Lifting a Metadata Model to the Semantic Multimedia World

GAËTAN MARTENS
Multimedia Lab – ELIS – IBBT
Ghent University, BELGIUM
Gaetan.martens@ugent.be

Ruben Verborgh
Multimedia Lab – ELIS – IBBT
Ghent University, BELGIUM
Ruben.verborgh@ugent.be

CHRIS POPPE
Multimedia Lab – ELIS – IBBT
Ghent University, BELGIUM
Chris.poppe@ugent.be

RIK VAN DE WALLE
Multimedia Lab – ELIS – IBBT
Ghent University, BELGIUM
Rik.vandewalle@ugent.be

Abstract

This paper describes best-practices in lifting an image metadata standard to the Semantic Web. We provide guidelines on how an XML-based metadata format can be converted into an OWL ontology. Additionally, we discuss how this ontology can be mapped upon the W3C's Media Ontology. This ontology is a standardization effort of the W3C to provide a core vocabulary for multimedia annotations. The approach presented here can be applied to other XML-based metadata standards.

Keywords: *Multimedia, Metadata Annotation, Semantic Web Technologies*

1 Introduction

In the last decade, digital imaging has experienced a worldwide revolution of growth in both the number of users and the range of applications that have replaced traditional film photography. The amount of digital image content produced on a daily basis is still increasing drastically. As from the very beginning of photography, those who took pictures tried to capture as much information as possible about the photograph and in today's digital age, the need for appending metadata is even bigger. The growth in digital technology has created the desire among users to manage and exchange their images in a variety of ways, including storage, e-mail exchange, World Wide Web postings and photo frames. However, multimedia metadata has, despite its great value, yet to find its way into standard use [1]. Different image metadata standards currently exist, but most of them are, from a consumer perspective, too simple to satisfy the needs or too complex for actual use.

The DIG35 standard is a metadata standard to describe, from an end-user perspective, a set of public metadata for digital still images [2]. As this standard is modeled in XML, the semantic interpretation of the XML-tags can be subjective which interferes with exchanging and managing photo collections as well as reasoning about the annotations. An approach to overcome those difficulties is the use of Semantic Web technologies as they permit reasoning about and relating custom defined concepts. Due to the increasing availability of web-based data sources, the need for integrating heterogeneous data and machine understanding of metadata has grown in importance.

In this paper, we describe the modeling process of an OWL DL [3] ontology for the DIG35 standard for digital images. The OWL DL language is chosen for the following reasons:

- it supports inference,
- it is flexible enough and fairly simple,
- it is a W3C recommendation.

Additionally, we provide means to convert instances of the (XML-based) metadata format to the DIG35 ontology. Lastly, we show how this metadata ontology can be mapped with the standardization efforts in the W3C. The Media Annotations WG is creating a multimedia ontology (called the MA ontology) that can be used as a pivot for multimedia related ontologies [4]. The goal is to improve the interoperability between media metadata schemas. The proposed approach is to provide an interlingua ontology and an API designed to facilitate cross-community data integration of information related to media resources in the web, such as video, audio, and images. This will provide a uniform vocabulary for multimedia applications originating from different communities (e.g., cultural heritage, news aggregation, and broadcasting). We will show how the DIG35 ontology can be linked with the MA ontology. This representation will consequently automatically be supported by the API that is under development [5].

Section 2 gives an overview of existing image metadata standards and the DIG35 standard is compared with the other image metadata standards. Section 3 elaborates on the difficulties that come up when converting the XML Schema into an OWL DL ontology and provides some solutions. In Section 4 the obtained ontology is presented. Section 5 will deal with the conversion of XML instances to OWL instances. Next, Section 6 holds a description of the Media Annotation Ontology. We will show that there exist some overlap with this ontology and DIG35 and discuss some of the relationships in more detail in Section 7. Section 8 evaluates the mappings. Finally, a use case and some conclusions appear in Section 9 and Section 10, respectively.

2 Overview of Image Standards

2.1 EXIF

Probably the best-known image metadata standard is the Exchangeable Image File Format (EXIF [6]) as it is the specification used by (almost) any digital camera nowadays. The metadata tags provided by the EXIF standard cover metadata related to the capturing process of the image. Recently, there have been efforts to represent the EXIF metadata tags in an RDF Schema [7] ontology [8, 9]. As these are mainly technical metadata, other standards are required to describe supplementary categories of metadata.

2.2 MPEG-7

MPEG-7, developed by the Motion Pictures Expert Group (MPEG), offers a comprehensive set of audiovisual description tools [10]. The metadata elements, their structure and relationships are defined by the standard in the form of Descriptors and Description Schemes. The latter specify the structure and semantics of the relationships while the former define the syntax and the semantics of each metadata element. MPEG-7 provides rich and general purpose multimedia content description capabilities, including both low-level features and high-level semantic description constructs. However, the lack of formal semantics in MPEG-7 makes the interpretation of high-level descriptions difficult to cope with. Consequently, low-level features are common, as they can be easily extracted from the content, but there is a deficiency of high-level descriptions.

