official egta logo
3.1.12. Liaison between EBU and egta

The European Broadcasting Union as the largest association of national broadcasters in the world promotes cooperation between broadcasters and facilitate the exchange of audiovisual content. The EBU works to ensure that the crucial role of public service broadcasters is recognised and taken into consideration by decision-makers. It provides Technical and Legal expertise. It also coordinates the co-production of audiovisual programmes for television and radio. The EBU is also the Eurovision and Euroradio networks for which it bids for sports and events rights, and operates as a key news provider in Europe.

egta is the Brussels-based trade association of television and radio sales houses that market the advertising space of both public and private broadcasters across Europe and beyond. egta's objective is to address the needs of advertising sales houses, whether individual or common, to help them secure, develop and/or diversify their revenue, and to work hand in hand both with the European institutions and with all actors of the advertising industry to help build a promising future for commercial communications in a fast-changing technological environment.

The business-to-business exchange of advertising spots is following the trend of file based (tapeless) production. It is obvious that the capacity of the contribution networks (including over IP) now support migration to file exchange, realising substantial economy in the current process of transport tapes between facilities. To be efficient this process requires accurate metadata about the advertising spot, the associated rights, the conditions of exploitation and the different parties involved in the process (production, post-production, sale and diffusion).

egta and EBU share a significant cross-membership. It is therefore natural that regular collaboration occurs between the two organisations. Twelve months ago, egta contacted the EBU to seek technical support in the development of a metadata specification. Therefore, a joint working group co chaired by EBU (EU Technical, Jean-Pierre Evain) and egta (Christophe Scherer, FranceTV Publicité) was established. egta members have contributed to the definition of the requirements and EBU has developed the technical metadata specification for advertising spot exchange.

The specification is in its final phase fo review before version 1.0 is co-published by egta and EBU as one EBU Technical Specification (Tech series) in March 2010.

3.1.13. Status of the current schema

The current xml schema for advertisement is under development and revision. Therefore the following explanations are based on this preliminary version of the schema. The focus of the EBU/egta schema is the business to business exchange process of advertisement descriptions. Currently no unified data model is available to store and to exchange the information about advertisement. Therefore the EBU/egta schema should be used to simplify and to ensure the exchange process of advertisement information within and between different IT-based systems.

During the work in NoTube the elaborated requirements for describing advertisements will be reflected into the ongoing development process of the schema. The following chapter describes the main properties of the schema and gives some examples (a complete example xml file for describing an advertisement based on the EBU/egta schema is attached in the annex [see Annex A]).

The schema consists of four different main parts to describe an advertisement. Each part has a different focus to describe an advertisement and each of these main parts has to be present within an instance:
**Descriptive Metadata**: Holds high level information about an advertisement such as a title, keywords, rights, spot text and a description. Furthermore the product could be described by an identifier, a product name, a brand and a product description. Within the following table some elements of the descriptive part are listed with an example (Source: http://www.teslamotors.com/).

<table>
<thead>
<tr>
<th>Descriptive Metadata</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Tesla Roadster - Long distance travel - Down Under</td>
</tr>
<tr>
<td>Description</td>
<td>The new electric driven car Tesla Roadster drives 313 miles (501 km) on single charge during Australians Global Green Challenge.</td>
</tr>
<tr>
<td>Creation date</td>
<td>2010-02-02</td>
</tr>
<tr>
<td>Commercial spot type</td>
<td>TV-Ad</td>
</tr>
<tr>
<td>Product name</td>
<td>Tesla Roadster</td>
</tr>
<tr>
<td>Product description</td>
<td>The all-electric sports car is faster than Porsche 911 or Audi R8 yet is six times as efficient as conventional sports cars. Tesla services cars in its galleries and through “house calls” so owners can enjoy hassle-free service without leaving their home or office.</td>
</tr>
<tr>
<td>Keyword</td>
<td>Car, Electric, Sports car</td>
</tr>
<tr>
<td>Background music title</td>
<td>music title</td>
</tr>
</tbody>
</table>

*Table 3: Example Descriptive Metadata (Excerpt) – Schema for Advertisement*

**Exploitation Metadata**: Holds publication information about the advertisement spot. This includes for example the period of the permitted exploitation and moreover the planned date of publication of the spot.

<table>
<thead>
<tr>
<th>Exploitation Metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validity period</td>
</tr>
<tr>
<td>Planned publication date</td>
</tr>
<tr>
<td>Target delivery Platform</td>
</tr>
<tr>
<td>Authorisation information</td>
</tr>
</tbody>
</table>

*Table 4: Exploitation Metadata (Excerpt) – Schema for Advertisement*

**Credits Metadata**: Holds information about the production of the advertisement spot. This part of the advertisement schema holds e.g. information about the actor, about the producer and also about the post production company. The following table lists some of the elements of the credits information.

<table>
<thead>
<tr>
<th>Credits Metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postproduction company name</td>
</tr>
<tr>
<td>Creative agency company name</td>
</tr>
<tr>
<td>Producer name</td>
</tr>
<tr>
<td>Advertiser company name</td>
</tr>
<tr>
<td>Background music producer</td>
</tr>
<tr>
<td>Background music interpret</td>
</tr>
<tr>
<td>Actor name</td>
</tr>
<tr>
<td>Actor role</td>
</tr>
</tbody>
</table>

*Table 5: Credits Metadata (Excerpt) – Schema for Advertisement*

**Technical Metadata**: Holds technical information about the advertisement spot. This part includes information about the technical video and audio properties. Furthermore subtitles can be part of these information and optional metadata about quality validation. The following table shows some elements with examples of this technical part of the schema for advertisement.
### Technical Metadata

<table>
<thead>
<tr>
<th>Technical Metadata</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>File name</td>
<td>TeslaRoadsterAd-2009-11-09.mxf</td>
</tr>
<tr>
<td>File format</td>
<td>MXF (Material exchange format)</td>
</tr>
<tr>
<td>Audio codec</td>
<td>PCM</td>
</tr>
<tr>
<td>Audio channel main language</td>
<td>English</td>
</tr>
<tr>
<td>Audio loudness level</td>
<td>-23</td>
</tr>
<tr>
<td>Video codec</td>
<td>H.264</td>
</tr>
<tr>
<td>Aspect ratio</td>
<td>16:9</td>
</tr>
<tr>
<td>Picture height</td>
<td>1080</td>
</tr>
<tr>
<td>Picture width</td>
<td>1920</td>
</tr>
<tr>
<td>Scanning format</td>
<td>progressive</td>
</tr>
</tbody>
</table>

*Table 6: Technical Metadata (Excerpt) – Schema for Advertisement*

### 3.2. Consumer Data Models

#### 3.2.1. TV-Anytime

##### 3.2.1.1. General description

TV-Anytime (TVA), a worldwide pre-standardization body, has developed open specifications for the PVR products. The two main objectives of TV-Anytime are:

- To ensure that users have access to personalised content (i.e. according to their specific interests) from the largest possible variety of content providers who, themselves, will benefit from these exchanges.

- To add value by allowing the users to access and use this content when and where they like, without being bound by restrictive access and usage rules.

The recommendations of TV Anytime are to provide:

- Metadata for program descriptions and content referencing:
  - EPGs and other information to help the viewer choose before acquiring content
  - Mechanisms for locating, tracking and record content to help the PVR to effectively acquire the content

- System aspects: Mechanisms for efficient transmission & processing of data to minimize network, CPU and memory resources

- Content Rights and Home Networking

##### 3.2.1.2. Keys of TV Anytime system: the attractors

In TV-Anytime, the metadata is descriptive information about content. It is generically qualified as attractors, as this data is used to attract the user towards content of interest to him/her.

As shown in Figure 15 below, the TV-Anytime schemas form different clusters:

- The program information and segment information schemas are related to a program, and interrelated via the Content Referencing Identifier, the CRID.

- The program location information is also linked to a program via the CRID in particular through the instance description of a broadcast, schedule or on demand event.

- The service information is related to the program location information via the Service Identifier.
• The User information is connected through the User Identifier. Program-specific information is identified by the CRID.

3.2.1.3. Structure

TV-Anytime has defined a unique document structure (see Figure 16 below) to aggregate program descriptions (programme, services, etc.), user descriptions or classification schemes.

In order to help in the development of attractive services, TV-Anytime has defined the following minimum interoperability restrictions:

• A valid description of a programme shall always at least contain a Title.
• All other TV-Anytime metadata is optional. However, TV-Anytime recommends the thorough documenting of the following elements and attributes of a programme description:
  ◦ Synopsis
  ◦ genre (intended audience, content origin, content intention and atmosphere)
  ◦ language
  ◦ member of, and
  ◦ credits list (director, provider, key talent, key character, writer)
3.2.1.4. **Program description**

The TV-Anytime model has been designed to describe the following content concepts:

- Simple programs
- Program versions (e.g., different edits)
- Commercials (program trailers)
- Groups of related programs to describe a program being broadcast in two parts, or a series (ordered or unordered), or related programs sharing common concepts, etc.
- A program that is a concatenation/group of sequences of other programmes identified as an aggregated programme
- The publication of a program that may have publication-dependent attributes (e.g., a film tribute to a recently deceased actor which would have a different description)

In this context, the TV-Anytime CRID links the descriptions to content referenced by this CRID.

