Towards a Unified Multimedia Metadata Management Solution

Samir Amir, Ioan Marius Bilasco, Md. Haidar Sharif, Chabane Djeraba
University of Sciences and Technologies of Lille (USTL– Lille1), France

ABSTRACT
With increasing use of multimedia in various domains, several metadata standards appeared these last decades in order to facilitate the manipulation of multimedia contents. These standards help consumers to search content they desire and to adapt the retrieved content according to consumers' profiles and preferences. However, in order to extract information from a given standard, user must have a pre-knowledge about this latest. This condition is not easy to satisfy due to the increasing number of available standards. In this book chapter we introduce some of the main de facto multimedia standards which cover the description, by means of metadata, of the content and of the use context (profiles, devices, networks…). We discuss then the benefits of proposing an integrated vision of multimedia metadata standards through the usage of a generic multimedia metadata integration system and we expose the challenges of its implementation.

Keywords
Multimedia, metadata integration, Semantic Web, ontology matching.

INTRODUCTION
Nowadays, with the vast expansion of the World Wide Web, several standards (such as MPEG-7(Chang, 2001), MPEG-21(Pereira, 2001), TV-Anytime (TV-Anytime Forum, 2003), etc.) have appeared for enhancing the retrieval, the usage and the delivery of multimedia data over a variety of channels (Web, TV, mobile). Those standards introduce descriptions of the content itself and of the context in which the content was created or for which the content was designed. We call these descriptions *metadata* as they bring new knowledge about the content and the context seen as regular *data*. The metadata presented in various multimedia standards describe different kinds of multimedia contents (e.g., video, image, audio, etc.), devices consuming or transmitting these contents (e.g., networks, TV, mobile, etc.), services processing or dealing with them (e.g., search, adaptation, etc.) and finally environment of user consuming these contents (e.g., user profile, user preference, etc.).

The first category of metadata presented here above, which is about content, can be found in multimedia standards such as (MPEG-7, Dublin Core (WEIBEL, 1998) or TV Anytime), as well as in different ongoing research projects such as the one defended by the *CAM4Home Consortium*[1] which proposes a dedicated content description schema called *CAM Core Metadata* (ITEA2-CAM4Home, 2008). This kind of metadata provides explicit knowledge about the features of the content (genre, appearing concepts, etc.) and about the physical properties of the content (required bandwidth, required decoders, file size, resolution, etc.). This knowledge improves the search processes as it enriches the signal-based characterization of content, with explicit knowledge closer to user criteria (such as, meaning-full keywords). The content-related metadata can also be used in order to propose adequate adaptation processes as, depending on the type of content, specific techniques might apply better.

---

[1] www.cam4home-itea.org

The last three categories presented above are about the context in which the delivery of multimedia content takes place. Standards such as MPEG-21, CC/PP (KLYNE, 2004) or description schemes like CAM Supplementary Metadata schema proposed by the *CAM4Home Consortium*[1] cover context-related information. These metadata offer knowledge that can also be injected in search, retrieval and delivery processes. While doing search, systems could benefit from the information about the user access device in order to propose content that are compatible. While doing delivery, systems can interpret the capacity of the access device and the capacity of the delivery network and it can use this information in order to adapt by simplifying accordingly the content (doing transcoding or transrating for videos, doing resolution reduction for images, doing filtering for complex documents such as web pages, etc.).

Considering the current state of art with regard to multimedia content and context descriptions, standards and consortium initiatives, taken all together, cover fairly well all aspects of the multimedia delivery problem. However, in order to take advantage of these entire standards one must have a strong and a very diversified pre-knowledge about a part or all of them. Besides the specific encoding proposed by each solution, those standards that are often created by specific multimedia communities (such as Multimedia Pictures Experts Groups – MPEG[2], World Wide Web Consortium - W3C[3], Dublin Core – DC[4], CAM4HOME Consortium – C4H[5],…) have led to the availability of multiple terminological and syntactical resources in numerous domain of multimedia. These different types of metadata are encoded using existing description languages (e.g., XML Schema (Thompson, 2009) for MPEG-7 and TV-Anytime or RDF Schema (Brickley, 2004) for CAM Core Metadata and CAM Supplementary Metadata), different vocabularies and different ontologies depending on the community that has created them. These multimedia contents can also be enriched by other kind of metadata such as consumers' annotations (e.g., comments, social tags, etc.) which is free text added by consumers having different point of view and different understanding about multimedia content.

Dealing with knowledge from multiple independent metadata standards is one of the most important challenges in multimedia domain due to the semantic heterogeneity of information as mentioned here above. The creation, the delivery and the consumption of rich multimedia experiences between different entities in multimedia community (e.g., multimedia content consumers, commercial content providers, simple producer, etc.) requires that each of these entities must be able to interpret all metadata standards used by the community. However this requirement is not easy to satisfy due to numerous standards which appeared and will appear in multimedia community.

An important prerequisite solution for this problem is to establish interoperability between the various metadata representations. This can be done by defining a framework for mapping and converting between the different metadata representations. The framework described in this chapter is a multimedia metadata integration system we planned to integrate as part of CAM4Home project to achieve both syntactical and semantic interoperability between different multimedia metadata standards used by multimedia community.

This chapter first gives an overview about existing multimedia metadata standards and CAM4Home project initiative (introduced above) that covers a wide area of information related to multimedia delivery and includes multimedia content description, user preference and profile description and devices characteristic description. Then we relate about multimedia and generic integration issues by discussing the work done by W3C working group in order to integrate heterogeneous metadata and some generic approaches providing mapping between ontologies (Kalfoglou, 2002) (Doan, 2002) (Doan, 2000).

[2] www.mpeg.org
[3] www.w3.org
[4] www.dublincore.org
[5] www.cam4home-itea.org

The second part of the chapter is consecrated to the illustration of the proposal of a new architecture for the multimedia metadata integration system and discuss about challenges of its realization. We first deal with the homogenization of different encoding styles (XML Schema, RDF Schema, and OWL (McGuinness, 2004)). Then we propose a solution for semantic disambiguation of metadata descriptions by explicitating the meaning of a concept with regard to its definition context. Several meanings can be associated with the same concept and a (definition) context-based filtering process is proposed. Finally, we expose the validation of mappings suggested by semantic proximity measured between concepts after the disambiguisation phase. We use the expressive power of Description Logic (Baader, 2003) in order to describe assumption on the possible mapping. The validation is proven or invalidate by using a version of Tableau algorithm (Baader, 2001).

MULTIMEDIA METADATA STANDARDS

In this section we briefly present some standards and consortium initiatives related to content description (MPEG-7, TV-AnyTime, CAM Core Metadata model) and context description (MPEG-21, CC/PP, CAM Supplementary Metadata model (Bilasco, 2010)). The aim of this section is two-fold. It allows unfamiliar reader to get acquainted with these standards and it also illustrates the heterogeneity of multimedia metadata encoding using specific community semantics and using specific schema languages (XML Schema and RDF Schema). We start by presenting the content-related metadata standards.

Content-related Metadata Standards

We have selected in this section three XML Schema based standards (MPEG-7, TV-Anytime and Dublin Core) as well as a description schema designed within the CAM4Home project that is defined using RDF Schema.

We have chosen three **XML Schema** based standards in order to differentiate between standards with a very complex structural organization such as MPEG-7 composed of hundreds of descriptors defined by inheritance and yielding complex compositions, light-weight standard such as Dublin Core having moreover a simple and linear structure and a small set of concepts and intermediate with regard to schema complexity such as TV-Anytime. The choice of TV-Anytime is also guided by the fact that the TV community that is slightly different from MPEG community, and we are interested in how similar concepts were brought in standards by different communities having different cultural backgrounds.

We have selected XML Schema based standards and **RDF Schema** in order to illustrate the heterogeneity that can arouse with respect to the underlying specification schema languages and the need of migrating these underlying schema languages to a common format.

MPEG-7

From the standardization efforts of the Working Group *Moving Picture Experts Group* (MPEG) emerged in 2001 the MPEG-7 specifications. MPEG-7 is a standard that covers the semantic description of media resources. Although the descriptors proposed in MPEG-7 specifically cover the multimedia resources of type audio and video, MPEG-7 is extensible and can include other types of media. MPEG-7 provides a series of audio-visual description tools that are grouped into the following categories: descriptors (**D**), description schemas (**DS**) and description definition language (**DDL**).

A descriptor is a unit of indexing features describing the primary visual, audio or object semantic. The description schemes, which are high-level descriptors, include several D and other DS on structured and semantic units. The DDL defines the syntax for creating new DS. Since DDL is derived from XML Schema, DDL allows the extensibility of the MPEG-7 (Bilasco, 2005).