For MPEG-7, there is no commonly agreed mapping to RDF/OWL. There are some existing approaches translating (parts of) MPEG-7 into RDF/OWL. The first one was created by Hunter [11]. This MPEG-7 ontology was firstly developed in RDF Schema, then converted into DAML+OIL [12], and is now available in OWL Full. The ontology covers the upper part of the Multimedia Description Scheme part of the MPEG-7 standard. Starting from the ontology developed by Hunter, the MPEG-7 OWL DL ontology by Tsinaraki covers the full Multimedia Description Scheme part of the MPEG-7 standard [13]. The MPEG-7 ontology by Rhizomik has been produced fully automatically using an implemented generic mapping tool (XSD2OWL) and is an OWL Full ontology, [14]. According to [15] there exists an OWL DL version of this ontology, but at the time of writing, the ontology available at the specified location is OWL Full.

2.3 IPTC

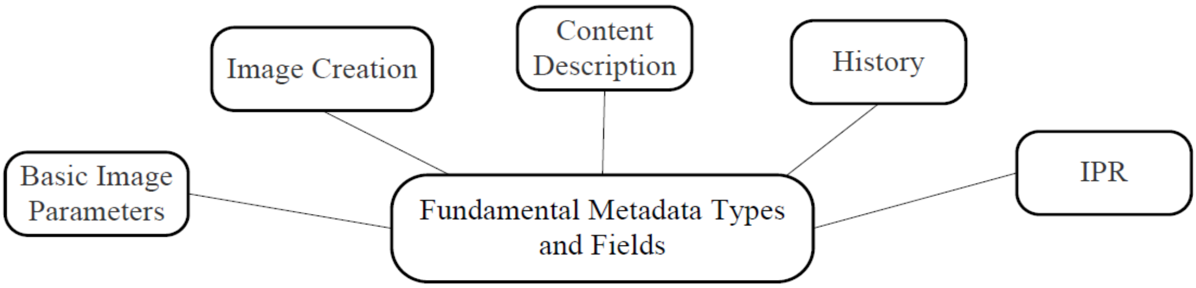
The International Press Telecommunications Council (IPTC) Core Schema is a metadata set primarily for photographer's use and aligns with the IPTC Headers, [16]. These IPTC photo metadata categorize the metadata fields into four groups regarding to the semantics they describe: Administrative, Content Descriptive, Rights, and Technical metadata [17]. The Administrative metadata are data about the content that cannot be inferred from the content, such as a free text field for describing the event at which the photo was taken, creation time, and any number of instructions from the provider to the receiver of the photo. The Content Descriptive metadata consists of 12 fields, including a field keywords to express the subject of the content in free text, and two fields to describe the most prominent subjects of the photo by one or more codes. However, the use of free text fields decreases the queriability of data collections. The Rights, and Technical metadata describe the rights (e.g., a copyright notice), and metadata about the hardware and creational parameters, respectively.

2.4 PhotoRDF and Dublin Core

Other standards for describing metadata of digital images using RDF, are Dublin Core (DC [18]) and PhotoRDF [19]. DC consists of a rudimentary set of 15 elements describing common properties of resources, such as title and creator. PhotoRDF is an attempt to standardize a set of categories and labels for personal photo collections using the DC schema as well as an additional schema for technical and content data. The content schema contains 10 fixed keywords to be used in the subject property of the DC schema, such as "Portrait" and "Baby".

2.5 VRA

Describing large collections of (annotated) slides, images and other representations of works of art, the VRA Core has been defined. VRA Core is a set of metadata elements used to describe works of visual culture as well as the images that represent them. In the context of VRA Core, a work is a physical entity that exists, has existed at some time in the past, or that could exist in the future. Similarly to DC, VRA Core defines a set targeted especially at visual



resources. There currently exists no commonly accepted mapping from VRA Core to RDF/OWL; at least two conversions have been proposed: the RDF/OWL representation of VRA by M. van Assem [20] and the VRA ontology by SIMILE [21].

Figure 1 Composition of the DIG35 standard.

2.6 DIG35

Most standards explained above allow full text annotations, provide a fixed set of descriptive keywords, and MPEG-7 lacks high-level semantics. A multimedia ontology should be rich enough to describe the document and its content, as well as the spatio-temporal relationships between the depicted entities. The DIG35 standard, a specification of the International Imaging Association (I3A), defines a set of public metadata for digital images covering a broad spectrum of metadata fields [2].

Figure 1 shows the composition of the DIG35 metadata definition. It consists of five logical blocks with a separate common definition, i.e., the fundamental metadata types and fields that is referred to by the other blocks. While each block is logically partitioned, they may be linked to each other to form additional semantics.

The first metadata block covers basic image parameters to specify generic information about the image, such as file name, format, and size. The creational metadata is grouped in a second block. It defines the metadata that are relevant to the creation of the digital image data. Many fields defined in this block are actually the same as defined by EXIF. The third container comprises the content description of an image. While many image metadata standards only provide limited support to describe the content, e.g., by keywords or a free text description, DIG35 offers fine grained description methods to model and relate depicted concepts within an image. The history metadata block holds partial information about how the image got to the present state; however, it is not designed to be used to reverse image editing operations. Finally, the last block defines the intellectual property rights metadata (IPR) which are designed to protect the contents of an image file from misuse and must preserve both moral rights and copyrights.