3.2.1.5. **Program Location**

The same program may be found in one or more locations according to the results of the location resolution process.

TV-Anytime supports the definition of schedules that will bind programs (using the CRID as the link) to a service. The link to a service is established using the service ID.

TV-Anytime allows the defining of on-demand services that propose a collection of on-demand programs.

3.2.1.6. **Program Segmentation**

Segmentation metadata is an important innovative feature of the TV-Anytime specification. It will allow the creation of richer content. It is believed that its simplicity, in comparison with MPEG-7 segmentation tools, better responds to the physical and operational limitations of broadcasting.

Segmentation refers to the ability to define access and manipulate temporal intervals (i.e., segments) within an audio-visual (AV) stream. It is possible to restructure and re-purpose an
AV stream to generate alternative consumption and navigation modes (e.g. a summary of the content with highlights, or a set of bookmarks, or re-purposing the content for educational purposes) by associating metadata to segments and segment groups.

3.2.1.7. User Description

Usage history enables the tracking and monitoring of user actions while consuming content e.g. record, pause, fast-forward, etc. This can be done to develop a user profile that will be used to select content that matches the user's personal interests. This information can also be shared with a service provider as a means to designing the content better and targeting the programme offers. Different users accessing the same content at different times can be identified separately, if required, by one of the possible applications having access to this information.

User preferences are used to filter content from a large variety of sources according to a rich combination of criteria such as genre, time, date, channel, etc. This operation may require user interaction to confirm a selection. It is also possible for software agents to accurately map the user preferences to media descriptions. It is possible to personalise a profile or a set of preferences by weighting the different selection criteria.

3.2.1.8. System aspect

TV-Anytime defines a set of mechanisms to enable and optimize the delivery of data.

The main technical requirements can be summarized as follows:

- bandwidth efficiency
- capacity for this data to be delivered asynchronously using a carousel
- modularization of the information carried to allow partial and targeted updates, and to enable a certain prioritization in the way the information is cyclically transmitted
- improvement of navigability within this information to provide, when needed, an efficient way to retrieve pieces of priority access

Fragmentation is the generic decomposition mechanism of a TV-Anytime metadata description into self-consistent units of data, called TVA fragments. A fragment is the ultimate atomic part of a metadata description that can be transmitted independently to a terminal. A fragment is self-consistent in the sense that it is capable of being updated independently from other fragments. The way it is processed and accessed is independent of the order in which it is transmitted, relative to other fragments.

Encapsulation is the process that enables the grouping of encoded fragments in containers, ready for transmission. It associates further information to these fragments such as a unique identifier and a version number that enable the monitoring and management of updates.

Indexing is an optional mechanism to deliver TV-Anytime metadata to receivers with limited processing and storage capabilities. It provides a mechanism for locating information from within a fragment stream, forming a TVA metadata description. Indexing structures accompanying a fragment stream provide direct access to each fragment by listing the values of a particular node and describing where the matching fragment(s) can be found within the delivery layer.

3.2.1.9. Contents rights

RMPI stands for “Rights Management & Protection Information” or, in other words, “the minimum set of usage rules and conditions required to enable protection of broadcast digital television content within a TVA
Principles are based on the Positive Assertion of Rights:

- TV-Anytime RMPI-MB rights are positively asserted and never implied.
- These rights are granted to the RMP System component or entity (EG: a domain) and not to a person.
- When a right is exercised, asserted conditions are validated.
- If those asserted conditions are not met, then the right cannot be exercised.
- If conditions are not asserted, then they do not constrain the rights.

3.2.1.10. **Advantage of TV Anytime**

The main advantage of TV Anytime is its independence:

- It is network independent (content/services on DVB, IP, …)
- It is middleware independent: TV-Anytime applications can be built-in the device or delivered as interactive applications

3.2.1.11. **TV Anytime around the world**

The following groups are testing, using or promoting TV Anytime:

- ARIB: ARIB-STD B38: Coding, Transmission and Storage Specification for Broadcasting System Based on Home
- DVB
  - MHP extension for PVRs
  - DVB GBS: transport and delivery of TVA data (ETSI TS 102 323)
  - DVB IPI: Broadband Content Guide
  - DVB CBMS based on TVA datatypes
- Other initiatives
  - UK consortium testing TVA for Freeview
  - Korea producing TVA demonstrators and editing tools
  - IPTC has now based its ProgrammeGuideML metadata specification on TVA
  - US’ Consumer Electronic Association considering the use of TVA
  - Asian and European user groups

3.2.1.12. **Applications of TV Anytime**

TV Anytime is used for several applications such as:

- EPG (Search, Selection, Consumption)
- Trailer booking and recording
- Record related content

3.2.2. **BBC Programmes Ontology**

This ontology aims at providing a simple vocabulary for describing programmes. It covers brands, series (seasons), episodes, broadcast events, broadcast services, etc. Its development was funded by the BBC, and is heavily grounded on previous programmes data modelling work done there. It is described using W3C's RDF technology.
It was developed by Yves Raimond, Patrick Sinclair, Nicholas J Humfrey, Michael Smethurst. It's available on the Web under a Creative Commons license from http://www.bbc.co.uk/ontologies/programmes/2009 (see [32])

A simplified version of the model is as follows:

![Simplified model of Programmes Ontology](image)

*Figure 17: Simplified model of Programmes Ontology*

Data about all BBC programmes is made available using the programmes ontology. For every programme, if you know its 8-digit unique identifier you can get an RDF representation of the data about it (as well as html versions).

For example:

http://www.bbc.co.uk/programmes/b008s9l8 gives an html representation of metadata for episode 2 of Citizen Smith

http://www.bbc.co.uk/programmes/b008s9l8.rdf gives an RDF representation of it.
3.2.3. **EPG / ESG**

3.2.3.1. **General description**

Electronic Service Guide (ESG) is the direct link between the operator, the broadcaster, the content owner and the user. Fast and smart metadata management is the enabling technology.

ESG is a digital guide to scheduled broadcast television or radio programs, typically displayed on-screen with functions allowing a viewer to navigate, select, and discover content by time, title, channel, genre, etc. by use of their remote control, a keyboard, or other input devices such as a phone keypad.

In conjunction with Programme Delivery Control (PDC), content can also be scheduled for future recording by a Digital Video Recorder (DVR) or Personal Video Recorder (PVR).

The on-screen information may be delivered by a dedicated channel or assembled by the receiving equipment from information sent by each program channel. To facilitate the latter method, the European Telecommunications Standards Institute (ETSI) has published the standard ETS 300 707.

3.2.3.2. **EPG Data Model**

EPG will be used to provide programme listings information for both audio and data services and as a mechanism for the user to select services, programmes and related content. A flexible structure has been defined, as shown in Figure 18. The EPG data is broken down into service information (ensembles and services) and programme information (schedules, programmes, groups and events). Additionally programmes and events can be linked together into groups (e.g. for grouping programmes together into serials or series).

Schedule information describes a schedule and its programmes on one or more services for a defined time period. Programmes can also include programme events.

Group information allows programmes to be put into groups. These may be series, serials or just general themes. A hierarchical approach also allows groups to belong to other groups.

*Figure 18: The flexible structure of the EPG Data Model*
Service information includes the structure of and information about the broadcast channel and its associated services.

### 3.2.3.3. ESG Data Model

Electronic Service Guide (ESG) contains information about the services available. Through the information in the ESG, the user can select the services and items he/she is interested in and find stored items on the terminal.

ESG operations take place after the DVB-H receiver has been started and the terminal is synchronized to a particular transport stream carrying IPDC services.

Based on the ESG information rendered to a user through an ESG application, a specific service can be selected. The ESG also provides information which enables the terminal to connect to the related IP stream in the DVB-H transport stream.

![Figure 19: Block diagram of the ESG data model, specified in TS 102 471](image)

The data model is described by an XML Schema. The ESG is subdivided into ESG Fragments, which can be instantiated as parts of the ESG. Figure 19 depicts the ESG Fragments specified in this clause and the relations between them. The indicated cardinalities of the references correspond to the cardinalities specified in the XML Schema definition of the ESG data model.

The ESG is primarily user attractor data, enabling the user to select and acquire content (acquisition information and purchase information have some exceptions to identify the related content respectively the related purchase information).

The **service bundle fragment** is used to group one or more services. The service bundle fragment can then be referenced by a purchase fragment enabling the purchase of rights to the set of services forming the service bundle.

The **service fragment** enables the definition of a service offering. The service fragment provides descriptive and possibly visual information which may be used to attract a user to consume the service, and control information to enable a terminal to configure itself for consumption of the service (this is achieved via a reference to one or more acquisition fragments).

The **schedule event fragment** defines a period of time during which one content item is transmitted on the referenced streaming service. It can also define a period of time during which one or more content items are transmitted on the referenced file download service.
The **content fragment** provides metadata about a content item. This content item may be a video and/or audio asset, or an interactive application. A content item is the smallest entity in the understanding of a non technical person e.g. a video composed from audio and video or an interactive application.