DS defined in MPEG-7 currently cover the following categories: visual description (VDS), audio description (ADS) and structural description of multimedia content (MDS). VDS and ADS describe the physical, logical or semantic structure of multimedia document. These structures are built using the DS offered by MDS. Figure 1 shows a fragment of MPEG-7 standard specification corresponding to the structural definition of the `VideoSegmentType` that is part of the MDS scheme. The construction of mapping between several standards must take into account the eventual lexical proximity (such as synonymies) between label names (such as `VideoSegment`*)* as well as the specificities of each underlying definition language. By examining the XML Schema fragment presented in Figure 1 we can observe that the `VideoSegmentType` is an extension of a generic `SegmentType` and it is composed by a sequence of properties corresponding to other instances of the MPEG-7 Structural (MDS) and Visual DS (VDS). Each property has a given cardinality. All these information constitutes the context of the `VideoSegment` type definition. This kind of information must be taken into account during the mapping process as it would allow the filtering among “false friends” mapping candidates.

```
<complexType name="VideoSegmentType">
 <complexContent>
  <extension base="mpeg7:SegmentType">
   <sequence>
    <choice minOccurs="0">
     <element name="MediaTime" type="mpeg7:MediaTimeType"/>
     <element name="TemporalMask" type="mpeg7:TemporalMaskType"/>
    </choice>
    <choice minOccurs="0" maxOccurs="unbounded">
     <element name="VisualDescriptor" type="mpeg7:VisualDType"/>
     <element name="VisualDescriptionScheme"
              type="mpeg7:VisualDSType"/>
     <element name="VisualTimeSeriesDescriptor" type="mpeg7:VisualTimeSeriesType"/>
    </choice>
    <element name="MultipleView" type="mpeg7:MultipleViewType" minOccurs="0"/>
    <element name="Mosaic" type="mpeg7:MosaicType" minOccurs="0" maxOccurs="unbounded"/>
    <choice minOccurs="0" maxOccurs="unbounded">
     <element name="SpatialDecomposition" type="mpeg7:VideoSegmentSpatialDecompositionType"/>
     <element name="TemporalDecomposition"
              type="mpeg7:VideoSegmentTemporalDecompositionType"/>
     <element name="SpatioTemporalDecomposition”
              type="mpeg7:VideoSegmentSpatioTemporalDecompositionType"/>
     <element name="MediaSourceDecomposition"
              type="mpeg7:VideoSegmentMediaSourceDecompositionType"/>
    </choice>
   </sequence>
  </extension>
 </complexContent>
</complexType>
```

Figure 1: Example of MPEG-7 description scheme: Definition of VideoSegmentType

TV-Anytime

TV-Anytime Forum is an association of organizations which seeks to develop specifications to enable audio-visual and other services based on mass-market high volume digital storage in consumer platforms. They have promoted a homonym standard, TV-Anytime which is a multimedia metadata standard that allows consumers to select and acquire content of interest.

TV-Anytime standard facilitates user browsing, selection and acquisition of content independently from their system of distribution, even if they are enhanced TV, interactive TV (ATSC, DVB, DBS and others) or Internet. TV-Anytime standard is a specification linking the creators and content providers to consumers. TV-Anytime allows the access to different views, versions or editions of a particular topic and it gives a short description for contents. User is able to add personal metadata, as annotation, comments, tags which can be useful for search services. To ensure compatibility, TV-Anytime has adopted a common representation format for the exchange of metadata. The adopted representation format is XML Schema.

The definition of the `ProgramInformationType` is presented in the TV-Anytime fragment shown in Figure 2. The structural organization of this concept is notably less important than the `VideoSegment` type of MPEG-7 presented in the previous section. No inheritance information about eventual parents is provided. The properties are defined by a single sequence construct with no alternative constructs. The mapping of this concept to other concepts in other standards would be more straightforward than the mapping of a `VideoSegment`, however ambiguities might arise as many “mapping” candidates could be identified if less constraints on the structure are defined.

```
<complexType name="ProgramInformationType">
    <sequence>
     <element name="BasicDescription" type="tva:BasicContentDescriptionType"/>
     <element name="OtherIdentifier" type="mpeg7:UniqueIDType" minOccurs="0"
              maxOccurs="unbounded"/>
     <element name="AVAttributes" type="tva:AVAttributesType" minOccurs="0"/>
     <element name="MemberOf" type="tva:BaseMemberOfType" minOccurs="0"
              maxOccurs="unbounded"/>
     <element name="DerivedFrom" type="tva:DerivedFromType" minOccurs="0"/>
     <element name="EpisodeOf" type="tva:EpisodeOfType" minOccurs="0"/>
     <element name="PartOfAggregatedProgram" type="tva:CRIDType" minOccurs="0"/>
     <element name="AggregationOf" type="tva:AggregationOfType" minOccurs="0" />
   </sequence>
   <attribute name="programId" type="tva:CRIDType" use="required"/>
   <attributeGroup ref="tva:fragmentIdentification"/>
   <attribute name="metadataOriginIDRef" type="tva:TVAIDRefType"use="optional"/>
   <attribute ref="xml:lang" default="en" use="optional"/>
</complexType>
```

Figure 2: Example of TV-Anytime specification: Program Information type

Dublin Core

The Dublin Core (WEIBEL, 1998) is one of the more compact metadata standards which have very simple structure. It is used to describe information resources. It defines conventions for describing things online in ways that make them easy to find. Officially it contains fifteen elements (title, creator, subject, description, publisher, contributor, date, type, format, identifier, source, language, relation, coverage, and rights).

The Dublin Core is already used to describe digital materials such as video, sound, images, text, and composite media, like web pages. These fifteen elements were deliberately made simple so that non-library catalogers could provide basic information for resource discovery. Because of its simplicity, the Dublin Core has been used with other types of materials, and for applications demanding increased complexity. Its design, which allows for a minimum set of shareable metadata in the Open Archive Initiative-Protocol for Metadata Harvesting, prove its efficiency as thousands of projects worldwide, use the Dublin Core for cataloging, or collecting data from the Web. Figure 3 shows the fifteen elements of Dublin Core standard which are encoded using XML Schema.

```
<xs:element name="title" substitutionGroup="any"/>
<xs:element name="creator" substitutionGroup="any"/>
<xs:element name="subject" substitutionGroup="any"/>
<xs:element name="description" substitutionGroup="any"/>
<xs:element name="publisher" substitutionGroup="any"/>
<xs:element name="contributor" substitutionGroup="any"/>
<xs:element name="date" substitutionGroup="any"/>
<xs:element name="type" substitutionGroup="any"/>
<xs:element name="format" substitutionGroup="any"/>
```

```
<xs:element name="identifier" substitutionGroup="any"/>
<xs:element name="source" substitutionGroup="any"/>
<xs:element name="language" substitutionGroup="any"/>
<xs:element name="relation" substitutionGroup="any"/>
<xs:element name="coverage" substitutionGroup="any"/>
<xs:element name="rights" substitutionGroup="any"/>
```

Figure 3: Example of Dublin Core specification

The structural and the linguistic simplicity of Dublin Core make it very easy for the mapping. Its semantic is clearly described, calling only external resources without any need to take into account the structure of the concept, is sufficient to link between proprieties defined in other schemas and those specified in the Dublin Core standard.

CAM4Home Core Metadata

CAM4Home Core Metadata is a part of the CAM4Home Metadata model proposed by the CAM4Home Consortium in order to deal with multimedia content in the era of Web, TV and mobile convergence. The objective of the CAM4Home consortium is to create a metadata enabled content delivery framework to allow end users and commercial content providers to create and deliver rich multimedia experiences. These multimedia experiences are based on a novel concept of collaborative aggregated multimedia. The Collaborative Aggregated Multimedia (**CAM**) refers to aggregation and composition of individual multimedia contents (called objects) into a content bundle that may include references to content based services and can be delivered as a semantically coherent set of content and related services over various communication channels. The consortium develops one common metadata framework for CAM content that can be applied for both personal and commercial applications and is interoperable with relevant standard metadata and content representation technologies.

CAM4Home Core Metadata is part of the CAM4Home Metadata Framework and offers complete descriptions of the structure and behavior of core CAM entities which are necessary to represent and manipulate simple or aggregated multimedia content. CAM Core Metadata supports the representation of a wide variety of Multimedia content in CAM Objects: downloadable applications, software services, images, video, etc. Specific metadata is attached to different types of Multimedia entities. This metadata describes both the content file or service deployment method and the actual content or service that is provided. This metadata model also describes the mechanisms by which CAM Bundles aggregate CAM Objects.

A fragment of the CAM4Home Core Metadata specification done using **RDF Schema** is shown in Figure 4. The example illustrates the generic RDF Schema templates for introducing simple and structured core metadata constructs to specialize the abstract concepts into concrete metadata elements. Any simple metadata used for content feature description is directly associated as a property which specializes the `simpleFeaturePropertyrdf:property` (lines 06-08 in Figure 4). Any structured metadata used for content feature description extends the `ContentFeatureMetadata` class (lines 9-11 in Figure 4). It is associated with a given concrete feature container by introducing a property that links the container and the newly created structured metadata. A naming convention has been adopted in order to make the property name reflect the metadata construct they are introducing. Complex properties use *has* prefix before the actual class name. In the example shown in Figure 4, `hasAppearingConcept` is a complex property in that it uses another class definition in the metadata model (the `AppearingConcept` class).