An example of a DIG35 instance is shown here:

```
<METADATA>
  <BASIC_IMAGE_PARAM>
    <BASIC_IMAGE_INFO>
      <IMAGE_ID>
        <UID>098f2470-bae0-11cd-b579-08002b30bfed</UID>
        <ID_TYPE>http://www.imaging.org//UUID</ID_TYPE>
      </IMAGE_ID>
    </BASIC_IMAGE_INFO>
  </BASIC_IMAGE_PARAM>
  <IMAGE_CREATION>
    <IMAGE_CREATOR>
      <PERSON_NAME>
          <NAME_COMP TYPE="Given">
                  Yosh
          <NAME_COMP>
          <NAME_COMP TYPE="Family">
                  Shibata
          <NAME_COMP>
      </PERSON_NAME>
    </IMAGE_CREATOR>
    <SOFTWARE_CREATION>
          <SOFWARE_INFO>
                  <Model>Wizzo Extracto</MODEL>
                  <VERSION>2</VERSION
          </SOFTWARE_INFO>
    </SOFTWARE_CREATION>
  </IMAGE_CREATION>
  <CONTENT_DESCRIPTION>
      <LOCATION>
        <ADDRESS>
          <ADDR_COMP TYPE="City">
                  Tokyo
          </ADDR_COMP>
          <COUNTRY>jp</COUNTRY>
        </ADDRESS>
        <CAPTURE_TIME>
          <EXACT>2000-10-10T19:45:00+09:00<EXACT>
        </CAPTURE_TIME>
      </LOCATION>
  </CONTENT_DESCRIPTION>
  <IPR>
      <IPR_IDENTIFICATION>
          <IPR_IDENTIFIER>
                  <IPR_ID>RID#</IPR_ID>
          </IPR_IDENTIFIER>
      </IPR_IDENTIFICATION>
  </IPR>
```

	PhotoRDF	Exif	VRA Core	IPTC Photo	MPEG-7	DIG35
Low level	-	-	-	-	✓	-
Creational	limited	✓	limited	✓	✓	✓
Technical	-	✓	-	✓	-	✓
Content	*keywords*	*title*	✓	✓	✓	✓
History	-	-	-	-	✓	✓
IPR	*copyright*	*copyright + holder*	*description &rights*	✓	✓	✓

</METADATA>

Table 2 Comparison of image metadata standards.

Table 1 lists a comparison between DIG35 and other image metadata standards about the different metadata types they cover. For example in PhotoRDF and VRA Core, the IPR metadata is described by only one field, and they offer only limited support for creational metadata. The application of the DIG35 fundamental types and fields permits a more detailed and conceptualized description than by simply using keywords or by a full text explanation.

3 Ontology creation

As the plethora of metadata formats, DIG35 typically consists of an XML Schema accompanied with plain text. An effect of using XML as underlying technology is that it is difficult to deduce the semantic information from the XML structure which makes it hard to be machine interpretable. An additional intricacy for converting the XML Schema to an OWL ontology is caused by the fact that not all semantic information can be deduced from the XML structure; the accompanying plain text describes the semantical meaning of the XML and even some supplementary semantics. Another consequence related to the use of XML are interoperability problems between different metadata standards [7]. A solution is to use formal representations by applying Semantic Web Technologies.

The use of Semantic Web Technologies for the representation of multimedia metadata is also proposed by working groups of the W3C, like the W3C Multimedia Semantics Incubator Group [22] and the W3C Media Annotations WG in which the authors of this paper have actively participated. In the remainder of this section, we give an overview of some common difficulties that can arise when converting an XML Schema into an OWL ontology.
Furthermore, some solutions we applied during the modeling of the DIG35 ontology, are proposed. As will be shown, the current W3C recommendation, i.e., OWL 1.0, is insufficient to tackle many common problems for converting an XML Schema with their accompanying text into an ontology without losing formal semantics.

3.1 Identifiers

One of the major problems of OWL 1.0 is its well-known inability to customize the standard data types4 to deal with data ranges. According to the DIG35 specification, many properties should have a range of a non-negative double, e.g., the subject distance, the focal number, color temperature, and the iso saturation. Due to this incapability, we were obliged to use the whole range, i.e. *xsd:double*. In OWL 2.0 (previously known as OWL 1.1), “complex” data ranges can be constructed from the simpler ones using the *dataOneOf*, the *dataComplementOf*, or the *datatypeRestriction* constructor, combined with one of the available facets to express a restriction (e.g., *minInclusive*, *maxExclusive*), [23].