The **acquisition fragment** provides information required for the consumption of the content referenced by a service or schedule event fragment. Specifically the acquisition fragment contains the SDP or the location of an SDP stream which carries the SDP.

The **purchase fragment** is an entity which provides information necessary to gain access to a particular set of services. This information is composed of both attractor information and data to enable the purchase of rights to consume the set of services.

The **purchase channel fragment** provides information about a service provider offering service bundles for purchase. This information includes a description of the provider and its contact points (e.g. telephone number, URL) that can be displayed to the user, or used by the terminal for an automatic purchase. Additional information that is specific to the key management system used by the provider can also be included in this fragment.

### 3.2.3.4. Advantages of EPG/ESG

EPG and ESG allow to:

- Describe the service or content,
- Attach / link media content to it,
- Transmit the service description,
- Help locate the content,
- Market and collect consumption usage,
- Infer user preferences to promote the proper content,
- Insert advertisements / targeted advertisements.

While, satisfying the end user:

- Privacy
- Quality of service
- Awareness of consumption

### 3.2.3.5. Applications of ESG/EPG technology

- Electronic TV-Program guide
  - Receive every morning the TV-Program guide for the next days
  - Select, find your favourite program
  - Download trailers, related material
  - View advertisements
  - Record video clips/ Send email to your PVR
- Video clips electronic guide
  - Access to a large video clips database
  - Sport events, music, news, …
  - Find the clips that best suites to your taste / preview
3.2.4. **Digital Audio: MP3 with ID3-Tags**

The probably most widely used container format holding metadata is MPEG-1 Audio Layer3, better known as MP3. The initial format did not provide a mechanism to denote metadata to a single file (other than the filename itself). Most vendors of hard- and software players do support the additional ID3-Tags. In the latest version, ID3v2.4, more than 80 fields are predefined, including commercial and copyright information and metadata about the essence like album, artist, and song title. There is also a field which can hold an attached picture, e.g. the cover art from an album. The picture should be supplied as PNG or JPG according to the ID3v2.4 specification. (see [33])

If users contribute music or other audio data, there is a high probability that they will use the MP3 format and that the ID3-Tags will hold some information, though not all fields might be used or filled with the correct values.

3.2.5. **Digital Images: Exif and IPTC**

Another widely used consumer format which holds metadata are digital images. Most digital cameras add the technical Exif (Exchangeable Image File Format) metadata, a standard of the Japan Electronic and Information Technology Industries Association (JEITA). Among the data saved are date and time information as well as GPS coordinates (if supported by the camera). This can be particularly interesting for an automated integration into location-based services. (see [34])

If descriptive metadata are desired, the IPTC-NAA-Standard (short: IPTC) can be added to digital images. While Exif data is added automatically, IPTC data has to be manually added by the user. This metadata format contains multiple text fields that can be filled by the user to his/her needs like “Caption/Abstract”, “Keywords”, “Headline” (title). (see [35])

While Exif data is automatically inserted by the camera and a basic set of the values are likely to be present and correct, the contents of the IPTC metadata may or may not be present and correct.
4. Taxonomy management

This chapter will give information about how to define and control any set of vocabulary used in the NoTube system as well as lists of languages and countries. The control mechanisms will be derived from existing and widely used standards to ensure a maximum interoperability. Sources for these standards are the EBU, SMPTE and ISO, which are all used extensively in the broadcasting market.

While WP1 will concentrate on taxonomy for semantic models to enable personalized access of distributed TV content, this deliverable will focus on taxonomy for TV metadata models and content specific descriptions.

4.1. EBU ESCORT 2007

The ESCORT 2007 specification has been developed by the EBU as a system of classification of radio and television programmes. The intention of the specification is to remain a reference for the exchange of comparable data in the area of broadcasting and audience research. Moreover it addresses the fields of finance and accounting, marketing and compliance reporting. (see [36])

The main focus of ESCORT 2007 is the definition of television and radio programme and service concepts and genres. These definitions are used to describe and to classify an service or a programme (e.g. news, social, politics, comedy etc.). The classification for programme and service conceptual data comprises different dimensions:

ESCORT 2007 classification (programme and service conceptual data) (see [36]):

- Intention (e.g. “Entertain”, “Advertise”)
- Format (e.g. “Reality”, “Magazine”)
- Content (e.g. “News and current affairs”, “Education”)
- Intended Audience / Target Group (e.g. “Children”, “Adults”)
- Origination (e.g. “TV”, “Radio”)
- Content Alert (e.g. “Language”)

In general the classification of ESCORT 2007 is build by a hierarchical schema that contains a series of dimensions to describe material (object of multimedia audiovisual experience) in a rising level of detail. This hierarchical schema contains “Required level”, “Recommended level”, and the “Individual level”. The “Individual level” is defined as informative and could be used as additional keywords or replaced by other terms of the user. Table 7 illustrates an example of a hierarchical classification for the conceptional classification of “Content”. (see [36])

<table>
<thead>
<tr>
<th>Required level</th>
<th>Recommended level</th>
<th>Individual level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non fiction / Information</td>
<td>News and current affairs</td>
<td>Daily News</td>
</tr>
<tr>
<td>Sports</td>
<td>Football</td>
<td>Football / Soccer</td>
</tr>
</tbody>
</table>

Table 7: Example of ESCORT 2007 “Content” hierarchical classification

Another important issue of ESCORT 2007 is to describe the languages in which programmes or services delivered. Therefore the specification supports the language description of audio, caption and signs. In this context it is recommended to use the ISO 639-2 language codes for the documentation of the language, which is described in chapter 4.4.. (see [36], [37])
Beside the concepts and genres, ESCORT 2007 provides other elements for description: administrative service information, administrative programme information, programme acquisition data, scheduling data, transmission data, viewing and listening data and financial data. Further information can be found in the specification of ESCORT 2007.

4.2. SMPTE UL (Universal Label)

The “Society of Motion Picture and Television Engineers” (SMPTE) has defined numerous documents, that deal with identifications and labelling. The technique used is called “Universal Labels”, defined in SMPTE 298M. These labels consist of 16 bytes which is used to identify the type and encoding of data and are attached to the data they identify. All labels must be approved by SMPTE before they can be used in data interchange to ensure maximum interoperability between different manufacturers and devices.

The first seven bytes of SMPTE Universal Labels are fixed and defined by SMPTE 400M:

<table>
<thead>
<tr>
<th>Byte position</th>
<th>Description</th>
<th>Value</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Object identifier</td>
<td>06h</td>
<td>Fixed start byte</td>
</tr>
<tr>
<td>2</td>
<td>UL length</td>
<td>0Eh</td>
<td>Byte length of the following label</td>
</tr>
<tr>
<td>3</td>
<td>UL code</td>
<td>2Bh</td>
<td>ISO organization</td>
</tr>
<tr>
<td>4</td>
<td>UL subcode</td>
<td>34h</td>
<td>SMPTE</td>
</tr>
<tr>
<td>5</td>
<td>Registry category</td>
<td>04h</td>
<td>Labels</td>
</tr>
<tr>
<td>6</td>
<td>Registry designator</td>
<td>01h</td>
<td>Labels registries</td>
</tr>
<tr>
<td>7</td>
<td>Structure designator</td>
<td>01h</td>
<td>Labels structure (SMPTE 400M)</td>
</tr>
<tr>
<td>8</td>
<td>Version number</td>
<td>01h-7fh</td>
<td>Incrementing labels registry version number</td>
</tr>
<tr>
<td>9-16</td>
<td>Item designator</td>
<td>Defined by labels registry</td>
<td>See the labels registry</td>
</tr>
</tbody>
</table>

*Table 8: Definition of SMPTE Universal Labels (from SMPTE 400M)*

Bytes 9 through 16 are set up in a tree-like structure and can be differentiated in “nodes” and “leaves”. Nodes may contain subnodes, both of which provide a structure to the registry only and are not to be used for interchange. Only leaves are designed for interchanging data. The Material Exchange Format (MXF) uses the SMPTE Universal Labels defined in SMPTE RP224 which can be used a good example of this structure.

The first 7 bytes are fixed, byte 8 is the version of the registry at the point of the specific labels introduction into the registry. In the example given in Table 9, the tree-like structure for defining an D-10 type video is shown. Nodes are marked in red to show that these entries in the registry are not to be used for file interchange when labelling data, only the last row of Table 9 holds a leaf and can be used.
Table 9: Example of Universal Labels as defined by SMPTE RP224 and used by MXF

In the Material Exchange Format (MXF), Universal Labels are used in every KLV (Key-Length-Value) packet. All data inside MXF (except an optional “Run-In”) is encoded in KLV packets. The “Key” is the 16-byte UL from SMPTE RP224 while the “Length” specifies the length (in bytes) of the following “Value” which holds the data to be interchanged. This ensures unique and common labelling of all data supported by the normative standards and ensures maximum interoperability.