The main difference with the XML Schema based standards is that the class definition can be done in separate places over an RDF Schema document as well as for XML Schema language all knowledge about a concept is contained within the main definition of the concept (`xsd:complexElement`). For instance, all properties of a concept are defined independently (line 04-08 for `core: title` and line 12-16 for `core:hasAppearingConcept`). Hence there is a need to have a global approach in describing the complete definition of a concept and not considering only a local approach (with regard to the specification document definition). Another specificity of RDF Schema is that it may introduce refined type of composition. In XML Schema world, each property of a complex element can be seen *isPartOf* relation. However, this might not always apply. In RDF Schema properties can be typed and hence more elaborate relation can be transposed without any loss of information into the RDF Schema descriptions.

```
01: <rdfs:Class rdf:about="&core;CAMElementMetadata"  rdfs:label="core:CAMElementMetadata">
02:  <rdfs:subClassOf rdf:resource="&abstract;ContentFeatureContainer"/>
03: </rdfs:Class>
04: <rdf:Property rdf:about="&core;title rdfs:label="core:title">
05:  <rdfs:domain rdf:resource="&core;CAMElementMetadata"/>
06:  <rdfs:subPropertyOf rdf:resource="‶&abstract;simpleFeatureMetadata"/>
07:  <rdfs:range rdf:resource="&xsd;string"/>
08: </rdf:Property>
09: <rdfs:Class rdf:about="&core;AppearingConcept"  rdfs:comment="A class for representing an
                         AppearingConcept"   rdfs:label="core:AppearingConcept">
10:  <rdfs:subClassOf rdf:resource="&abstract;ContentFeatureMetadata"/>
11: </rdfs:Class>
12: <rdf:Property rdf:about="&core;hasAppearingConcepts ‶ rdfs:label="core:hasAppearingConcepts">
13:  <rdfs:range  rdf:resource="&core;AppearingConcept"/>
14:  <rdfs:domain rdf:resource="&core;MultimediaElementMetadata"/>
15:  <rdfs:subPropertyOf rdf:resource="&abstract;hasFeatureMetadata"/>
16: </rdf:Property>
```

Figure 4: RDF-S definitions of concrete classes and properties

Context-related Metadata Standards

In the previous section we introduced some metadata standards related to the description of multimedia content. The previous section offered a view on how different encoding technologies impact on the structure and the organization of metadata standards. In the following we discuss some metadata standards which introduce concepts for characterizing the delivery context. The aim of this section is to illustrate how different communities (MPEG, CAM4Home and W3C) perceive and models through metadata the delivery context. In all the standards which we present here after (respectively MPEG-21, CAM4Home Supplementary Metadata, CC/PP) concepts like user device, access network, environmental context and theirs properties are represented using various constructs depending on the encoding scheme considered.

MPEG-21

MPEG-21 *Multimedia Framework* proposes solutions for the use of multimedia resources across a wide variety of networks and access devices. To achieve this, MPEG-21 addresses the standardization of content management, reuse content in new contexts of use, protection of rights of privacy of consumers and providers of multimedia content, etc. MPEG-21 is built on top of the family of MPEG standards (MPEG-1, MPEG-2, MPEG-4, MPEG-7) drawing particularly on MPEG-4 regarding the dissemination of information and MPEG-7 semantic description for content.

```
<complexType name="TerminalCapabilitiesType">
 <complexContent>
  <extension base="dia:DIAUsageEnvironmentBaseType">
   <sequence>
    <element name="TerminalCapabilities" type="dia:TerminalCapabilitiesBaseType minOccurs="0"
     maxOccurs="unbounded"/>
   </sequence>
   </extension>
 </complexContent>
</complexType>
```

```
          .....
10: <complexType name="InputOutputCapabilitiesType">
11:  <complexContent>
12:   <extension base="dia:TerminalCapabilitiesBaseType">
13:    <sequence>
14:     <element name="Display" type="dia:DisplayCapabilitiesType"  minOccurs="0"
         maxOccurs="unbounded"/>
15:     <element name="AudioOut"  type="dia:AudioOutputCapabilitiesType" minOccurs="0"/>
16:     <element name="UserInteractionInputSupport" type="dia:UserInteractionInputSupportType "
         minOccurs="0"/>
17:    </sequence>
18:   </extension>
19:  </complexContent>
20: </complexType>

21: <complexType name="DisplayCapabilitiesType">
22:  <complexContent>
23:   <extension base="dia:DIABaseType">
24:    <sequence>
25:     <element name="Resolution" type="ResolutionType" minOccurs="0" maxOccurs="unbounded"/>
26:     <element name="DisplayDevice"type="mpeg7:ControlledTermUseType"  minOccurs="0"
         maxOccurs="unbounded"/>
27:    </sequence>
28:   </extension>
29:  </complexContent>
30: </complexType>
```

Figure 5: Example of MPEG-21 DIA specification of TerminalCapabilities and DisplayCapabilities

MPEG-21 is organized into several parts that can evolve independently. Each part covers one of the aspects related to the **management** and **dissemination** of multimedia information. To date, there are twelve parts which make up the platform MPEG-21, among which the *Digital Item Adaptation (DIA). Authors* in (Vetro, 2005) define sets of tools for describing the **environment** of use and properties of media resources which can affect the diffusion process (terminals, networks, users profile and preferences).

The fragment of MPEG-21 DIA presented in Figure 5 illustrates the specification of the `TerminalCapabilities` (lines 1 to 09 in Figure 5), `InputOutputCapabilities` concept specification (lines 10 to 20 in Figure 5) that embeds together with audio information and `DisplayCapabilities` of the device (lines 21-30 in Figure 5) such as resolution or type of display device, etc. Hence we have a four level description for introducing characteristics of a device seen as a **terminal** giving user access to content.

In the following paragraph we will discuss how the same concept Device or Terminal is modeled and encoded with the metadata initiative of the CAM4Home consortium.

CAM4Home Supplementary Metadata

CAM4Home Supplementary Metadata is a part of the CAM4Home Metadata model proposed by the CAM4Home Consortium that was briefly introduced in the previous section.

CAM Supplementary Metadata schema provides information required to enable **interoperability** of the platform services and supplement the Core metadata. CAM Supplementary Metadata provides the structures for **profiling** of users, communities, devices, network and platform services. For each one of these entities several profiles can be associated in order to support time-dependant (in the morning, in the afternoon) and usage-dependent (at home, at work) characteristics.

A fragment of the schema describing the device profiles and device-related metadata within the CAM4Home Supplementary Metadata schema is shown in Figure 6. A first `rdf:property` construct introduces the `hasC4HDeviceProfile` property (lines 1-4 in Figure 6) that links `C4HDeviceProfile` description to a given `C4HDevice`. Further on, the `hasDeviceCapabilities` property (defined by lines 5-8 in Figure 6) attaches a `DeviceCapabilitiesDescription` to a given C4HDeviceProfile. Other descriptions not included in this fragment such as software or hardware description can be attached to a `C4HDeviceProfile` in order to describe a specific device **configuration**. The `DeviceCapabilitesDescription` (lines 13-15 in Figure 6) is composed of Display Capabilities descriptions (lines 13-23 in Figure 6), Audio Capabilities description and Browser Capabilities (not included in the fragment). Properties such as display Size (lines 16-19 in Figure 6) and type Of Display (lines 20-23 in Figure 6) describe the display capabilities of a device.

From a logical point of view, many similarities can be observed between the MPEG-21 DIA `TerminalCapabilities` and C4HDevice descriptions. `DeviceCapabilitiesDescription` (from CAM4Home) element could be partially mapped onto the `TerminalCapabilities` (from MPEG-21 DIA) element. `DisplayCapabilities` concept is available in both specification, however they encoding is quite different with regard to the labels used for introducing display-related properties (e.g. `displaySize` in CAM4Home vs. `Resolution` in MPEG-21 DIA) and the way the properties are encoded (e.g. `displaySize` as a string vs. `Resolution` as a complex element having a vertical and horizontal resolution components).

However, the MPEG-21 DIA addresses the issue of characterizing a wide variety of **access devices**. The CAM4Home project only considered a small set of devices such as 3G mobiles phones, TV set-top boxes, laptop and desktop computers.

```
01: <rdf:Property rdf:about="&supplementary;hasC4HDeviceProfile" ...>
02:  <rdfs:domain rdf:resource="&supplementary;C4HDevice"/>
03:  <rdfs:range rdf:resource="&supplementary;C4HDeviceProfile"/>
04: </rdf:Property>
05: <rdf:Property rdf:about="&supplementary;hasDeviceCapabilities" ...>
06:  <rdfs:range rdf:resource="&supplementary;C4HDeviceCapabilitiesDescription"/>
07:  <rdfs:domain rdf:resource="&supplementary;C4HDeviceProfile"/>
08: </rdf:Property>
09: <rdf:Property rdf:about="&supplementary;hasDisplayCapabilities" ...>
10:  <rdfs:domain rdf:resource="&supplementary;C4HDeviceCapabilitiesDescription"/>
11:  <rdfs:range rdf:resource="&supplementary;DisplayCapabilities"/>
12: </rdf:Property>
13: <rdfs:Class rdf:about="&supplementary;DisplayCapabilities" …>
14:  <rdfs:subClassOf rdf:resource="&abstract;DeviceMetadata"/>
15: </rdfs:Class>
16: <rdf:Property rdf:about="&supplementary;displaySize" ...>
17:  <rdfs:domain rdf:resource="&supplementary;DisplayCapabilities"/>
18:  <rdfs:range rdf:resource="xsd:string"/>
19: </rdf:Property>
20: <rdf:Property rdf:about="&supplementary;typeOfDisplay" ...>
21:  <rdfs:domain rdf:resource="&supplementary;DisplayCapabilities"/>
22:  <rdfs:range rdf:resource="xsd:string"/>
23: </rdf:Property>
```

Figure 6: Example of CAM4Home Supplementary specification of Device and DisplayCapabilities

A fine grain analysis of MPEG-21 DIA and CAM4Home Supplementary Metadata suggests that the CAM4Home Supplementary Metadata can be partially mapped on a subset of MPEG-21 DIA descriptions. A coherent metadata integration approach should be able to identify similarities like the one we have presented above by going beyond the logical structure imposed by XML Schema or RDF Schema constructs. The computation of partial mappings between parts of the metadata specifications seems also to be very useful when considering specifications having complex structures like the ones exhibited by MPEG-21 DIA.

