How to Align Media Metadata Schemas? Design and Implementation of the Media Ontology

Florian Stegmaier[1], Werner Bailer[2], Tobias Bürger[3], Mario Döller[1], Martin Höffernig[3], Wonsuk Lee[4], Véronique Malaisé[5], Chris Poppe[6], Raphaël Troncy[7], Harald Kosch[1] and Rik Van de Walle[6]

[1] Chair of Distributed Information Systems, University of Passau, Germany
[2] Institute of Information Systems, JOANNEUM RESEARCH, Graz, Austria
[3] Semantic Technology Institute (STI), University of Innsbruck, Austria
[4] Standards Research Center, ETRI, Daejeon, Korea
[5] Web and Media Group, VU University, Amsterdam, Netherlands
[6] Department of Electronics and Information Systems, Ghent University - IBBT, Belgium
[7] EURECOM, Sophia Antipolis, France

Abstract. Multimedia data is generated, shared, stored and distributed worldwide at an ever increasing rate. This huge amount of content comes with metadata represented in different formats which hardly interoperate although they partially overlap. The W3C Media Annotations Working Group is chartered to recommend a Media Ontology compatible with most of these schemas. In this paper, we present the process for modeling this ontology and we discuss various approaches for explicitly representing the mappings between the core set of annotation properties defined in the Media Ontology and some major deployed metadata standards. We highlight the benefits and drawbacks of each approach and conclude on future work for the implementation of the Media Ontology.

1 Introduction

The publication and consumption of multimedia data on the Web has grown heavily thanks to the multiplicity of photo and video sharing platforms, usually embedded within social networks, along with the spread of multimedia enabled mobile devices. This huge amount of content can be generally accessed either via standardized and proprietary metadata formats, or more directly via APIs attached to web sites. As a result, the content is often locked in within silos preventing an effective search across these sites and making it complicated to create mashable applications.

While the multimedia metadata formats used on the web largely overlap in their coverage, they are at the same time dissimilar in many ways. **Coverage**: MPEG-7 [9] for example aims to be domain independent while DICOM [10] focuses on medical images, videos and workflows; **Comprehensiveness**: For example, MPEG-7 aims to provide comprehensive descriptions of multimedia content ranging from low-level features that can be extracted automatically to fine-grained semantic description of a scene, while Dublin Core [6] provides a simple list of general annotation properties and EXIF focuses on the technical aspects of the media; **Complexity**: Metadata formats also differ in the complexity of their description syntax. For example, the Dublin Core `dc:creator` property is a simple name or an URI identifying an agent whereas the creator's name in MPEG-7 is divided into a complex nested structure of `Title`, `FamilyName` and `GivenName` along with the definition of his or her `Role`.

Designing multimedia systems nowadays often amounts to choose a subset of these various formats and implements manually their correspondence which severely hampers their interoperability. In this paper, we report on the design and implementation of the Media Ontology developed by the W3C Media Annotations Working Group (MAWG)[8] which aims at defining a set of minimal annotation properties for describing multimedia content along with a set of mappings between the main metadata formats in use at the moment. This ontology being described in prose, we investigate and discuss different options of formalization and implementation of its core annotation properties and the defined mappings with other standard formats.

The remainder of this paper is organized as follows. Section 2 presents multimedia metadata formats between which interoperability is necessary, and an overview of interoperability approaches for XML or RDF/OWL-based schemas. Section 3 presents the Media Ontology and the process of its elaboration. Section 4 discusses various implementation approaches for representing the ontology itself and the mappings between multimedia formats. Finally, Section 5 concludes the paper and outlines some future work.

2 Related work

Several standards have been created to improve the interoperability between different systems within one domain or application type. In this section, we describe some image and video metadata standards (*i.e.* schemas), and discuss some approaches for combining them. An exhaustive list of multimedia metadata formats has been produced by the W3C Multimedia Semantics Incubator Group[9].