Organizations, vendors, manufacturers and users can propose additions to the registry which will go through the SMPTE standardization process before being added to the document.

4.3. ISO country codes

The International Organization for Standardization has developed a standard for coding countries and other areas of geographical interest. The standard (ISO 3166) is being modified as needed, e.g. after German reunification or after the division of the former Czechoslovakia into the Czech Republic and the Slovak Republic. The standard does not define the country or area names but rather defines an abbreviation for them. The names are taken from UN sources or (for local areas) from relevant official national information sources.(see [38])

The standard consists of three parts, where the latter two are only rarely used. Part one (“Country codes”) contains a two-letter code, a three-letter code and a numeric code. The two-letter code is used as the top-level domain on the internet, for example. There are some exceptions for this rule, Great Britain, for instance, has the ISO 3166 code “GB” but uses the top-level domain “.uk”. The numeric code was mainly designed to aid the usage of the country codes in scripts.

Part two (“Country subdivision code”) holds information about principal subdivisions such as states or provinces. The code defined in ISO 3166-1 is extended by up to three alphanumeric characters. Part three (“Code for formerly used names of countries”) is a kind of log of the standard. It holds a four-letter code for countries that have been deleted from ISO 3166-1. (see [38])
The country codes could be used by the NoTube system to specify the origin of content.

### 4.4. ISO language codes

Similar to the ISO country codes, the International Organization for Standardization also developed a standard to describe languages with a short code. The standard 639 consists of six parts, but not all parts of the standard have yet been approved.

Part one was last updated in 2002 and holds a two letter code for the described languages. New languages are adopted only if ISO has information about enough documents written in that language and terminology exist. Computer-programming languages are not in the scope of this document. Part two of ISO 639 holds a three letter code for the languages and holds a larger number of entries than part one, but also requires certain criteria to be met in terms of usage.

Part three has just been approved in 2007 and holds three letter codes for languages. Different than part one and two, part three attempts to include all known languages. The three letter codes are compatible between part two and three. (see [37])

The last published standard is ISO 639 part five, which holds informations about language families and groups, e.g. all Germanic languages are group together into three groups: East, North and West Germanic languages. (see [37])

The language codes could be used by the NoTube system to specify the languages that are contained in a content item.
5. Assessment (validation of data models)

This chapter reflects the elaboration of the three use case and gives an overview about the requirements for the metadata exchange and the assessment of data models. The collected requirements have been extended according to the 1st version of this deliverable. The different explored data models have been verified on the basis of the elaborated requirements.

5.1. Assessment for use case “Personalised Semantic News” (7a)

The following requirements for a unified NoTube metadata model have been accumulated from the current state of the use case “Personalised Semantic News” (7a). These requirements may change and must be further expanded as the work continues on the use case and may need to be differentiated between Service Provider and Home Ambient.

Within the use case “Personalised Semantic News” the essential role for the metadata exchange has the so called News Item Container (NIC). The NIC is defined as “the object which goes through the system and is enriched along the workflow” (see Deliverable of WP 7a). To clarify which data model can be used to describe this NIC object, the characteristics of it have been analyzed and the adequate requirements collected. This work has been done in close collaboration with the work package 7a members.

The NIC object should contain a univocal identifier and information about legacy information (detected during the ingestion and item detection phase), semantic enriched information and should give the ability to link to user generated information.

Inside the NIC the management of content (e.g. video, audio or text) should be light as possible to avoid an overload of the system. Therefore the metadata format which describes the NIC must have the ability to describe and to reference different physical content. Only text content could be directly present inside the NIC. Therefore a requirement for the data model is to hold and to describe text content. This could be the text which is generated by a speech-to-text analysis tool. Moreover the NIC should hold the technical information to describe how the content is physically made.

The NIC has also the ability to hold information about content which is related to main content inside the same NIC. For that reason the metadata should describe the relationships in the context of physical derivation (original, version), relationships following equivalence of use criteria and also relationships following representativeness criteria. For example, a text that was generated by a speech-to-text software can be considered as equivalent to the original audio, it merely is a different form of representation. A physical derivation could be a cropping and scaling approach. These relationships could be described inside the NIC by using internal different index reference mechanism.

In the following table all elaborated requirements for the use case “Personalised Semantic News” are listed. These requirements have been detected by analysing the characteristics of the NIC. From the use case “Personalised Semantic News” perspective the NIC is composed of three main parts. At first you have the part that holds the generic information about the NIC itself. The second part is describing the main topic of the news (so called Attractor part). The third part holds one or more descriptions about content which is related directly or indirectly to the main topic of the news. This includes e.g. a video as main content and different contents such as radio interviews and press text as main related content. Additionally there are descriptions about the secondary content (indirectly related to the main topic).
The data model which should be used to describe the NIC object should also have the ability to hold the information for the described three main parts. For that reason the requirements for the metadata model are depicted in three tables according to these three main parts of the NIC.

Note: The numbers in the following table are for reference only, they do not intend to be a weighting.

Within the following table, an example of metadata to hold the generic information about the NIC is given.

<table>
<thead>
<tr>
<th>Number</th>
<th>The metadata model should have the ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1.1</td>
<td>Univocal ID for the NIC. (Identifier for the metadata instance)</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Hold the title of the NIC object</td>
</tr>
<tr>
<td>5.1.3</td>
<td>Hold the date of the NIC creation</td>
</tr>
<tr>
<td>5.1.4</td>
<td>Hold the date of the NIC expiration</td>
</tr>
</tbody>
</table>

Table 10: Examples of data inside metadata model by the use case “Personalised Semantic News” – NIC generic information

Describing the main topic (Attractor) of the NIC is given as an example in the following table.

<table>
<thead>
<tr>
<th>Number</th>
<th>The metadata model should have the ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1.5</td>
<td>Hold the title of a News Item</td>
</tr>
<tr>
<td>5.1.6</td>
<td>Describe the subject of a News Item</td>
</tr>
<tr>
<td>5.1.7</td>
<td>Describe the classification (genre) of a News Item (e.g. &quot;Sports news&quot;, &quot;Weather forecast&quot;)</td>
</tr>
<tr>
<td>5.1.8</td>
<td>Hold information about location covered by the News Item</td>
</tr>
<tr>
<td>5.1.9</td>
<td>Hold an abstract (synopsis) of the News Item</td>
</tr>
<tr>
<td>5.1.10</td>
<td>Hold person related data, such as forename, surname etc. This could be also a link to a suitable person description.</td>
</tr>
</tbody>
</table>

Table 11: Examples of data inside a metadata model by the use case “Personalised Semantic News” – Attractor information (main topic)

The NIC container could hold several descriptions for different content. To describe a content the following metadata can be used.

<table>
<thead>
<tr>
<th>Number</th>
<th>The metadata model should have the ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1.11</td>
<td>Describe the type of the content (e.g. &quot;video&quot;, &quot;audio&quot;, &quot;picture&quot;, &quot;text&quot;)</td>
</tr>
<tr>
<td>5.1.12</td>
<td>Hold information about the location covered by the content</td>
</tr>
<tr>
<td>5.1.13</td>
<td>Hold person related data, such as forename, surname etc. This could be also a link to a suitable person description.</td>
</tr>
<tr>
<td>5.1.14</td>
<td>Describe the role of a person (e.g. &quot;creator&quot;, &quot;anchorman&quot;, &quot;subject&quot;)</td>
</tr>
<tr>
<td>5.1.15</td>
<td>Hold the date of content creation</td>
</tr>
<tr>
<td>5.1.16</td>
<td>Hold the date of content expiration</td>
</tr>
<tr>
<td>5.1.17</td>
<td>Hold information about the transmission time</td>
</tr>
<tr>
<td>5.1.18</td>
<td>Hold information about the transmission channel</td>
</tr>
<tr>
<td>5.1.19</td>
<td>Describe awards. This could be also a link to suitable award description.</td>
</tr>
<tr>
<td>5.1.20</td>
<td>Set relations between different descriptions (e.g. link role to a person related data)</td>
</tr>
<tr>
<td>5.1.21</td>
<td>Hold information about the language for audio, captions, subtitles</td>
</tr>
<tr>
<td>5.1.22</td>
<td>Describe timelines for video and audio</td>
</tr>
<tr>
<td>5.1.23</td>
<td>Assign descriptions to a specific timeline event for a specific duration</td>
</tr>
<tr>
<td>5.1.24</td>
<td>Hold information about (copy-, digital-)rights</td>
</tr>
<tr>
<td>5.1.25</td>
<td>Hold information about and to link to external referenced resources</td>
</tr>
<tr>
<td>5.1.26</td>
<td>Reference to the physical content</td>
</tr>
</tbody>
</table>
5.1.27 Hold and describe text content (e.g. text result from speech-to-text tool)
5.1.28 Hold relationships to related content (for physical derivation) which is described in the same NIC.
5.1.29 Hold relationships to related content (for equivalence of use criteria) which is described in the same NIC.
5.1.30 Hold relationships to related content (for representativeness criteria) which is described in the same NIC.