CC/PP

As part of a framework for **adaptation** and **contextualization** of content, a format; called CC/PP that can describe the possibilities of a user agent (Web browser, modem, etc.) was proposed by W3C. The CC/PP is a language based on profile description. Each CC/PP profile is a description of the possibilities of access device and user preferences that can be used to guide the adaptation of contents.CC/PP is based on RDF, which was developed by the W3C as a language for describing metadata. RDF provides a basic framework for the **scalability** of CC/PP vocabulary, through the usage of XML namespaces.

CC/PP Profile contains attributes and values which are processed by a server to determine the most appropriate form of a media resource relative to this profile. It is structured to enable a client and/or an intermediate server to describe their capabilities by referring: a) a standard profile, accessible to origin server or other sender and b) a smaller set of properties outside the standard profile, but understood by both parties. CC/PP profile is structured in a hierarchy with 2 levels with a number of components, each component having at least one or more attributes.

The main difference with the previous standards is that CC/PP specifications provide only very few predefined concepts (Profile, Component, Property, Structure and Attribute). A profile is composed of one or several components. Components are described either using CC/PP natives' properties such as CC/PP structures and CC/PP attributes or using properties defined by external schemas. The CC/PP standard was conceived as a generic approach for describing any kind of **client profiles**. Constructing the mapping between other standards and CC/PP should be envisioned in a new way as CC/PP is more over a container (having a structure that conforms the profile-component pattern) of descriptions. When mapping CC/PP to other standards we should take into account the RDF Schema included in CC/PP description through specific XML namespaces. For CC/PP-related the mapping one should consider the instances of CC/PP profiles and the mapping will be done among the RDF Schema associated with the XML namespaces. So, here we have to consider a specific schema mapping process driven by instances.

```
01: <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:ccpp="http://www.w3.org/2002/11/08-ccpp-schema#">
02:  <rdf:Description rdf:about="http://www.example.com/profile#MyProfile">
03:   <cpp:component>
04:    <rdf:Description rdf:about="http://www.example.com/profile#TerminalHardware">
05:     <!-- TerminalHardware properties here -->
06:    </rdf:Description>
07:   </ccpp:component>
08:   <ccpp:component>
09:    <rdf:Description rdf:about="http://www.example.com/profile#TerminalSoftware">
10:     <!-- TerminalSoftware properties here -->
11:    </rdf:Description>
12:   </ccpp:component>
13:   <ccpp:component>
14:    <rdf:Description rdf:about="http://www.example.com/profile#TerminalBrowser">
15:     <!-- TerminalBrowser properties here -->
16:    </rdf:Description>
17:   </ccpp:component>
18:  </rdf:Description>
19: </rdf:RDF>
```

Figure 7: CC/PP profile definition using a specific namespace for component properties

For instance, the CC/PP fragment in Figure 7 introduces three components: `TerminalHardware` (lines 3-7), `TerminalSoftware` (lines 8-12) and `TerminalBrowser` (lines 13-17). These components where defined as belonging the `http://www.example.com/profile` namespace. Hence, if one considers mapping the fragment in 7 onto other standards, one should compute a mapping between the RDF Schema corresponding to the `http://www.example.com/profile` namespace and the given standard.

EXISTING METADATA INTEGRATION SOLUTIONS

In the previous chapter we introduced some multimedia standards and we briefly discussed some issues that can arise while trying to compute mapping between existing multimedia standards, usually characterized by highly complex schema introduced numerous concepts and inter- or intra-concepts relations.

In the following, we briefly discuss some solutions available in the current state-of-art that addressed the problem of multimedia metadata **integration**, Afterwards, we discuss the W3C initiative, called Ontology for Media Object (WonSuk, 2009), that opens the way to providing specific mapping solutions closely related to *de facto* multimedia standards as those presented in the previous section. This section shows also some works done in order to **semi-automatically** map between heterogeneous descriptions structured as generic ontologies.

The problem of integrating heterogeneous multimedia metadata interested researcher this last decade. (Garcia, 2005) proposed a framework to integrate three different music ontologies. He used the generated MPEG-7 OWL ontology as an upper-ontology (mediated schema) to integrate other music metadata (MusicBrainz schema, Simac music ontology and a music vocabulary to describe performances). This music metadata are mapped manually to MPEG-7 ontology. (Hunter, 2003) proposed a core top-level ontology for the integration of information from different domains. A core top-level is an extensible ontology that expresses the basic concepts that are common across a variety of domains and media types and that can provide the basis for specialization into domain-specific concepts and vocabularies. It allows the construction of well-defined mappings between several domain-specific knowledge representations (i.e., metadata vocabularies).
Some other frameworks have been proposed to integrate multimedia metadata (Doerr, 2003) (Tsinaraki, 2004) (Troncy, 2003). However, these frameworks are limited and cover usually a small number of standards. The Ontology for Media Object mentioned previously is, at our knowledge, the framework that considers the integration of the largest number of metadata standards. Moreover, all exiting works, including Ontology for Media Object are based on manual mapping which requires experts having deep knowledge about each considered standard and, hence, an important time for the acquiring the knowledge and processing hundreds of metadata definitions. A semi-automatic framework for mapping between different formats of metadata facilitates the integration process; this is the reason that drove us to propose a new framework described later in this book chapter. Our solution is inspired by the existing generic ontology matching approaches. We have studied these approaches in order to select the most adequate approach for multimedia metadata (Amir, 2009).

Ontology for Media Object (W3C)

In order to enhance the integration of heterogeneous metadata, since 2008 a new W3C working group has been working at the definition of a new system called *Ontology and API for Media Object* (WonSuk, 2009). The ontology obtained following the integration process is expected to support cross-community data integration of information related to multimedia content on the Web. The API will provide **read access** and potentially **write access** to media objects, relying on the definitions from the ontology.

The *Ontology for Media Object* addresses the **inter-compatibility** problem by providing a common set of properties to define the basic metadata needed for media objects and the semantic links between their values in different existing vocabularies. It aims at circumventing the current proliferation of video metadata formats by providing full or partial translation and mapping between the existing formats. The ontology is to be accompanied by an API that provides **uniform access** to all elements defined by the ontology, which are selected elements from different formats. Table 1 shows a part of the mapping result between some multimedia metadata standards (METS, LOM2.1, CableLabs 1.1, YouTube and TV-Anytime) and a set of mediated ontology properties (API).

Notwithstanding the efforts of *Ontology for Media Object* working group, we think that the integration could benefit from the existing approaches to find a semi-automatic semantic mapping between a common set of properties which a mediator between user and metadata standards. At the given time, the W3C working group constructs the mapping manually.

Currently the W3C Working Group considers standards such as: Dublin Core, EXIF (EXIF, 2004), ID3 (ID3, 1999), IPTC (IPTC, 2008), Media RSS (Media RSS, 2008), MPEG-7 and XMP (XMP, 2008). The **manual mapping** is more precise then the semi-automatic one, but the first one is hard to realize for large standards and need to be updated manually after each modification or addition of new properties.

In order to deal with this problem, and to link **semi-automatically** proprieties and concepts over various multimedia standards we are working in order to design and implement a multimedia metadata integration system described in the next section.

Standards / API	METS	LOM 2.1	*CableLabs 1.1*	*YouTube*	TV-Anytime
contributor	metsHdr/agent	contribute	Actors Actors_Display Advisories Director	credit@role	CreditsList/CreditsItem + role
language	X	language	Languages	media:content@lang	Language, CaptionLanguage, SignLanguage
copyright	Copyright	X	Provider	copyright	DRMDeclaration / CopyrightNotice
format	X	format	Encrypting_System_Info Encryption Encryption_Algorithm Encryption_Date Encryption_Key_Block Encryption_Time Encryption_Time Audio_Type	content@type	AVAttributes
genre	ImageDescription, INAM	Learningresourcetype	Genre Category	content@medium	Genre

Table 1: A fragment of mapping table (Ontology for Media Object)

Ontology metadata mapping

A **mapping** is a sort of **alignment** between ontology constructs as defined in (Kalfoglou, 2002). That is, given two ontologies, one should be able to map concepts found in the first ontology onto the ones found in the second one. Several mapping algorithms have been proposed during the last few years, each of them has been implemented for a **specific domain**. Several surveys (Kalfoglou, 2003) (Shvaiko, 2005) (Rahm, 2001) about different existing method for ontologies mapping show that their applicability depend on utilization domain.

When doing ontology mapping, some authors focus only on syntax and structure (Doan, 2002) (Doan, 2000) others use specific semantics associated with concepts in the ontology and some of them extract information from instances (Madhavan, 2001) to give more precision during schema mapping operation.

Kalfoglou and Schorlemmer (Kalfoglou, 2002) developed an automatic method for ontology mapping, IF-Map, based on the Barwise-Seligman theory of information flow. Their method built on the proven theoretical ground of Barwise and Seligman's channel theory, provides a systematic and mechanized way for deploying it on a distributed environment to perform ontology mapping among a variety of different ontologies. They consider a mapping between two ontologies as a **logic informorphism** and they apply tools based channel theory to establish the semantic mapping.

Doan et al. (Doan, 2002) developed a system, GLUE, which employs machine learning techniques to find mappings. Given two ontologies, for each concept in one ontology, GLUE finds the most similar concept in the other ontology using probabilistic definitions of several practical similarity measures. It exploits information about the concept instance such as the frequencies of words in the text, instance names, the value formats, or the characteristics of value distributions. GLUE is an interesting approach and is applicable to semi-structured metadata such as those defined by XML Schema that implies a tree structure. However, it does not take into consideration the type of relationship between concepts. All mappings are considered as equivalence relationships.