According to the DIG35 specifications, many concepts can have an associated identifier. Expressing the concept of an identifier (cf. a key in a relational database system) in OWL, can be accomplished using the *owl:InverseFunctionalProperty*. However, there is a restriction: the latter property type is only applicable for object properties. Consequently, if one wants to use a literal value as identifier (ID), an additional ID-class should be created. This ID-class holds a data type property uid to refer to the value of the ID. A drawback of this approach is that some extra complexity is added since one ends up with the definition of 2 properties and an extra class for modeling one single concept.

3.2 Order of Appearance

The Image Processing Hints in the History metadata section of DIG35 specifies a list of the operations applied when editing an image. Consequently, the order in which the operations are listed is of great importance. RDF provides two container mechanisms to encapsulate data in a user defined order: the *rdf:Seq* class which is used for representing ordered lists of literals or resources and *rdf:List* which is a class for representing a closed list of items. However, those predefined classes are not OWL DL compatible. To formalize the order in which an object or literal appears in a list, there are two approaches.

A first method is to create a (double) linked list. The linked list is navigated by the two properties next and previous while the object is referred by the property value. When a new item is inserted, the corresponding properties of the previous, next, and current item need to be adjusted. The latter, however, cannot be formalized in the ontology and consequently, should be carried out at a higher level (e.g., by the software).

A second approach, the one which is implemented in the DIG35 ontology, is to create an item-class with two properties representing the order and the value. The first property, a data type property (order), defines the order of the object and the second (value) refers to the object itself. The item-class is to be referred to by an *owl:ObjectProperty* with no cardinality restrictions. However, the constraint that the value of the order property should be unique within the scope of the list, cannot be expressed in OWL. So, in OWL DL only a partial solution for this problem can be elaborated.

3.3 Reusing existing ontologies

Many existing ontologies are OWL Full which renders them difficult to reuse and import in existing ontologies [24]. Examples of some popular OWL Full ontologies, are Friend Of A Friend (FOAF [25]), DC, and Simple Knowledge Organization Schema (SKOS [26]). In most cases, ontologies are rather OWL Full due to syntactic errors or accidental misuse of the vocabulary, such as the use of RDF-properties instead of OWL-properties. On the other hand for SKOS, the restrictions on OWL DL prevent treating SKOS concepts as OWL classes. Since a SKOS concept is defined as an OWL class, an instance of a concept should also be an OWL instance, but according to SKOS, an instance must be treated as a class. The latter statement is only supported by OWL Full.

Meta-modeling, i.e., the treatment of classes, properties and other entities as individuals, is partially allowed by OWL 2.0: a name can be used for any or all of an individual, a class, or a property. For DC, there is an OWL DL version available at the website of Protégé. However, a DC property is modeled using an *owl:AnnotationProperty*. The range of the latter must be an individual, literal, or URI which makes it impossible to refine an existing OWL DL ontology using this version of DC. For example, it is not possible to express that the datatype property *fileFormat* is an equivalent property or sub-property of *dc:format*, or that the range of *dc:creator* is the class *Person*. As a result of the OWL Fullness of many ontologies we did not reuse any existing ontologies during the modeling process in order to keep the DIG35 ontology OWL DL. For practical implementations, an extra mapping ontology to an existing OWL Full ontology can always be created by importing both ontologies and employing the typical OWL or RDF constructs (e.g., *owl:equivalentClass* and *rdfs:subPropertyOf*) between the concepts of the two ontologies.

Today, many reasoners accept to a certain extend OWL Full ontologies. In fact, by adding a few extra, clarifying statements, a large proportion of the available OWL Full Ontologies could be validated as OWL DL. However, any OWL Full feature that is not supported by the reasoner, will be ignored, resulting in a loss of semantics.

4 The ontology

The obtained DIG35 ontology is created manually and permits to deduce semantic relationships that are not defined in the provided XML Schema. According to the different parts that are covered by the metadata specification (see Sect. 2.6) and the semantics they describe, the DIG35 ontology consists of five subontologies: Basic Image Parameters, Image Creation, Content Description, Image History, and IPR. Additionally, a number of fundamental ontologies are created to represent basic concepts:

- Address
- Audio stream,
- Date & Time,
- Direction (describes the direction to specify a 3D heading),
- Email & Web Address,
- Event,
- Location,
- Organization,
- Person,
- Phone number,
- Position (of an object within an image),
- Product details (specifies details about hardware or software)

- Tangible thing.

These ontologies are domain independent and can be reused for different purposes. The legitimacy of the produced OWL ontologies is validated with the WonderWeb OWL Ontology Validator which is recommended by W3C. The metrics of the DIG35 ontology are listed in Table 1.

Table 3. DIG35 Ontology Metrics

Number of classes	152
Number of data type properties	151
Number of object properties	146
Number of restrictions	185
Expressivity	*ALUOIN*(D)

The DIG35 ontology is available at http://multimedialab.elis.ugent.be/users/gmartens/Ontologies/DIG35.

4.1 Discussions

We have explained the modeling of the DIG35 standard for digital still images into an OWL DL ontology. The obtained DIG35 ontology offers fine grained description methods to model and relate depicted concepts, intellectual property rights, history metadata, basic parameters as well as creational metadata. We have emphasized that the current W3C recommendation for Semantic Web technologies, i.e., OWL 1.0, is insufficient to model many important semantics. The application of OWL 2.0 eliminates some, but certainly not all impediments that can arise in ontology modeling, such as additional meta-modeling capabilities or support for containers.