Table 12: Examples of data inside a metadata model by the use case "Personalised Semantic News" – Content description

5.2. Assessment for use case “Personalised TV Guide with Adaptive Advertising” (7b)

The following requirements have been collected from the current state of the use case “Personalised TV guide with Adaptive Advertising” (7b). These requirements may change and must be further expanded as the work continues on the use case.

In “Personalised TV guide with Adaptive Advertising” use case, we need metadata to describe general content and ad (Table 12). These metadata are essential to ensure that users have access to personalised content (i.e. according to their specific interests).

<table>
<thead>
<tr>
<th>Number</th>
<th>Requirement: The metadata model should have the ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.1</td>
<td>Hold the title of the program/ad</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Describe briefly the program/ad</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Hold the program/ad type (Video/Audio/Ad)</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Hold the program/ad genre</td>
</tr>
<tr>
<td>5.2.5</td>
<td>Hold the keywords that describes the program/ad (can be a link to the profile)</td>
</tr>
<tr>
<td>5.2.6</td>
<td>Hold the start time</td>
</tr>
<tr>
<td>5.2.7</td>
<td>Hold the end time</td>
</tr>
<tr>
<td>5.2.8</td>
<td>Hold the duration</td>
</tr>
<tr>
<td>5.2.9</td>
<td>Hold the language (Spoken/subtitle)</td>
</tr>
<tr>
<td>5.2.10</td>
<td>Hold the production year</td>
</tr>
<tr>
<td>5.2.11</td>
<td>Hold the production country</td>
</tr>
<tr>
<td>5.2.12</td>
<td>Hold the credits list (One or more persons/companies related to the TV program (e.g. Director/Producer/Presenter/Author/Actor/Actress)</td>
</tr>
<tr>
<td>5.2.13</td>
<td>Hold the resources list (One or more resources related to the TV program (e.g. trailers, links to related websites)</td>
</tr>
</tbody>
</table>

Table 13: Requirements to a metadata model by the use case "Personalised TV Guide with Adaptive Advertising" – Content description

Technical metadata are essential to technical management of this content (Region Of interest algorithm, Sequence of interest algorithm, Loudness harmonization).

<table>
<thead>
<tr>
<th>Number</th>
<th>Requirement: The metadata model should have the ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.14</td>
<td>Hold the information about the file (name/format)</td>
</tr>
<tr>
<td>5.2.15</td>
<td>Hold the Information about the audio used in the ad (e.g. codec, format, language)</td>
</tr>
<tr>
<td>5.2.16</td>
<td>Hold the information about the video used in the ad (e.g. codec, format, aspect ratio)</td>
</tr>
<tr>
<td>5.2.17</td>
<td>Hold the information about program/ad type (Video/Audio/Ad)</td>
</tr>
<tr>
<td>5.2.18</td>
<td>Hold the information about the text used in the ad (e.g. language, encoding)</td>
</tr>
</tbody>
</table>

Table 14: Requirements to a metadata model by the use case "Personalised TV Guide with Adaptive Advertising" for programme – Technical metadata
For ads, some more information like intended audience (defined by age, cultural/ethnic background, profession …) (Table 13) can be very useful to target the best advertisement to the correct consumer.

<table>
<thead>
<tr>
<th>Number</th>
<th>Requirement: The metadata model should have the ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.19</td>
<td>Hold the general audience</td>
</tr>
<tr>
<td>5.2.20</td>
<td>Hold the age groups (children, age 4-7, …)</td>
</tr>
<tr>
<td>5.2.21</td>
<td>Hold the social groups</td>
</tr>
<tr>
<td>5.2.22</td>
<td>Hold the occupational groups</td>
</tr>
<tr>
<td>5.2.23</td>
<td>Hold the other special interests</td>
</tr>
<tr>
<td>5.2.24</td>
<td>Hold the gender (for men, for women, for both)</td>
</tr>
<tr>
<td>5.2.25</td>
<td>Hold the geographical location (territory for which the programme is primarily intended)</td>
</tr>
<tr>
<td>5.2.26</td>
<td>Hold the education level</td>
</tr>
<tr>
<td>5.2.27</td>
<td>Hold the lifestyle stages (single, couple, divorced, family …, retired)</td>
</tr>
<tr>
<td>5.2.28</td>
<td>Hold the income</td>
</tr>
</tbody>
</table>

Table 15: Requirements to a metadata model by the use case "Personalised TV Guide with Adaptive Advertising" for programme – Target audience metadata

5.3. Assessment for use case “Internet TV in the Social Web” (7c)

The following requirements have been collected from the use case 7c "Internet TV in the Social Web". This use case uses MythTV, an open source software media centre, together with a DVB-T stick and a smart phone used as a remote and 'companion device'. The metadata for the personal programme planning is taken from the DVB stream. Additionally, it is important in this use case to link programmes to and from the web. Thus, special attention needs to be paid to a unique referencing of TV stations and their programmes. Examples for this can be found in the following sections.

5.3.1. General Principles

In general, UC 7c follows the principles and detail of the BBC programmes Ontology. The detail below is not perfect yet and may change.

5.3.1.1. Distinguish Brand, Series, Programme, Version and Broadcast

The Brand (e.g. 'Doctor Who') and the Series (‘Doctor Who season 8’) must be distinguishable in some fashion from the specific programme (‘Doctor Who season 8 programme 1’), and the specific version (‘Doctor Who season 8 programme 1 signed version’) that was broadcast on a specific date (a link between a version, a channel and a date-time). These different conceptual variants must be linked together.

The NoTube aims of this are:

- to be able to link similar but not identical versions of programmes (e.g. my annotation on an English subtitled version of the film 'Intacto' might still be interesting to a Spanish speaker).

- to be able to make simple recommendations based on brands and series

5.3.1.2. Brand, Series, Programme, Version and Broadcast must have unique dereferenceable Web identifiers

This is to link the programme with the Web. Where no web location is available we can for the moment drop the dereferencing requirement, but there still should be a unique URL for every programme.
The requirement for uniqueness is that the NoTube system must be able to identify whether a user has watched a specific programme or not.

The requirement for de-referenceability is that the NoTube system must be able to find out information about a programme even when the video content is not available.

5.3.1.3. **Channels must have unique de-referenceable Web identifiers**

One of the biggest benefits of NoTube could be a set of globally unique identifiers for channels, as that's the part where short names are often used and there is potential for confusion and failed lookups.

The NoTube aim here is to be able to do broadcast lookups in legacy systems.

5.3.1.4. **Good timekeeping**

By good timekeeping we mean care with timezone offsets; precision with date-times of broadcast, and agreement on date-time formatting.

The NoTube aim here is to be able to do broadcast lookups in legacy systems.

5.3.1.5. **Be able to identify points and time-based segments in video**

The NoTube aim here is to be able to annotate parts of Versions of Programmes with tags.

5.3.1.6. **Be extensible**

The NoTube aim is to be able to use other metadata formats where they have specific domain knowledge that we may not want to replicate, for example to be able to include a specific thesaurus for searching.

5.3.2. **Fields**

<table>
<thead>
<tr>
<th>Number</th>
<th>Requirement: The metadata model should have the ability to</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.1</td>
<td>Hold a URL</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Hold a Title</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Hold a Description</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Hold a Detailed Description</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Hold a Category</td>
</tr>
<tr>
<td>5.3.6</td>
<td>Hold a Brand</td>
</tr>
<tr>
<td>5.3.7</td>
<td>Hold a Series</td>
</tr>
<tr>
<td>5.3.8</td>
<td>Hold a Version</td>
</tr>
<tr>
<td>5.3.9</td>
<td>Hold a Duration</td>
</tr>
<tr>
<td>5.3.10</td>
<td>Hold Contributors</td>
</tr>
<tr>
<td>5.3.11</td>
<td>Hold Contributors unique identifiers</td>
</tr>
<tr>
<td>5.3.12</td>
<td>Hold a Contributor role (author, anchor, actor etc.)</td>
</tr>
<tr>
<td>5.3.13</td>
<td>Hold a Broadcaster</td>
</tr>
<tr>
<td>5.3.14</td>
<td>Hold a Microsite</td>
</tr>
<tr>
<td>5.3.15</td>
<td>Hold a Depiction</td>
</tr>
<tr>
<td>5.3.16</td>
<td>Hold a Clip</td>
</tr>
<tr>
<td>5.3.17</td>
<td>Hold Links to external web-based resources</td>
</tr>
</tbody>
</table>

*Table 16: Requirements to a metadata model by the use case "Internet TV in the Social Web" for programme*
Table 17: Additional requirements to a metadata model by the use case "Internet TV in the Social Web" for a brand

<table>
<thead>
<tr>
<th>Number</th>
<th>Requirement: The metadata model should have the ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.18</td>
<td>Hold the Series</td>
</tr>
<tr>
<td>5.3.19</td>
<td>Hold an Episode</td>
</tr>
</tbody>
</table>

Table 18: Additional requirements to a metadata model by the use case "Internet TV in the Social Web" for a series