Structural and syntactic solutions such Cupid (Madhavan, 2001) can be applied with success for semi-structured metadata but it does not offer a good mean to resolve heterogeneity problem that can arise while considering semantic resources.

Nevertheless, some of these approaches are purely semantic and give interesting results in specific domain depending on the complexity of semantic and structure of resources. In the next section we will present the approach that we consider it as a compatible with the complexity of multimedia metadata and it is applicable for our suggested framework.

MULTIMEDIA METADATA INTEGRATION SYSTEM

Firstly, we present the overview of our approach. Secondly, we discuss some issues and solutions regarding the **homogenization** of schema languages used before illustrating how we envision computing **mappings** between various concepts using formal tools such as Description Logic (Baader, 2003).

Overview of the approach

Figure 8 shows the general architecture of multimedia metadata integration system. This system helps users to interpret metadata encoded by external formats (other than the one adopted as the mediated format). The specification of external formats is done by other communities which use different technologies to encode multimedia contents and contexts. The user of the integration system is not able de understand the meaning of this external metadata if it does not have a pre-knowledge about it. The framework described below is a solution to deal with this problem. It allows user to access existing metadata requiring knowledge only about the mediated schema.

With a view to making a good understanding of the role of this framework, we present in Figure 9 an integrated mapping which we are experimenting within the CAM4HOME ITEA2 project. Figure 9 shows a fragment of CAM4Home metadata model (that plays the role of the mediated schema) mapped to two other metadata standard (DIG35 and MPEG7) fragments. The role of mapping consists of facilitating the access to information of users regardless of the encoding used by external standards. For instance, if a CAM4HOME user asks the following query *get c4h:CreatorReference of MediaObject#1* and if this information is not natively available in CAM4HOME format, the integration system exploits existing mapping in order to rewrite queries and to retrieve metadata from MPEG-7. Hence, the user does not have to look by itself for equivalent queries tailored for MPEG-7 constructs. The integration system must find the equivalence relationship between concepts in the mediated schema and external metadata standards considered for integration.

The framework that we present here enables the construction of single *virtual integrated metadata* sources that regroups information required by services or users dealing with metadata. With this framework user does need any **pre-knowledge** about all standards to be able to interpret their meaning. End-user requests metadata from virtual integrated metadata source which is the interface between end user and all available metadata resources.

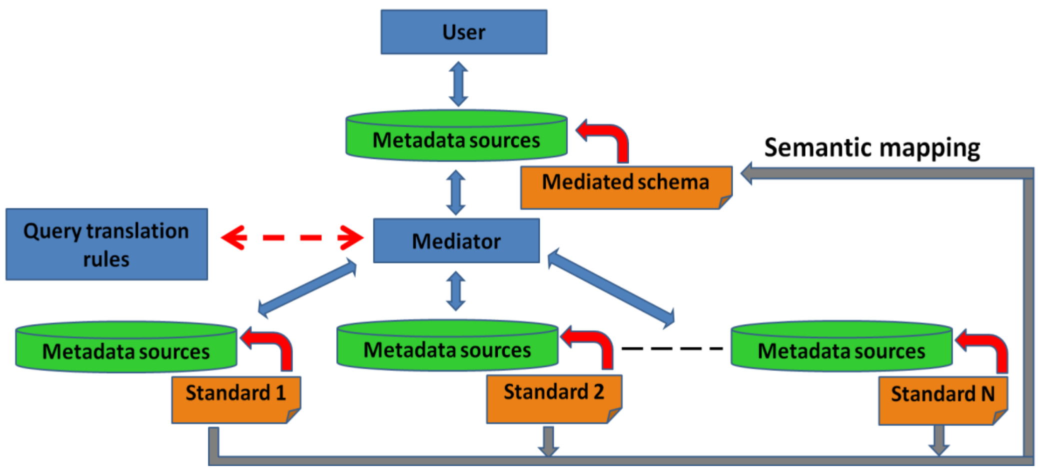

Figure 8: Metadata multimedia integration system architecture

Virtual integrated metadata interpreter sends the query to metadata source mediator which will use translation rules to transform the query so it matches the available metadata sources containing the metadata corresponding to the query. **Translation rules** are defined according to the semantic mapping done between the mediated schema and different schemas describing metadata standards. Finally, metadata sources mediator translates query results via rules back language used by virtual integrated metadata source. The solution we propose follows a **Local As View** approach. As each source is integrated independently of the other, the system is flexible and does not require recomputing all mappings when a new standard is integrated. In contrast to Local As View, adding, deleting, updating one standard integrated in the system requires a total change of mapping because the properties of the mediated schema elements were extracted according to the global mapping previously done.

To implement the framework described here above, four main challenges have to be taken into consideration:

- Syntactic heterogeneity due to the diversity of existing description language. This kind of heterogeneity is resolved by the homogenization of all standards to a unique format able to express all semantics available in each standard. The conversion must capture all semantics in original schemas.
- Finding the **semantic alignment** between mediated schema and metadata multimedia standards. This step is done after the homogenization.

- **Query translation** rules definition according to the complex alignment done between mediated schema and metadata standards.
- Identification of all information encoded by different metadata format and describing a given multimedia object.

In the following we will only talk about two first challenges which are described in the next sections.

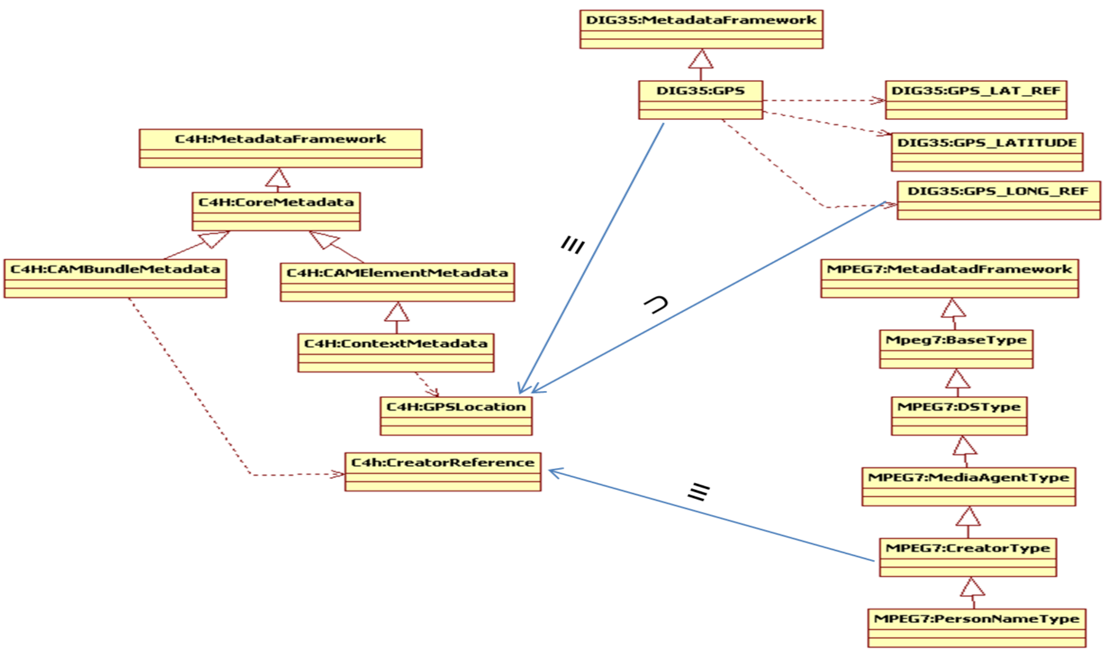

Figure 9: Example of CAM4Home metadata fragment mapped to DIG35 and MPEG7

Multimedia metadata homogenization

Homogenization is a necessary stage to implement our suggested framework. (Kalfoglou, 2003) consider that a common representation language is often necessary in order to create a semantic mapping between heterogeneous resources. Therefore, the challenge at this stage of the work is to specify a set of translation rules to convert between different encoding schema languages (XML Schema, RDF Schema, OWL).This step must be done with a minimum loss of structural and semantic information.

Several research works have been done during the last few years in order to resolve heterogeneity problem. Some of them focus on the syntactic level of resources to resolve semantic ambiguity (Chaudhri, 1998). Others have showed that using different syntax or structure formalization is source of various errors when transforming a representation into another (Bowers, 2000).

Since the appearance of XML, several multimedia standards have adopted XML to describe their metadata (MPEG-7, TV-Anytime, MPEG-21, IEEELOM (LOM, 2002), SCORM (SCORM, 2004), etc.). However, XML Schema provides support for explicit structural, cardinality and datatyping constraints, but offers little support for the semantic knowledge necessary to enable flexible mapping between metadata domains. RDF schema has appeared to cope with XML Schema semantic limitations (Antoniou, 2008). It provides support for rich semantic descriptions but provide limited support for the specification of local usage constraints. Finally, OWL appeared as extension of RDF Schema, it allows maximum semantic expressiveness. These three description languages mentioned previously cover a large amount of multimedia metadata standards (Hausenblas, 2007). For this reason we will restrict our study to them within the integration framework we plan to implement.