In order to be OWL DL compliant, we worked out solutions for the faced problems. However, some solutions often introduce some extra complexity since additional properties and classes are required to model certain concepts. The fact that many popular ontologies are OWL Full certainly reduces their integration in new ontologies. The latter constatation may result in an expansion of ontologies modeling the same concepts and consequently in a fragmentation of the Semantic Web. However, the lack of additional modeling capabilities by OWL will certainly favor the creation of many OWL Full ontologies as well, and consequently in its turn, compelling additional reasoning support for certain OWL constructs.

Note that the ontology that was constructed holds all information present in the current DIG35 standard. However, many annotations are available that still use the XML format. Consequently, a mechanism is needed that converts the XML instances into instances of the ontology. This will be discussed in the next section.

5 XML to RDF

Today, many metadata standards are expressed in XML Schema and there are numerous multimedia documents with XML-based annotations. Moreover, several software agents have specific XML-based parsers to interpret and produce multimedia metadata. Consequently, a conversion tool is needed to automatically generate the RDF/OWL triples out of the XML fragments. For this purpose, two major approaches exist: fixed XML to RDF mappings and ontology dependent XML to RDF mappings. The former uses a mapping based on the XML schema and disregards the actual ontology [27-29]. The ontology-dependent mappings specify a XML to RDF mapping document that is specifically tailored for the actual ontology [30-33]. They mostly differ in the way the mapping document is created.

In our previous we created a generic XML to RDF convertor [30]. This convertor uses an XML document as mapping document that defines specific mapping rules between an XML instance document and the resulting RDF document. A generic XMLtoRDF tool takes this mapping document and the used ontology as input and then automatically transforms corresponding XML documents to RDF instances, as shown in Figure 2.

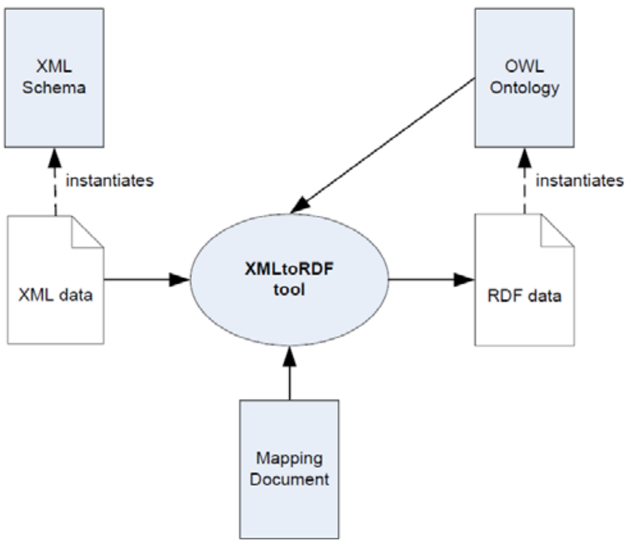


Figure 2 XML to RDF conversion.

The XML-based mapping document is built from specific elements which form a mapping language and can be interpreted by the XMLtoRDF tool. This language allows creating a simple mapping of XML nodes to corresponding OWL classes or properties.

Conditional mappings are available in case a mapping not always holds. In that case, a condition can be made of XPATH (XML Path Language) or SPARQL ASK expressions. Finally, value processing is included which specifies different ways to infer the value of a resulting OWL property. These specific language constructs ease the development of such XML to RDF mappings.

Lifting a metadata standard to the Semantic Web allows for more straightforward reasoning en inferencing. Additionally, relationships between concepts of the metadata model can be more formally described. However, to foster re-use the ontology should be coupled to other multimedia initiatives in the Semantic Web. This is the topic of the following sections.

6 Media Annotations

The W3C Media Annotations Working Group (MAWG) has the goal of improving the interoperability between media metadata schemas. The proposed approach is to provide an interlingua ontology and an API designed to facilitate cross-community data integration of information related to media resources in the web, such as video, audio, and images.

A first ontology was created, Media Ontology 1.0, with the goal to find common properties in multimedia annotations. Additionally, an API is being defined that allows uniform access to the underlying Media Ontology.

The set of core properties that constitute the Media Ontology 1.0 is based on a list of the most commonly used annotation properties from media metadata schemas currently in use. This set is derived from the work of the W3C Incubator Group Report on Multimedia Vocabularies on the Semantic Web and a list of use cases [34], compiled after a public call. The use cases involve heterogeneous media metadata schemas used in different communities (interactive TV, cultural heritage institutions, etc.). In this section, we describe the content of this ontology and how it is related to other formats.

The set of core properties is defined in the *ma* namespace and consists of 20 descriptive and 8 technical metadata properties. This distinction has been made as the descriptive properties are media agnostic and also apply to descriptions of multimedia works that are not specific instantiations, e.g. the description of a movie on IMDB in contrast to a particular MPEG-4 encoded version of this movie broadcasted of the RAI Italian TV channel. The technical properties, specific to certain media types, are only useful when describing a certain instantiation of the content.