<table>
<thead>
<tr>
<th>Number</th>
<th>Requirement: The metadata model should have the ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.20</td>
<td>Hold an episode</td>
</tr>
<tr>
<td>5.3.21</td>
<td>Hold a position in a series</td>
</tr>
</tbody>
</table>

Table 19: Additional requirements to a metadata model by the use case "Internet TV in the Social Web" for service

<table>
<thead>
<tr>
<th>Number</th>
<th>Requirement: The metadata model should have the ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.22</td>
<td>Hold the channel</td>
</tr>
<tr>
<td>5.3.23</td>
<td>Hold an episode</td>
</tr>
</tbody>
</table>

Table 20: Additional requirements to a metadata model by the use case "Internet TV in the Social Web" for version

<table>
<thead>
<tr>
<th>Number</th>
<th>Requirement: The metadata model should have the ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.24</td>
<td>Hold the broadcast date and time</td>
</tr>
<tr>
<td>5.3.25</td>
<td>Hold the date and time of next broadcasts</td>
</tr>
<tr>
<td>5.3.26</td>
<td>Hold the channel unique identifier</td>
</tr>
<tr>
<td>5.3.27</td>
<td>Hold the expiry date</td>
</tr>
<tr>
<td>5.3.28</td>
<td>Hold the availability by country</td>
</tr>
</tbody>
</table>

Table 21: Additional requirements to a metadata model by the use case "Internet TV in the Social Web" for a segment

<table>
<thead>
<tr>
<th>Number</th>
<th>Requirement: The metadata model should have the ability to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.29</td>
<td>Hold the start time offset</td>
</tr>
<tr>
<td>5.3.30</td>
<td>Hold the end time offset</td>
</tr>
</tbody>
</table>
6. Metadata handling and format transformation

6.1. Basic principles of model transformations

A model transformation is defined as mapping of a given source model into a target model. One model is appointed the target model, thus setting the direction of the transformation. The properties of source and target model are defined by a source and target meta model, respectively. (see [39], [40])

A meta model defines the structure and semantics of models (see [41]; [39]). A compiled class model can be seen as an example: this class model defines the elements of the model (e.g. classes and attributes) that may be present in an object model. An object model can be contrived as the concrete specificity of a class model, i.e. object models are instances of class models. Instances of classes are objects that have attributes with their concrete values. (see [39])

The transformation of models results from the usage of transformation rules. These transformation rules specify the procedure of creating new model elements in the target model by deriving them from elements of the source model. (see [39]; [42])

Subsequently, a transformation is to be understood as a mapping of instances from a source metadata model to instances of a target metadata model. The resulting instances of a transformation must conform to the structure and semantics of the particular metadata model.

6.2. Metadata model transformations in NoTube

In this section, the metadata transformation in the NoTube system architecture will be explored. In accordance with the current level of knowledge, the needed adaptations of the metadata will be discussed, that are required within the process chain of NoTube. For these considerations, two main components of the NoTube architecture will be examined: Service Provider and Home Ambient. A metadata transformation will have to be done in both components.

From a high level point of view the NoTube System contains two main components. The first one is the Service Provider architecture that is focused on the preparation and delivering of services to the user. The second main component is the Home Ambient architecture which is receiving the delivered content. The Architecture of NoTube is described in detail in deliverable 6.1.

For NoTube it is necessary to select on the one hand a metadata model for the Service Provider side and on the other hand a metadata model for the Home Ambient side which fulfil the requirements from the use cases.

The requirements concerning the metadata models will be different for the Service Provider and the Home Ambient, respectively. In the scope of the end-user, no production specific metadata will be needed, for example. At the current level of knowledge, this suggests that the metadata needed in the Home Ambient could be a subset of the metadata needed in the Service Provider architecture. While implementing the first prototype of the transformation services, it has to be verified, if the chosen metadata formats can handle all requirements.
For further explanations, the following terms will be defined for the metadata models that will be used in the NoTube architecture:

- **NoTube metadata model for Service Provider (NMSP):** This metadata model will be used in the Service Provider ambient and should hold all necessary data to be used in professional TV productions.
- **NoTube metadata model for Home Ambient (NMHA):** This metadata model will be used in the Home Ambient. This model could also be a subset of the model used in the Service Provider domain (NMSP).

In which way the transformations between the formats that will be used in NoTube, will be embraced on subsequently. This focuses on the adaptation between NMSP and NMHA. Furthermore, transformations from external sources and formats into the internal NoTube formats will be explored.

### 6.2.1. Transformation of the selected NoTube metadata models

This chapter will emphasize on the needed transformations of the metadata models, which will be used within the NoTube architecture, thus it will cover an examination of the adaptation from the NMSP format to the NMHA.

**6.2.1.1. Transformation within Service Provider Ambient**

The transformation within the Service Provider Ambient is needed when the metadata should be transmitted from the Service Provider side to the Home Ambient side. During this transformation the relevant information for the user is extracted from professional TV metadata (NMSP) and mapped to the metadata model (NMHA) selected for the Home Ambient environment. In case that within both environments the same metadata model is used the transformation has only to filter the relevant information.

**6.2.1.2. Transformation within Home Ambient**

The transformation within the Home Ambient is needed if the received information submitted by the Service Provider has to be further filtered or adapted according to user preferences that are available only in the Home Ambient. In this case, the transformation or filtering will be done directly on the Home Ambient metadata model (NMHA).

Additionally, information being collected by the user must be taken into consideration. In that case, an enrichment of the Home Ambient metadata format (NMHA) has to be done.

### 6.2.2. Transformation of external sources

In the NoTube architecture and use cases it is foreseen that NoTube-based applications will be receiving the main content from external sources (e.g. broadcasters) and also be able to retrieve additional content from external sources that is correlated to the main content being streamed from the Service Provider to the Home Ambient. The main content may typically be a TV show or a radio broadcast, while additional content may typically be retrieved from Web 2.0 applications such as social networks. For instance, while the user is watching a piece of news about a person the application may present (as retrieved content or a simple link) the Wikipedia page about that person, as well as pictures from Flickr and clips from YouTube. This on the fly mash-up process implies on the one hand a semantic engine that is able to correlate content, and on the other hand an enrichment of the unified reference metadata model, both on the Service Provider side (NMSP) as well as in the Home Ambient (NMHA). The enrichment will be in form of links to the additional content as derived by the semantic engine.

Traditionally metadata has been created and managed by dedicated professionals, namely cataloguers, who possess education and training required to classify content and categorise...
it into predetermined schemes. Metadata created by professionals have high quality, but are expensive in terms of time. This kind of professional metadata can be expected to be delivered together with the main content provided by broadcasters and must be mapped to the internal NoTube metadata format. This will mainly occur on the Service Provider side.

This effort cannot scale and keep up with the vast amounts of new content being produced on the internet. A more viable option is author generated metadata: the creator of the content also provides metadata to describe it. Both professional generated metadata and author generated metadata have a limitation: the intended and unintended content consumers are left out of the process.

A third approach, user created metadata, has become popular with the growth of Web 2.0 applications and social networks and can be described as a folksonomy. Folksonomy is the practice and method of collaboratively creating and managing tags to annotate and categorize content (see [43]). The tagging is typically done in a bottom up fashion in a social environment (shared and open to others) by the user at the time of producing or consuming content. The result of personal free tagging of content for one’s own retrieval and based on author and user generated metadata. Unlike formal taxonomies and classification schemes where terms come from a hierarchy and are explicitly correlated, in folksonomies there is no structure: terms come from a flat namespace, there is no and no relationships among them.

Popular social networks and Web 2.0 applications, such as Facebook, Flickr, YouTube, Panoramio, Del.icio.us, Last.fm etc. are based on folksonomies. While some applications may expose some known metadata fields (for instance Flickr shows part of the EXIF information embedded in uploaded pictures), the bulk of metadata comes from author or user generated tagging. This metadata has to be handled with care, as it need not be correct or complete. It might be necessary to enrich the content gathered from these sources with the aid of the Metadata Acquisition and Enrichment Service in NoTube (see Deliverable 4.1).

Once additional external sources are identified via the services provided by WP4 (see Deliverable 4.2, 4.3 and 4.4) and the Semantic Broker (see Deliverable 5.1), the WP2 metadata service will provide an interface to add these sources to the metadata describing the main content. This metadata instance will be enriched with the URL to the related content. This process will happen on both the Service Provider side (mapping to NMSP) and in the Home Ambient (mapping to NMHA) and allows the provider to send all content retrieved from external sources to the user. In the Home Ambient WP2 services will again allow metadata based content retrieval from external sources, but will also leverage on WP3 profiling services for further user customisation (e.g., considering tags assigned by the individual user who is consuming content, see also D3.1 and D3.2).

6.3. Metadata models selected by NoTube

6.3.1. Models for Personalized Semantic News (7a)

In the use case “Personalized Semantic News” within the NoTube project, the input metadata format will be PrestoSpace and is provided by RAI. This metadata model must be mapped into the NMSP. However, in order to be future proof, a more generic approach has to be considered instead of a simple one-to-one mapping. To enable other broadcasters to connect to the NoTube-system, a uniform interface is needed. This interface must be able to handle various broadcast metadata formats. Therefore, the Broadcast Metadata Exchange Format (BMF) was chosen by WP2 as this interface between the NoTube Service Provider side and external broadcast sources.