In order to implement our suggested framework we chose to use **OWL-DL** (McGuinness, 2004) which is sub-language of OWL as a common language. OWL-DL has a sufficient expressiveness for all types of multimedia metadata. Besides, existing mapping approaches based on logic description methods cannot be applicable to OWL-Full ontology which is sometimes **undecidable.**

In order to convert XML Schema to OWL several approaches have been proposed during the last few years to alleviate the lack of formal semantics in XML documents by converting this latest to OWL ontology. Authors in (Matthias, 2004) made a conversion only between XML Schema and OWL without considering XML instance metadata, (Battle, 2004) has proposed an approach to convert XML document to OWL ontology using XML Schema. If this latest is not available they generate it from XML data instance. Bohring et al. (Bohring, 2005) proposed a framework for converting a single XML instance document to an OWL model. Authors in (Garcia, 2005) used an approach called XSD2OWL to transform an XML Schema into a set of OWL ontologies.

Authors of XSD2OWL introduced an adjustment to first versions of XSD2OWL in order to alleviate some lacks relative to some implicit semantic after the conversion of XML Schema into OWL ontology. The XSD2OWL produces three ontologies: one describing the domain concepts of the XML Schema, a second one describing the user-defined value types used by the XML Schema and finally a third one, keeping track of the structure information (position of properties in complex elements descriptions). The latest ontology is necessary as all syntactic information (like the order of properties) is lost when passing to OWL encodings. The XSD2OWL mapping has been successfully applied to the MPEG-7 XML Schemas producing a MPEG-7 Ontology. The result got from this conversion is interesting from our point of view. However, the ontology produced is **OWL-Full** which is undecidable.

- We think that the XS2OWL algorithm described in (Chrisa, 2007) is more suitable for the homogenization step as it responds to the following requirements: It outputs OWL-DL ontologies, which gives us the opportunity to use inference based on description logic to find semantic mapping. It conserves all structural information which can be used for the mapping. Besides, the XS2OWL model and its implementation has been successfully applied to several *de-facto* multimedia standards (IEEE LOM, MPEG-21, MPEG-7 and others) and its results showed that the captured semantics is similar with the manually created one.

Concerning the conversion from RDF Schema to OWL, we know that OWL is an extension of RDF Schema. OWL inherits and extends RDF Schema characteristics. Therefore, any document which is valid according to RDF Schema is valid as an OWL-Full ontology. However, this condition is not always satisfied for **OWL-DL** (Patel-Schneider, 2004). For this reason we must distinguish between OWL-Full and **OWL-DL**. Besides, the ontology result must be syntactically compatible with the result of XS2OWL algorithm mentioned before.

Before presenting in detail the method for computing mappings between concepts in various multimedia metadata standards, we recall the main conclusion of the homogenization step. A common schema description language is needed for computing coherent mappings. OWL-DL, notably due to its decidability feature, was selected for playing the role of common schema description language within the integration framework. XS2OWL tool will be used to convert XML Schema to OWL-DL. A tool for converting RDF Schema to OWL-DL will be developed within the framework.

Semantic method for multimedia metadata mapping

The mapping process is a two step process. The first one consists in associating explicit meaning to concepts issued from various standards created by communities having specific backgrounds. The explicit meanings are defined with regard to a neutral external semantic resource. In the following, we use WordNet to do explicitation. However other resources such as DBpedia (Auer, 2008) or YAGO (Suchanek, 2008) can be considered as external resources in order to increase the neutrality of the explicitation process.

Multimedia metadata explicitation

Because of the richness of natural language, different multimedia ontologies maybe use different terminologies to denote the same concept. The **explicitation** of concept meanings is a necessary step to offer adequate solution for facing the linguistic heterogeneity (Bouquet, 2003). Explicitation is done by giving all possible interpretations of each concept via an external resource which contain a set of synonyms for each concept.

We use **WordNet** (Fellbaum, 1998) as an external resource for explicitation. WordNet is an on-line lexical reference system developed at Princeton University. It consists of synonym sets called **synsets**, and each synset represents a single distinct concept. For instance, the right-hand side of Figure 10 is a part of DIG35 standard ontology created manually. The explicitation of the embedded `Organization` concept is represented by seven synsets returned by WordNet: (`Organization#1`$\cup$ …$\cup$`Organization#7`).

Each synset corresponds to a specific meaning of the Organization concept. For instance, the `Organization#1`synset contains words (sister terms, hyponyms, etc) related to the notion of *a group of people who work together*. For instance, the `Organization#1`synset contains words (sister terms, hyponyms, etc) related to the notion of *an organized structure for arranging or classifying*.

Concepts explicitation is a necessary step because it gives to each concept all possible interpretations for it. However, concepts explicitation can also result in a wide enlargement of the field of meaning of concepts. The selection of pertinent classes of words considered (sister terms, hyponyms, etc.) for each synset is an open issue. Sense filtering is a needful operation which allows the selection of the real meaning of concept. In order to select the real meaning of a given concept C on the hierarchy H, sense **filtering** operation uses, n neighbors of C ($C_{1,}$…$_{,}$ C_1) with regard to the hierarchy H. Other resources can be also helpful for sense filtering operation (e.g. extraction of **semantic** from documentation, user comment, and social tags for instance).

Mapping discovering using DL

Notwithstanding the existing approaches for ontology mapping, **DL** based techniques are more appropriate for that, since they rely on the explicit and formal semantics represented by ontologies. When used for comparing ontologies, they ensure that the original semantics of resources is preserved. Besides, DL provides an explicit and a formal interpretation of concepts (Amir, 2009).

A valid semantic mapping between two concepts C_1 and C_2 belonging, respectively, to ontologies O1 and O_2 is a logical relationship ($\subseteq$ - left subsumption, $\supseteq$-right subsumption, $\equiv$ - equivalence and $\perp$- disjointness -). This logical relationship can be computed starting from the explicitation results obtained from external resource and depending on the nature of the relationship it must satisfy one one of the operations shown in Table 2

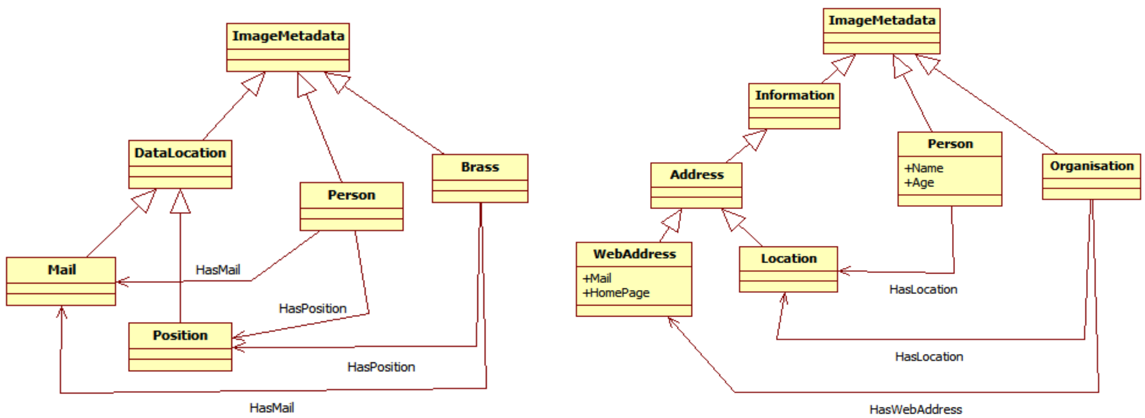

Figure10: (a) Mediated Schema (on the left-side), (b) fragment ofDIG35 (DIG35, 2002) standard ontology

DL Axioms		Set relations
$C_1 \subseteq C_2$	$\Leftrightarrow$	$C_1 \cap \neg C_2$ is unsatisfiable.
$C_1 \supseteq C_2$	$\Leftrightarrow$	$C_2 \cap \neg C_1$ is unsatisfiable.
$C_1 \equiv C_2$	$\Leftrightarrow$	$C_1 \cap \neg C_2$ is unsatisfiable and $C_2 \cap \neg C_1$ is unsatisfiable.
$C_1 \perp C_2$	$\Leftrightarrow$	$C_1 \cap C_2$ is unsatisfiable.

Table 2: DL axioms and corresponding set relations

In order to do inference between concepts, the relationships between concepts in mediated schema and multimedia metadata standards is computed by using the explicitation results provided by external resource. The results are transformed to the axioms as the premises of inference. To obtain this kind of axioms we analyze the relationship within the synsets provided by WordNet. WordNet organizes terms based on the semantic relations of them. So these relations are the origin of the generation of axioms. The table here bellows shows the transformation of relations of terms from WordNet to corresponding **subsumption** axioms.

Relations in WordNet	Subsumption axioms
term_1*meronym* term_2	$\text{term}_1 \subseteq \text{term}_2$
term_1*holonym* term_2	$\text{term}_2 \subseteq \text{term}_1$
term_1*hyponym* term_2	$\text{term}_1 \subseteq \text{term}_2$
term_1*hypernym* term_2	$\text{term}_2 \subseteq \text{term}_1$

Table 3:WordNet relations vs. subsumption axioms

Once the axiom is obtained by analyzing the retrieved WordNet synsets, we use **Tableau algorithm** that can be easily applied to DL as illustrated in (Baader, 2001). The Tableau algorithm establishes the relationships between concepts by observing one of the four conditions of satisfaction mentioned above, which corresponds to the axiom to be validated.

Tableau algorithm belongs to a class of algorithms that use subsumption expansion in order to prove the satisfiability of a relation (e.g. C1 $\subseteq$ C2). The Tableau algorithm expands each ontology concept by considering the subsumption axioms deduced from WordNet relations. Then it tries to prove the satisfiability of the considered relation by verifying the coherence of the set relations as indicated in Table 2. The relation itself can be seen as a *complex* concept (C) that is defined by the associated set relations.