All properties are defined within the *ma* namespace. Whenever the properties exist in other standards, it is explicitly shown how they are related. For many of the descriptive properties, subtypes are foreseen that optionally further qualify the property, e.g., qualify a title as main or secondary.

The descriptive properties contain identification metadata, such as identifiers, titles, languages and the locator of the media resource being described. Other properties describe the creation of the content (containing the creation date and location and the different kinds of creators and contributors), the content description as free text, the genre, a rating of the content by users or organizations, and a set of keywords. There are also properties to describe collections of which the described resource is part of and to express relations to other media resources, e.g., source and derived works, thumbnails or trailers. Digital rights management is considered out of scope, the set of properties only contains a copyright statement and a reference to a license (e.g., Creative Commons or MPEG-21 licenses). The distribution related metadata includes the description of the publisher and the target audience in terms of regions and age classification. Annotation properties can be attached to the whole media or to part of it, for example using the Media Fragments URI specification for identifying multimedia fragments. The set of technical properties has been limited to the frame size of images and video, the duration, the audio sampling and frame rate, the format (specified as MIME type), the compression type, the number of tracks, and the average bit rate.

It should be clear that DIG35 holds many correspondences with the MA ontology. In the next section we will provide ways to map both ontologies together, allowing an automated conversion of DIG35 instances to instances of the MA ontology.

7 Mappings

To make the DIG35 standard available through the media ontology we need to define a mapping. This mapping relates concepts and properties of the DIG35 ontology to those of the MA ontology. Using the mappings, DIG35 instance data can consequently be transformed to instance data of the MA ontology, which can be accessed through the API. The MA ontology is created for usage in different domains, not restricted to image annotation. It can be foreseen that browsers will provide implementations of the API and MA ontology to give access to the metadata properties of multimedia resources. The mapping to the ontology prevents that companies do not need to implement parsers specific for DIG35 (or other metadata standards).

The first step is to identify those concepts that correspond to the MA properties. For this purpose, an analysis is needed of the specifications. The ontology document for the Media Ontology, and the textual specification of DIG3 need to be searched for similar concepts. For DIG35 we created a so called mapping table [4]. This table holds the specific relationships between the concepts of the MA ontology and DIG35. Special care needs to be taken of the actual relationship, since concepts can have different semantics, or syntactic features.

Once these concepts are found, a way to define these relations, or mappings is needed. We propose to use OWL to express direct mappings between the Media Ontology and DIG35 ontology [35]. The mappings expressed in OWL are in fact by themselves an ontology, called a mapping ontology. Such an ontology typically consists of basic OWL or RDFS constructs (e.g., *owl:equivalentClass* and *rdfs:subPropertyOf*) with the sole purpose to relate concepts of different ontologies. The following example defines a formal semantic equivalence between the title property defined in DIG35 and in the Media Ontology:

```
dig:descriptionTitle owl:equivalentProperty ma:title
```

These constructs can be used for all properties that have the same semantic and structural characteristics. Note that, for practical implementations, a mapping ontology as presented above is not sufficient. Rules are needed to create advanced conditional relationships, e.g., to declare instance equivalence when certain properties match. Logical rules can be employed to do any type of conversion and transformation of values (e.g., convert bps to kbps). Following example expresses in SWRL [36] that the value of *ma:frameSize* property can be filled from the values of the *dig:width* and *dig:height*:

```
[r1: (?res rdf:type dig:BasicParam) ^
     (?res dig:imageWidth ?width ) ^
     (?res dig:imageHeight ?height )
                   ⇨  (?size1 rdf:type ma:Size ) ^
                      (?size1 ma:width ?width) ^
                      (?size1 ma:height ?height) ^
                      (?size1 ma:unit "pixels")      ]
```

First the rule searches for triples describing the basic parameters of DIG35. Then it stores the image width and height in two variables. Next, it creates an instance of the *ma:Size* class and fills in the width and height. Additionally, since image width and height, as defined by the DIG35 standard, is always expressed in pixels, we can include this in the rule by setting the *ma:unit* property to "pixels".

Another example of the use of rules is string manipulation through built-in functions. Following example shows how the name of the creator is created from DIG35 metadata:

```
[r2: dig:Person(?p1) ^
     dig:imageCreator(?x2,?p1) ^
     dig:givenName(?p1,?y1) ^
     dig:familyName(?p1,?y2)^
                  ⇨  strcat(?name, ?y1, "", ?y2)^
                     ma:creator(?resource,?p1)^
                     foaf:name(?p1,?name)       ]
```

In this example, we first search for an instance of the Person class in DIG35, more specifically the person should be the creator of the image. Next, the given and family name is stored in two variables. These variables are used to create the name by concatenation. Additionally, it is stated that the person is the creator of the media resource using the *ma:creator* property. Currently, the creator is represented as an identifier which might correspond to a FOAF Person. As such, by creating a *foaf:name* triple we can state that the person is a FOAF person and fill in the entire name.