As the metadata available during this project will be PrestoSpace, WP2 will first map this into BMF and then use the generic BMF-to-NMSP mapping. With this approach, other metadata formats (e.g. FESAD, a widely used format by the German Public Service Broadcasters) can
easily be added at a later time. A mapping between FESAD and BMF was already developed in the European research project Vitalas.

![Diagram](image)

**Figure 20: Generic transformations of broadcast sources provided by WP2**

After mapping external broadcast sources to the intermediate BMF (or accepting already externally transformed BMF data into the NoTube system), only one uniform transformation to the NMSP and later from NMSP to NMHA is required.

After analysing the requirements of use case 7a, it currently seems feasible that both internal formats (NMSP and NMHA) can be TV-Anytime. This will also ease the integration at the Home Ambient side, as a variety of end-user hardware already supports this format. The NMHA will be a subset of the NMSP, converting the transformation between the two ambients into a filtering.

The choice of TV-Anytime as the metadata format must be closely monitored to recognize promptly any missing elements within it and to expand it if necessary.

### 6.3.2. Models for Personalized TV Guide with adaptive Advertising (7b)

Currently, EBU-egta is the only available candidate for the business-to-business (B2B) exchange of descriptive metadata for advertisements. This professional metadata format has to be converted into NMSP and NMHA. It is likely that advertisement metadata from sources other than the current NoTube partners might be delivered in this format in the future.

As this metadata format is not a broadcast source, mapping into BMF as a generic interface to the NoTube system is neither feasible nor suitable. From the current NoTube perspective and the requirements from the Service Provider and the Home Ambient for adaptive Advertising, again TV-Anytime seems to fulfill the demands. Therefore, WP2 will offer a service to transform EBU-egta to TV-Anytime. Note that there will be no intermediate transformation.

### 6.3.3. Models for Internet TV in the Social Web (7c)

As of now, there is no separate metadata model required for use case 7c and the Internet TV in the Social Web. The only requirement existing is to query the metadata already available from other sources. Therefore, the metadata models selected in the two previous paragraphs are sufficient.
6.4. Service Implementation

6.4.1. Services Description

The analysis of the evolving use cases requirements during the first year of the project lead to an iterative collaboration between the application scenario partners and the WP2. In particular the Personalised News scenario (WP7.a) has been chosen as the starting point for this activity. The first outcome of this process is a collection of metadata-oriented services, part of the envisaged WP2 framework, designed to cover three main goals, namely:

1. Metadata extraction
2. Metadata injection
3. Metadata conversion

Such services are general enough to be made available to the whole platform providing interesting functionalities aimed at making the metadata format issues transparent to the application logic.

The following section provides an overview of the envisaged services for each one of the abovementioned categories. Please note that a detailed description, mainly from a technical perspective, is included in D6.2.

6.4.2. Metadata Extraction

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GetPublicationDate</td>
<td>This method, giving as input a NICid, returns the publication date of the video present into the correspondent NIC as main content. The service will parse the NIC xml file (TvAnytime file) and retrieve this information</td>
</tr>
<tr>
<td>GetPublicationInfo</td>
<td>This method, giving as input a NICid, returns the publication info of the video present into the correspondent NIC as main content. The service will parse the NIC xml file (TvAnytime file) and retrieve this information</td>
</tr>
<tr>
<td>GetGenre</td>
<td>This method, giving as input a NICid, returns a list of genre of the video present into the correspondent NIC as main content. The service will parse the NIC xml file (TvAnytime file) and retrieve this information</td>
</tr>
<tr>
<td>GetContentRef</td>
<td>This method, giving as input a NICid, returns the mms reference to the entire TvShow video which contains the video present into the correspondent NIC as main content. The service will parse the NIC xml file (TvAnytime file) and retrieve this information</td>
</tr>
<tr>
<td>GetTimecodeStart</td>
<td>This method, giving as input a NICid, returns the start timecode of the video present into the correspondent NIC as main content. The service will parse the NIC xml file (TvAnytime file) and retrieve this information</td>
</tr>
<tr>
<td>GetDuration</td>
<td>This method, giving as input a NICid, returns the duration in seconds of the video present into the correspondent NIC as main content. The service will parse the NIC xml file (TvAnytime file) and retrieve this information</td>
</tr>
<tr>
<td>GetLowQualityRef</td>
<td>This method, giving as input a NICid, returns the mms reference to the low quality version of the video present into the correspondent NIC as main content. The service will parse the NIC xml file (TvAnytime file) and retrieve this information</td>
</tr>
<tr>
<td>GetHighQualityRef</td>
<td>This method, giving as input a NICid, returns the mms reference to the high quality version of the video present into the correspondent NIC as main content. The service will parse the NIC xml file (TvAnytime file) and retrieve this information</td>
</tr>
<tr>
<td>GetSpeechToText</td>
<td>This method, giving as input a NICid, returns the speech transcription of the video present into the correspondent NIC as main content. The service will parse the NIC xml file (TvAnytime file) and retrieve this information</td>
</tr>
</tbody>
</table>
GetEnrichmentData: This method, giving as input a NICid, returns all the enrichment data (URI i.e. dbpedia links) associated to the video present into the correspondent NIC as main content. The service will parse the NIC xml file (TvAnytime file) and retrieve this information.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GetEnrichmentData</td>
<td>This method, giving as input a NICid, returns all the enrichment data (URI i.e. dbpedia links) associated to the video present into the correspondent NIC as main content. The service will parse the NIC xml file (TvAnytime file) and retrieve this information.</td>
</tr>
</tbody>
</table>

### Table 22: Methods for metadata extraction

#### 6.4.3. Metadata Injection

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>UploadNIC</td>
<td>This method receives as input the NIC.xml of a news item and save the contained information locally in order to accomplish Data Extraction tasks. The example input could change according to the chosen schema (Prestospace – TvAnytime)</td>
</tr>
<tr>
<td>SetPublicationDate</td>
<td>This method sets the publication date of the video present as main content into the NIC identified by the NICid given as input. The service will parse the correspondent NIC xml file and adds or updates this information.</td>
</tr>
<tr>
<td>SetPublicationInfo</td>
<td>This method sets the publication info of the video present as main content into the NIC identified by the NICid given as input. The service will parse the correspondent NIC xml file and adds or updates this information.</td>
</tr>
<tr>
<td>SetGenre</td>
<td>This method sets the list of genres of the video present as main content into the NIC identified by the NICid given as input. The service will parse the correspondent NIC xml file and adds or updates this information.</td>
</tr>
<tr>
<td>SetContentRef</td>
<td>This method sets the mms reference to the entire TvShow video which contains the video present as main content into the NIC identified by the NICid given as input. The service will parse the correspondent NIC xml file and adds or updates this information.</td>
</tr>
<tr>
<td>SetTimecodeStart</td>
<td>This method sets the start timecode value of the video present as main content into the NIC identified by the NICid given as input. The service will parse the correspondent NIC xml file and adds or updates this information.</td>
</tr>
<tr>
<td>SetDuration</td>
<td>This method sets the duration in seconds of the video present as main content into the NIC identified by the NICid given as input. The service will parse the correspondent NIC xml file and adds or updates this information.</td>
</tr>
<tr>
<td>SetLowQualityRef</td>
<td>This method sets the mms reference to the low quality version of the video present as main content into the NIC identified by the NICid given as input. The service will parse the correspondent NIC xml file and adds or updates this information.</td>
</tr>
<tr>
<td>SetHighQualityRef</td>
<td>This method sets the mms reference to the high quality version of the video present as main content into the NIC identified by the NICid given as input. The service will parse the correspondent NIC xml file and adds or updates this information.</td>
</tr>
<tr>
<td>SetSpeechToText</td>
<td>This method sets the transcription text correspondent to the video present as main content into the NIC identified by the NICid given as input. The service will parse the correspondent NIC xml file and adds or updates this information.</td>
</tr>
<tr>
<td>AddEnrichmentData</td>
<td>This method receives as input a NIC ID and a new enrichment data for the NIC itself. The service will parse the correspondent NIC xml file and adds this external link into the dedicated tag.</td>
</tr>
<tr>
<td>DeleteAllEnrichment</td>
<td>This method deletes all the enrichment data in the NIC identified by the NIC ID given as input. The service will parse the correspondent NIC xml file and deletes all the external link from the dedicated tag.</td>
</tr>
</tbody>
</table>

### Table 23: Methods for metadata insertion
6.4.4. Metadata Conversion

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ConvertMetadata</td>
<td>This method is used to convert a given input metadata format into a specified output metadata format. This interface is specified in a generic way to provide much flexibility according to the different needed transformations. For use case 7a PrestoSpace file will be transformed into a TV-Anytime representation. During this transformation the BMF format is used as the generic intermediate format.</td>
</tr>
</tbody>
</table>

Table 24: Methods for metadata conversions

6.5. Conclusion

Three different metadata transformations can be identified that will be needed in the scope of the NoTube architecture:

On one hand it was shown that a generic interface to the NoTube system is advisable for broadcast sources in order to ensure future extensibility. This interface will be the Broadcast Metadata Exchange Format. Conclusively a transformation between this intermediate uniform model and the model selected for NMSP must be implemented. As shown in chapter 6.3., the format for NMSP that seems to fulfil all current requirements is TV-Anytime. Thus, one transformation to be offered from WP2 Metadata Conversion service is the mapping of BMF to TV-Anytime.