The algorithm exhaustively applies **tableau rules** which decompose the syntactic structure of this *complex* concept in order to construct a so-called ***completion tree***. Each node (x) of a *completion* is labeled with a concept set $L(x) \subseteq sub(C)$ where sub(C) is a set of sub-concepts composing the concept C. Each edge $<x, y>$ is labeled by $L(<x, y>) = R$ for some role R occurring in $sub(C)$, where the $sub(C)$ is the set of sub-concepts of *C.*

The expansion rules are illustrated in Figure 11. The algorithm terminates when the graph is complete (no further inference is possible) and returns “mapping satisfiable”, or when contradiction have been revealed and return in this case “mapping unsatisfiable”; see (Baader, 2001) for details. This contradiction is called a clash and means that for some concepts C, $\{C_1, \neg C_2\} \subseteq L(x)$. This means that the concept C_1 and the negation of concept C_2co-occurs and hence the subsumption relation is **unsatisfiable.**

$\cap$- rule: if (1) $(C_1 \cap C_2) \in L(x)$
(2) $\{C_1, C_2\} \not\subset L(x)$
then $L(x) \rightarrow L(x) \cup \{C_1, C_2\}$
$\cup$- rule: if (1) $(C_1 \cup C_2) \in L(x)$
(2) $\{C_1, C_2\} \cap L(x) = \perp$
then $L(x) \rightarrow L(x) \cup \{C\}$ for some $C \in \{C_1, C_2\}$
$\exists$ - rule: if (1) $\exists S.C \in L(x)$
(2) x has no S-neighbor y with $C \in L(y)$
then create a new node y with $L(<x, y>) = S$
and $L(y) = \{C\}$
$\forall$ - rule: if (1) $\forall S.C \in L(x)$
(2) there is an S-neighbor y of x with $C \notin L(y)$,
then $L(y) \rightarrow L(y) \cup \{C\}$

Figure 11: Tableau algorithm expansion rules

In the following we illustrate how Tableau algorithm can be used to validate a mapping between two concepts belonging to two different schemas. We use a fragment of the DIG35 standard and the mediated schema proposed in Figure 10.

Firstly, we start to extract the **logic formula** for each concept in two schemas (Bouquet, 2003), this step is done after concepts explicitation step:

$$\begin{aligned} C_1 = C\ &(\text{Organization}) = \\ &(\text{Organization\#1} \cup \ldots \cup \text{Organization\#8}) \cap \\ &\forall \text{HasLocation.}(\text{Location\#1} \cup \ldots \cup \text{Location\#4}) \cap \\ &\forall \text{HasWebAddress.}(\,(\text{Web\#1} \cup \ldots \cup \text{Web\#7}) \cap \\ &(\text{Address\#1} \cup \ldots \cup \text{Address\#8})\,) \cap \\ &(\text{Image\#1} \cup \ldots \cup \cup \text{Image\#9}). \end{aligned} \tag{1}$$

$$C_2 = C\,(Brass) = \\ (Brass\#1 \cup \ldots. \cup Brass\#7) \cap \\ \forall HasPosition.(Position\#1 \cup \ldots. \cup Position\#16) \cap \ ) \qquad (2) \\ \forall HasMail.(Mail\#1 \cup \ldots.\cup Mail\#5) \cap (Image\#1 \cup \ldots.\cup Image\#9).$$

Secondly, we apply the sense filtering operations on available synsets. The aim of this filtering is to identify which one of the synsets is closer to the meaning of the concept in the ontology where he was initially defined. Hence, the sense filtering operation is taking into account the context of the source ontology by examining the neighboring concepts. For instance, if we consider the DIG-35 fragment considered, synsets associated with the Organization concept but containing terms which are not at all related with at least one of neighboring concepts: location, address or image will be left out. Same filtering can be applied to the Brass concepts in the mediated schema.

In the current example, the Organization#3 and Brass#3 synsets are the result of filtering operations. Both synsets refer to "the persons (or committees or departments etc.) who make up a body for the purpose of administering something". As they are belonging to the same synset then they can be represented by the same label, which means that there is an equivalence relationship between the two atomic concepts according to WordNet. However, this is not always true when considering complex structures because the meaning of a given atomic concept depends also on its hierarchy in the schema. Therefore, this relationship is considered as an axiom and has to be validated. The axiom stating that C_1 and C_2 are equivalent is explicitated here bellow:

$$C_1 \equiv C_2 \Leftrightarrow C_1 \cap \neg C_2 \text{ is unsatisfiable and } C_2 \cap \neg C_1 \text{ is unsatisfiable.}$$

To compute the validity of the **axiom** we introduce new concepts and relations starting from (1) and (2) equations. We also restrain the number of synsets corresponding the filtering applied to the mapped concepts (see C_{11} and C_{21} concept sets). The list of concept and relations are presented here bellow:

C_{11}= (Organization#3)
C_{12} = (Location#1 $\cup$ …. $\cup$ Location#4)
C_{13}= (Web#1 $\cup$ …. $\cup$ Web#7)
C_{14} = (Address#1 $\cup$ …. $\cup$ Adress#8)
C_{15} = (Image#1 $\cup$ …. $\cup$Image#9)
C_{21}= (Brass#3)
C_{22}= (Posotion#1 $\cup$ …. $\cup$ Position#16)
C_{23} = (Mail#1 $\cup$….$\cup$Mail#5)

R_1= HasLocation, R_2= HasWebAddress, R3= HasPosition, R_4= HasMail

Since Organization#3 and Brass#3 belong to the same synset according to WordNet, they can be represented by the same label $C_{11.}$ However, this is not necessary mean that the mapping is validated. We have to validate or invalidate the axiom by testing the satisfaction of the corresponding DL-axiom (from Table 2). Hence (1) and (2) equations become:

$$C_1 = C_{11} \cap \forall R_1.\, C_{12} \cap \forall R_2.\, (C_{13} \cap C_{14}) \cap C_{15}$$
$$C_2 = C_{11} \cap \forall R_3.\, C_{22} \cap \forall R_4.\, C_{23} \cap C_{15.}$$

We apply the De Morgan's law to transform the concepts to negation normal form

$\neg C_2 = \neg C_{11} \cup \exists R_3. \neg C_2 \cup \exists R_4. \neg C_{23} \cup \neg C_{15}$

The tableau algorithm initializes a tree T which contains initially only one node x0 labeled with the following concept set : L(x0) = {C1 ∩ ¬C2}, called the root node. The Tableau rules showed in Figure 11 in order to extend L(x0).If tableau algorithm returns unsatisfiable for C1 ∩¬C2 as well as for the second tree computed starting from the concept set C2 ∩¬C1, then the axiom C1≡ C2 can be considered as a valid mapping.

In order for the Tableau algorithm to expand rules in a coherent manner we need to consider equally the explicitation of relations and related concepts involved in the equations. Hence, this solution exhibits a recursive behavior. Currently, we are working on studying the depth of the recursion tree to consider. At the leaf levels, we only consider the mapping provided by WordNet synset comparisons without verifying the equivalence (or partial subsumption) by using Tableau algorithm. For the current experience, we have applied to all relations and related concepts the mapping obtained following the WordNet synset comparison. The results obtained from WordNet show that R_1 maps on R_3 and R_2 (partially) maps on R_3. Under this assumption, the Tableau algorithm validates the mapping of C_1 and C_2.

FUTURE RESEARCH DIRECTIONS

We described in the previous sections some existing tools which can be used to integrate heterogeneous multimedia metadata. We have shown DL characteristics and its advantages to discover a mapping between mediated multimedia metadata ontologies. Since DL is **monotonic**, we also consider getting other semantic resources involved especially in the explicitation process. They allow enriching schema semantic, sense filtering, enhancing semantic mapping between concepts. An external resource does not mean necessary thesaurus such WordNet but any kind of information that can be extracted from other corpus (e.g., user comment, social tags, documentation provided by communities that have created multimedia metadata standards, etc). Since these corpuses contain fuzzy information, we consider testing other methods based **fuzzy logic** or **probabilistic approaches** for discovering semantic mapping.

CONCLUSION

This paper has given an overview of multimedia metadata heterogeneity problems and discussed issues related to an integration system based on DL. We have shown how to homogenize different metadata format in order resolve syntaxic heterogeneity of existing schema description languages. Homogenization is achieved by translating all schema specifications (done mainly in XML Schema or RDF Schema) to OWL-DL. Homogenization methods that preserve the structural information of originals schemas and that capture source schema implicit semantics were presented. We have migrated all schema description to OWL-DL as this ontology language fits well with metadata semantics representation and it supports description logic reasoning. DL is a powerful logical framework issued from the domain of knowledge representation which exhibits high expressiveness constructs. In order to facilitate the mapping between concepts, the precise meaning of concepts within the given context must be first explicitated. The explicitation is done via external resources in order to give all possible interpretation of concepts in schemas. As we have shown before, explicitation is a necessary operation; it must be done in order to select the real meaning of each concept. Since all these steps are done, the mapping can be considered as a DL axiom which must satisfy some conditions. Conditions validation can be tested using Tableau algorithm and its rule-based extension mechanism.

ACKNOWLEDGEMENTS

This work was kindly supported by the ITEA2 CAM4HOME project.

REFERENCES

Amir, S. (2009). Un système d'intégration de métadonnées dédiées au multimédia. *In Informatique des Organisations et Systèmes d'Information et de Décision* (PP. 490-492). Toulouse, France.