At this point we can lift DIG35 XML instances to OWL instances, and through the mappings to instances of the MA ontology. The next section shows how this can be used.

8 Evaluation

A number of different approaches can be taken for creating the mappings.

Cruz et al. use a proprietary mapping table that links an upper ontology to local RDF ontologies [10]. A query upon this upper ontology is therefore first translated, using the mapping table, to different queries on the local RDF ontologies. Moreover, according to their architecture, these queries are translated to XML queries that are executed upon the XML sources. This introduces much overhead compared to the proposed system where only one query is executed on the RDF data to retrieve the same results. Garcia et al. used an automatically generated MPEG-7 ontology and presented an architecture to achieve semantic multimedia metadata integration and retrieval [22]. They use XSD2OWL to automatically create an ontology based on an existing XML schema. However, this tool is made to allow the automatic conversion from MPEG-7 XML schema to an OWL ontology (and only results for MPEG-7 conversions were presented). We noticed that this tool is not usable for the conversion of the DIG35 XML schema. For example, an XSD2OWL translation is used that translates an XML sequence to an intersection of classes (denoted by the *OWL:intersectionOf* construct). Within the DIG35 schema the element ContentDescription is defined as a sequence of different elements (Caption, Location, Person, Thing, Comment, etc.). Some of these would be translated to *owl:Class* constructs, others to *owl:DatatypeProperty* constructs, which would invalidate the intersection. Moreover, Garcia et al. do not define ways to map different ontologies on each other. By using the XSD2OWL conversions, making a mapping would only be possible if the different metadata schemes use the same names for the same concepts, which is obviously not always the case.

9 Use case

The MA ontology is foreseen to be accessible through a standardized API. Browsers will implement the API so that Web developers can request the metadata of multimedia resources. This allows creating richer search engines, news aggregators and so on. Using the guidelines described above, a photo collection with annotations in DIG35 metadata, can be lifted to the MA-space. Web-sites can consequently more easily extract information from the collection and show for instance all pictures created by a specific person. Such an example can be found at [37]. This prototype includes a client-side implementation of the API (in JavaScript). Upon requests by a user, the API interrogates with the MA ontology. Underlying this ontology are a number of different metadata formats (e.g., EXIF, MPEG-7, Dublin Core, and DIG35). The underlying information is represented as XML, and can be queried using the uniform interface.

10 Conclusions

In this paper, we have presented guidelines to lift an XML-based metadata standard to the Semantic Web. In a first step, we compared different multimedia metadata standards and have chosen DIG35 as the underlying format. We have discussed different modeling issues that arise when creating an ontology based on an existing metadata scheme. Ways have been presented to automatically convert instances of the XML-based standard to instances of the OWL-based ontology. Furthermore, the created ontology is mapped upon the MA ontology, an ongoing initiative of the W3C. We have shown how Semantic Web Technologies can be used to accomplish this mapping.

Our approach allows that current DIG35-compliant software can still produce XML-fragments, which are then lifted in the Semantic Web and coupled with the MA ontology. The same approach can be used for other multimedia metadata standards used in different communities.

References

[1] J. R. Smith and P. Schirling, Metadata Standards Roundup, *IEEE Multimedia*, Vol. 13, No. 2, 2006, pp. 84-88.

[2] Digital Imaging Group (DIG), DIG35 Specification - Metadata for Digital Images - Version 1.1 April 16, 2001, available at http://www.i3a.org/i dig35.html

[3] S. Bechhofer, F. van Harmelen, J. Hendler, I. Horrocks, D. L. McGuinness, P. F. Patel-Schneider, and L. Andrea Stein, OWL Web Ontology Language Reference, http://www.w3.org/TR/owl-ref/

[4] W. Lee, T. Bürger, F. Sasaki, V. Malaisé, F. Stegmaier, and J. Söderberg. Ontology for Media Resource 1.0. W3C Working Draft, 2009. http://www.w3.org/TR/mediaont-10/

[5] W. Lee, F. Stegmaier, and C. Poppe. API for Media Resource 1.0. W3C Working Draft, 2010. http://www.w3.org/TR/mediaont-api-1.0

[6] Japan Electronic Industry Development Association (JEIDA), "Digital still camera image file format standard (exchangeable image file format for digital still cameras: Exif) version 2.1", Technical report, 1998

[7] S. Decker et al, The semantic web - on the respective roles of XML and RDF, IEEE Internet Computing, 2000

[8] M. Kanzaki, "The Kanzaki Exif RDF Schema", http://www.kanzaki.com/ns/exif

[9] The Norm Walsh Exif RDF Schema, http://sourceforge.net/projects/jpegrdf

[10] MPEG-7 Overview, International Organization for Standardisation, Klagenfurt ISO/IEC JTC1/SC29/WG11, July 2002, http://mpeg.telecomitalialab.com/standards/mpeg-7/mpeg-7.htm

[11] J. Hunter, Adding Multimedia to the Semantic Web Building an MPEG-7 Ontology, In International Semantic Web Working Symposium (SWWS 2001)

[12] I. Horrocks, F. van Harmelen, P. Patel-Schneider, T. Berners-Lee, D. Brickley, D.