As described in chapter 6.2., the transformation service will offer a BMF interface to handle incoming broadcast metadata sources. Since the use case 7a (Personalized News) will offer PrestoSpace data, this format therefore has to be mapped by the WP2 Metadata Conversion service to the intermediate interface format BMF.

The third identified transformation currently identified to be offered by the Conversion service is from EBU-egta into TV-Anytime.

The adoption from the Service Provider Side to Home Ambient will be only a filtering, as TV-Anytime will be used in both ambients.
7. Conclusion

The deliverable gives an overview of file formats that are being used in the professional and the non-professional field. For simultaneous transport of essence (audio, video, data) and associated metadata, the professional MXF container format has considerable advantages when compared with the other file formats (see chapter 2.1) In which way the content (combined from essence and metadata) will be exchanged between the two “worlds” (Service Provider and Home Ambient) has to be specified in detail in the following work. Just as transmitting the essence combined in one format, it might just as well be transmitted as separate formats for essence and metadata.

In the scope of this deliverable, multiple metadata models have been introduced. The properties and utilisation of those metadata models have been discussed. The exploration of the different models distinguished between models mainly utilised in the professional domain and models mainly used in the consumer domain. Unsurprisingly, the professional metadata models have a higher complexity for describing data. However, due to different specific utilisation fields, even within the professional metadata models there are great differences with regard to level of detail, options for articulation and complexity.

It was shown in chapter 6. that a total of four internal metadata formats is going to be used for the NoTube architecture. On the service provider side, three of those four data models will be used as input formats and then transformed into the unified NoTube metadata format:

a) PrestoSpace as an input metadata format, provided by RAI for use case 7a.

b) BMF, which enables other broadcasters to connect to the NoTube-system. Such a more generic approach has to be considered in order to be future-proof instead of a simple one-to-one mapping of various metadata formats. Thus a uniform interface is needed, which must be able to handle various broadcast metadata formats.

c) EBU/egta metadata specification for advertising spot exchange in the B2B domain. egta members have contributed to the definition of the requirements and EBU has developed the technical metadata specification.

On the Home Ambient side, TV-Anytime, as an internationally agreed and accepted metadata schema in the TV consumer domain has been identified to meet the requirements of the NoTube use cases. To allow an exchange of information and thus the interoperability between those two sides, a metadata transformation is required. Chapter 6. covers this transformation as well as the transformation of external sources into the NoTube metadata formats.

Requirements to the internal metadata formats (Service Provider and Home Ambient) were decided upon, based on the description of the use cases (see chapter 6.3.). Based on these requirements, the appropriate metadata schema and required transformations have been selected by WP2 members and described in this second version of the deliverable.
8. References

4: Mücher, M., Broadcast Fachwörterbuch, 2004
5: Tudor, P., Advance Authoring Format, 2004
6: Fileformate für die vernetzte Fernsehproduktion - Die Bedeutung von Dateiformaten aus Sicht der TV Produktion, - 2002
7: 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects Transparent end-to-end packet switched streaming service (PSS); 3GPP file format, 2005
8: DVB Project Office, DVB Fact Sheet, 2009
9: Institut für Rundfunktechnik, BMF Broadcast Metadata exchange Format Description of the class model., 2009-02-05, http://www.irt.de/bmf
10: Austausch von Metadaten: Das "Broadcast Metadata exchange Format", - 2005
12: Special Issue on MPEG-7: IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE - 2001
14: ISO/IEC, Information technology - Coding of audio-visual objects - Part 20: "Lightweight Application Scene Representation (LASeR) and Simple Aggregation Format (SAF)", 2006
16: EBU, P_META 2.0 Metadata Library, EBU TECH 3295-v2, 2007
18: Wright, R., Some considerations on using P_META and Dublin Core, 2003
19: EBU, EBU Core Metadata Set, 2008
20: SMPTE380M (approved) Standard for Television - Material Exchange Format (MXF) - Descriptive metadata scheme-l (Standard, Dynamic), SMPTE - 2004
21: SMPTE RP210 Metadata Dictionary, SMPTE - 2005
24: NewsML 2 Architecture, IPTC - 2005
33: Nilsson, M., ID3 tag version 2.4.0 - Main Structure, 2000
39: Marschall, Frank, Modelltransformation als Mittel der modellbasierten Entwicklung von Softwaresystemen, 2005
41: Bichler, L., Codegeneratoren für MOF-basierte Modellierungssprachen, 2004
9. Annex A

This is a (non-authorized) example of a metadata instance in EBU/egta format.

```xml
<adsmeta:adsMeta>
  <adsmeta:documentId>001</adsmeta:documentId>
  <adsmeta:dateLastModified>2009-11-09</adsmeta:dateLastModified>
  <adsmeta:version>0.1</adsmeta:version>
  <adsmeta:xsi:schemaLocation>
    urn:ebu:metadata:schema:ebu_egta_adsmeta_20091012 EBU_EGTA_AdsMeta_20091012.xsd
  </adsmeta:xsi:schemaLocation>
  <xmlns:xsi=http://www.w3.org/2001/XMLSchema-instance>
    <xmlns:adsmeta=urn:ebu:metadata:schema:ebu_egta_adsmeta_20091012>
      <adsmeta:descriptiveMetadata>
        <adsmeta:title>Tesla Roadster - Long distance travel - Down Under</adsmeta:title>
        <adsmeta:description>The new electric driven car Tesla Roadster drives 313 miles (501 km) on single charge during Australia's Global Green Challenge.</adsmeta:description>
        <adsmeta:backgroundMusicTitle>Highway to hell</adsmeta:backgroundMusicTitle>
        <adsmeta:version>1.1</adsmeta:version>
        <adsmeta:duration>30</adsmeta:duration>
        <adsmeta:creationDate>2009-11-09</adsmeta:creationDate>
        <adsmeta:commercialSpotType>TV-Ad</adsmeta:commercialSpotType>
        <adsmeta:campaign>Global Green Challenge Campaign</adsmeta:campaign>
        <adsmeta:product>
          <adsmeta:productName>Tesla Roadster</adsmeta:productName>
          <adsmeta:productDescription>The all-electric sports car is faster than Porsche 911 or Audi R8 yet is six times as efficient as conventional sports cars. Tesla services cars in its galleries and through “house calls” so owners can enjoy hassle-free service without leaving their home or office.</adsmeta:productDescription>
        </adsmeta:product>
        <adsmeta:spotText>Tesla Roadster owner Simon Hackett and co-driver Emilis Prelgauskas completed 313 miles (501 km) in a Tesla Roadster on a single charge – a distance that appears to set a new record for a production electric vehicle.</adsmeta:spotText>
        <adsmeta:keyword>Car</adsmeta:keyword>
        <adsmeta:keyword>Electric</adsmeta:keyword>
        <adsmeta:keyword>Sports car</adsmeta:keyword>
        <adsmeta:keyword>Environmentally friendly</adsmeta:keyword>
        <adsmeta:keyword>Fuel saving</adsmeta:keyword>
      </adsmeta:descriptiveMetadata>
      <adsmeta:exploitationMetadata>
        <adsmeta:validityPeriod>
          <adsmeta:validFrom>2009-11-09</adsmeta:validFrom>
          <adsmeta:validTo>2009-12-24</adsmeta:validTo>
        </adsmeta:validityPeriod>
        <adsmeta:plannedPublicationDate>2009-11-10</adsmeta:plannedPublicationDate>
        <adsmeta:targetDeliveryPlatform>TV</adsmeta:targetDeliveryPlatform>
        <adsmeta:authorisationInformation>
          <adsmeta:clearanceFlag>true</adsmeta:clearanceFlag>
        </adsmeta:authorisationInformation>
      </adsmeta:exploitationMetadata>
      <adsmeta:technicalMetadata>
        <adsmeta:SoM>12:00:00.0Z</adsmeta:SoM>
        <adsmeta:EOm>13:00:00.0Z</adsmeta:EOm>
        <adsmeta:filename>TeslaRoadsterAd-2009-11-09.mxf</adsmeta:filename>
        <adsmeta:fileFormat>MXF</adsmeta:fileFormat>
        <adsmeta:audioCodec>PCM</adsmeta:audioCodec>
        <adsmeta:audioSystemName>PCM</adsmeta:audioSystemName>
        <adsmeta:channelDescription>
          <adsmeta:channelId>0</adsmeta:channelId>
        </adsmeta:channelDescription>
      </adsmeta:technicalMetadata>
    </xmlns:adsmeta>
  </xmlns:xsi>
</adsmeta:adsMeta>
```