Amir, S., Bilasco, I. M., & Djeraba, C. (2009). A Semantic Approach to Metadata Management in Sensor Systems. *COGnitive systems with Interactive Sensors*, (PP. 112-119). Paris, France.

Antoniou, G., & Harmelen, F. V. (2008). *A Semantic Web Primer*. Second edition. Cambridge, MA: The MIT Press.

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., & Ives, Z. (2008). DBpedia: A Nucleus for a Web of Open Data. *In International Semantic Web Conference:* Vol. 4825. (pp. 722-735).

Baader, F., & Sattler, U. (2001). An Overview of Tableau Algorithms for Description Logics. In Springer Netherlands, *Studia Logica*, (pp. 5-40). Springer.

Baader, F., Calvanese, D., McGuinness, D. L., Nardi, D., & Patel-Schneider, P. (2003). *The Description Logic Handbook, Theory, Implementation, and Applications*. Cambridge University Press.

Battle, S. (2004). Round-tripping between XML and RDF. In the 3rd *International Semantic Web Conference.*

Bilasco, I. M., Gensel, J., Villanova-Oliver, M., & Martin, H. (2005). On indexing of 3D scenes usin1g MPEG-7. *ACM Multimedia*, (pp. 471-474).

Bilasco, I. M., Amir, S., Blandin, P., Djeraba, C., Laitakari, J., Martinet, J., Gracia, E. M., Pakkala, D., Rautainen, M., Ylianttila, M., & Zhou, J. (2010). Semantics for intelligent delivery of multimedia content. *SAC ACM*.

Bohring, H. & Auer, S. (2005). Mapping XML to OWL Ontologies. *LeipzigerInformatik-Tage*, volume 72 of LNI, (pp. 147-156), GI Publisher.

Bouquet, P., Serafini, L., & Zanobini, S. (2003). Semantic Coordination: A New Approach and an Application. *International Semantic Web Conference*. LNCS 2870. (pp. 130-145).

Bowers, S., & Delcambre, L. (2000). Representing and transforming model-based information. *First Workshop on the Semantic Web at the Fourth European Conference on Digital Libraries*. (pp. 18-20). Lisbon, Portugal.

Brickley, D. (2004). RDF Vocabulary Description Language 1.0: RDF Schema, Retrieved February 10, 2004 from http://www.w3.org/TR/rdf-schema/

Chang, S. F., Sikora, T., & Puri, A. (2001). Overview of the MPEG-7 standard. *IEEE Transactions on Circuits and Systems for Video Technology:* vol. 11. (pp. 688–695).

Chaudhri, V. K., Farquhar, A., Fikes, R., Karp, P. D., & Rice, J. P. (1998). OKBC A programmatic foundation for knowledge base interoperability. *In Proceedings of the 15th National Conference on Artificial Intelligence and of the 10th Conference on Innovative Applications on Artificial Intelligence*, AAAI Press, (pp. 600-607).

Chrisa, T., & Stavros, C. (2007). Interoperability of XML Schema Applications with OWL Domain Knowledge and Semantic Web Tools. *In On the Move to Meaningful Internet Systems, OTM 2007*, Part I, LNCS 4803, (pp. 850–869).

DIG35: Metadata Standard for Digital Images (2002). Retrieved June 2002, from http://xml.coverpages.org/dig35.html

Doan, A. H., Domingos, P., & Levy, A. (2000). Learning source descriptions for data integration. *In the Workshop at the Conference of the American Association for Artificial Intelligence*, (pp. 81–92).

Doan, A., Madhavan, J., Domingos, P., & Halevy, A. (2002). Learning to map between ontologies on the semantic web. *In the 11th International WorldWide Web Conference*, (pp. 662-673), Hawaii, USA.

Doerr, M., Hunter, J., & Lagoze, C. (2003). Towards a Core Ontology for Information Integration. *Journal of Digital Information*, Vol. 4 (1).

Exif vocabulary workspace - *RDF Schema* (2004). Retrieved December 2003, from http://www.w3.org/2003/12/exif/

Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*. The MIT Press, Cambridge.

Garcia, R., & Celma, O. (2005). Semantic integration and retrieval of multimedia meta-data. *In the 5th International Workshop on Knowledge Markup and Semantic Annotation*. (pp. 69-80).

Hausenblas, M. (2007). Multimedia Vocabularies on the Semantic Web. Retrieved July 2007, from http://www.w3.org/2005/Incubator/mmsem/XGR-vocabularies-20070724/

Hunter, J. (2003). Enhancing the Semantic Interoperability of Multimedia through a Core Ontology. *IEEE Trans. on Circuits and Systems for Video Technology*: Vol.13. (pp. 49-58).

ID3 tag version 2.4.0 February (1999). Retrieved September 2007, from http://www.id3.org/Developer_Information.

IPTC Standard Photo Metadata (2008). IPTC Core Specification Version 1.1, IPTC Extension Specification Version 1.0, Document Revision 2, June 2008. Retrieved June 2008, from http://www.iptc.org/std/photometadata/2008/specification/IPTC-PhotoMetadata-2008.pdf

ITEA2-CAM4Home - Collaborative Aggregated Multimedia for Digital Home (2008). Retrieved May 2008, from http://www.cam4home-itea.org/

Kalfoglou, Y., & Schorlemmer, M. (2002). Information-Flow-Based Ontology Mapping. Irvine/CA, *In the 1st International Conference on Ontologies, Databases and Application of Semantics*, (pp. 98-127).

Kalfoglou, Y., & Schorlemmer, M. (2003). Ontology mapping: the state of the art. *The Knowledge Engineering Review*, 18 (1), (pp. 1-31).

Klyne, G., Reynolds, F., Woodrow, C., Ohto, H., Hjelm, J., Butler, M. H., & Tran, L. (2004). *Composite Capability/Preference Profiles (CC/PP): Structure and Vocabularies 1.0.* W3C Recommendation. Retrieved January 2004, from http://www.w3.org/TR/CCPP-struct-vocab/

WG12: Learning Object Metadata (2002). *Learning Object Metadata Standard*. Retrieved December 2002, from http://ltsc.ieee.org/wg12/

Madhavan, J., Bernstein, P. A. , & Rahm, E. (2001). Generic Schema Matching with Cupid. *In the 27th International Conference on Very Large Data Bases*, (pp. 49-58).

Matthias, F., Christian, Z., & David, T. (2004). Lifting XML Schema to OWL. *In the 4th International Conference of Web Engineering*, (pp. 354-358).

McGuinness, D. L., & Harmelen, F. V. (2004). *OWL Web Ontology Language Overview*. Retrieved February 2004, from http://www.w3.org/TR/owl-features.

Yahoo! Media RSS Module - RSS 2.0 Module *(2008).* Retrieved January 2004, from http://video.search.yahoo.com/mrss.

WonSuk, L., Tobias. B., Felix, S., & Véronique, M. (2009). *Use Cases and Requirements for Ontology and API for Media Object 1.0.* W3C Working Group.

Patel-Schneider, P. F., Hayes, P., & Horrocks, I. (2004). OWL Web Ontology Language Semantic and Abstract Syntax. W3C recommendation. Retrieved February 2004, from http://www.w3.org/TR/owl-semantics/

Pereira, F. (2001). The MPEG-21 standard: Why an open multimedia framework?. Springer, *Interactive Distributed Multimedia Systems*, (pp. 219-220).

Rahm, E., & Bernstein, P. A. (2001). A *survey of approaches to automatic schema matching*. *VLDB Journal* 10 (4), (pp. 334–350).

Advanced Distributed Learning (ADL) Technical Team (2004). *Sharable Content Object Reference Model (SCORM)*. Retrieved April 2004, from http://www.adlnet.org/Technologies/scorm/default.aspx

Shvaiko, P., & Euzenat, J. (2005). A Survey of Schema-based Matching Approaches. *Journal on Data Semantics* 5(1), (pp. 146-171).

Suchanek, F., Kasneci, G., & Weikum, G. (2008). Yago - A Large Ontology from Wikipedia and WordNet. *Elsevier Journal of Web Semantics*, 6(3), (pp. 203-217).

Thompson, H. S., Mendelsohn, N., Beech, D., & Maloney, M. (2009). *W3C XML Schema Definition Language (XSD) 1.1 Part 1: Structure*. W3C Candidate Recommendation April 2009. Retrieved April 2009, from http://www.w3.org/TR/xmlschema11-1/

Troncy, R. (2003). Integrating Structure and Semantics into Audio-visual Documents. *In the 2nd International Semantic Web Conference*, (pp. 566-581).

Tsinaraki, C., Polydoros, P., & Christodoulakis S. (2004). Integration of OWL ontologies in MPEG-7 and TVAnytime compliant Semantic Indexing. In *the 16th International Conference on Advanced Information Systems Engineering*, (pp. 299 - 325).

TV-Anytime Forum, Metadata Working Group (2003). *Metadata Specification Version 1.3*. Retrieved March 2003, from ftp://tva:tva@ftp.bbc.co.uk/Specifications/COR3_SP003v13.zip

Vetro, A., & Timmerer, C. (2005). Digital item adaptation overview of standardization and research activities. *IEEE Transactions on Multimedia*, 7 (3), (pp. 418-426).

Weibel, S., Kunze, J., Lagoze, C., & Wolf, M. (1998). Dublin Core Metadata for Resource Discovery. Retrieved April 1998, from http://www.ietf.org/rfc/rfc2413.txt

XMP Specification Part 2 (2008). Retrieved January 2008, from http://www.adobe.com/devnet/xmp/pdfs/XMPSpecificationPart2.pdf