Connolly, M. Dean, S. Decker, D. Fensel, P. Hayes, J. Heflin, J. Hendler, O. Lassila, D. McGuinness, and L. Stein, "DAML+OIL", 2001 http://www.daml.org/2001/03/daml+oil-index
[13] C. Tsinaraki, P. Polydoros and S. Christodoulakis, Interoperability support for Ontology-based Video Retrieval Applications, In Proc. of 3rd International Conference on Image and Video Retrieval (CIVR 2004), pp. 21-23, July 2004
[14] R. Garcia and O. Celma, Semantic Integration and Retrieval of Multimedia Metadata, In Proc. of the 5th International Workshop on Knowledge Markup and Semantic Annotation (SemAnnot 2005), 7 November 2005.
[15] R. Troncy, O. Celma, S. Little, R. Garcia and C. Tsinaraki, MPEG-7 based Multimedia Ontologies: Interoperability Support or Interoperability Issue?, 1st International Workshop on Multimedia Annotation and Retrieval enabled by Shared Ontologies (MAReSO), Italy, 2007
[16] IPTC Metadata, http://www.iptc.org/
[17] H. Löffler, W. Baranger, Photo Metadata White Paper 2007, IPTC, 2007, available at http://www.iptc.org/goto?phmdwp2007
[18] Dublin Core Metadata Initiative, "Dublin Core Metadata Element Set, version 1.1: Reference Description", Tech. Rep., 2004, available at http://www.dublincore.org/documents/dces/
[19] Y. Lafon, B. Bos, "Describing and Retrieving Photos using RDF and HTTP",W3C Note (19-April-2002), http://www.w3.org/TR/2002/NOTE-photo-rdf-20020419
[20] M. Van Assem. RDF/OWL Representation of VRA http://www.w3.org/2001/sw/BestPractices/MM/vra-conversion.html
[21]http://simile.mit.edu/2003/10/ontologies/vraCore3
[22] World Wide Web Consortium (W3C) Multimedia Semantics Incubator Group, http://www.w3.org/2005/Incubator/mmsem/
[23] OWL 1.1 Web Ontology Language Structural Specification and Functional-Style Syntax, http://www.webont.org/owl/1.1/
[24] S. Bechhofer, R. Volz, "Patching Syntax in OWL ontologies", International Semantic Web Conference 2004, pp. 668-682, 2004
[25] D. Brickley, L. Miller, "FOAF Vocabulary Specification", 2006, http://xmlns.com/foaf/0.1/
[26] A. J. Miles, N. Rogers, and D. Beckett, "SKOS Core RDF Vocabulary", 2004, http://www.w3.org/2004/02/skos/core
[27] R. Garcia, O. Celma, Semantic integration and retrieval of multimedia metadata, in: Proceedings of the Fifth Knowledge Markup and Semantic Annotation Workshop, 2006, pp. 69–80.
[28] S. Battle, Gloze, XML to RDF and back again, in: Proceedings of the First Jena User Conference, 2006.
[29] M. Ferdinand, C. Zirpins, D. Trastour, Lifting XML schema to OWL, in: Proceedings of the Web Engineering – Fourth International Conference, 2004, pp. 354–358.
[30] I.F. Cruz, H. Xiao, F. Hsu, An ontology-based framework for XML semantic integration, in: Proceedings of the International Database Engineering and Applications Symposium, 2004, pp. 217–226.
[31] M. Klein, Interpreting XML documents via an RDF schema ontology, in: Proceedings of the DEXA Workshop, 2002, pp. 889–894.
[32] B. Amann, C. Beeri, I. Funulaki, M. Scholl, Ontology-based integration of XML web resources, in: Proceedings of the First International Semantic Web Conference on The Semantic Web, 2002, pp. 117–131.
[33] D. Van Deursen, C. Poppe, G. Martens, E. Mannens, R. Van de Walle, XML to RDF Conversion: a generic approach, in: Proceedings of the Fourth International Conference on Automated Solutions for Cross Media Content and Multi-Channel Distribution, 2008
[34] W. Lee, T. Bürger, F. Sasaki, and V. Malaisé. Use Cases and Requirements for Ontology and API for Media Object 1.0. W3C Working Draft, 2009. http://www.w3.org/TR/media-annot-reqs/
[35] Chris Poppe, Ga¨etan Martens, Erik Mannens, and Rik Van de Walle. Personal Content Management System: A Semantic Approach. Journal of Visual Communication and Image Representation, 20(2):131–144, 2009.
[36] I. Horrocks, P. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, M. Dean, SWRL: a semantic web rule language-combining owl and rule ML, W3C Member Submission, 21 May 2004. Available from: http://www.w3.org/Submission/SWRL
[37] C. Poppe, MA annotation test implementation 2010 http://multimedialab.elis.ugent.be/users/chpoppe/MediaAnnotations.html