2.1 Many Standards for Different Needs

Photos taken by digital cameras come with Exchangeable Image File (EXIF[10]) metadata directly embedded into the header of image files. It provides technical characteristics such as the shutter speed or aperture, and contextual information (date and time) of the captured image. Two RDFS ontologies of this specification have been proposed by Kanzaki and Norm Walsh. The Extensible Metadata Platform (XMP[11]) is a specification published by Adobe for attaching metadata

[8] `http://www.w3.org/2008/WebVideo/Annotations/`
[9] `http://www.w3.org/2005/Incubator/mmsem/XGR-vocabularies/`
[10] `http://www.digicamsoft.com/exif22/exif22/html/exif22_1.htm`
[11] `http://www.adobe.com/devnet/xmp/`

to media assets in order to enable a better management of multimedia content. The specification standardizes the definition, creation, and processing of metadata by providing a data model, a storage model, and formal predefined sets of metadata property definitions. XMP makes use of RDF in order to represent the metadata properties associated with a document. The DIG35[12] specification of the International Imaging Industry Association (I3A) defines a standard set of metadata for digital images including basic image parameter, image creation (à la EXIF), content creation and intellectual property rights and represented in XML. The IPTC Photo Metadata standard[13] developed by the International Press Telecommunication Council (IPTC) provides also a set of metadata properties being administrative, descriptive or related to the image rights. Largely based on XMP, this specification allows to represent as well complex semantic descriptions of the subject matter (e.g. persons, organizations, events).

EBUCore[14] is an XML-based metadata standard created by the European Broadcasting Union (EBU) consisting in a set of metadata properties specializing Dublin Core for describing radio and television content. MPEG-7 [9] is the Motion Pictures Expert Group (MPEG)[15] standard for the description of audio, video and multimedia content designed for document retrieval. The standard is based on XML Schema but MPEG-7 ontologies expressed in OWL have been proposed and compared among each other [12]. The standard is composed of many descriptor tools for diverse types of annotations on different semantic levels, ranging from very low-level features, such as visual (e.g. texture, camera motion) or audio (e.g. melody), to more abstract descriptions. The flexibility of MPEG-7 is based on structuring tools, which allow descriptions to be associated with arbitrary multimedia segments or regions, at any level of granularity, using different levels of abstraction.

Numerous metadata standards exist for annotating multimedia resources, all with their own merits and community usage. It is undesirable to enforce a single multimedia metadata standard that would satisfy all use cases. Some additional steps are needed to combine these formats and interoperability can be achieved by the means of mappings or relationships between the different schemas. In the next section, we review approaches for structural (*i.e.* syntactic) and semantic integration of multimedia metadata schemas.

2.2 Interoperability Approaches between Metadata Schemas

JPSearch is a project issued by the JPEG standardization committee to develop technologies that enable search and retrieval capabilities among image archives, consiting of five parts. While the first part focus on describing use cases and the overall architecture of image retrieval systems, the part 2 introduces an XML-based core metadata schema and transformation rules for mapping descriptive

[12] `http://xml.coverpages.org/FU-Berlin-DIG35-v10-Sept00.pdf`
[13] `http://www.iptc.org/std/photometadata/2008/specification/IPTC-PhotoMetadata-2008_2.pdf`
[14] `http://tech.ebu.ch/docs/tech/tech3293-2008.pdf`
[15] `http://www.chiariglione.org/mpeg/`

information (e.g., core metadata to MPEG-7 or core metadata to Dublin Core) between peers [2]. Part 3 adapts a profile of the MPEG Query Format [3] for ensuring standardized querying. Part 4 adopts the well known image data formats (JPEG and JPEG 2000) for embedding metadata information. The benefit of such an integration and combination of metadata with raw data is the mobility of metadata and its persistent association with the image itself. By embedding the metadata into the image raw data file format, one improves the flexibility within the annotation life cycle. However, the interchange of image data between JPSearch compliant systems remains an open issue. For this purpose, Part 5 concentrates on the standardization of a format for the exchange of image or image collections and its metadata and metadata schema between JPSearch compliant systems.

Xing et al. [13] present a system for automating the transformation of XML documents using a tree matching approach. However, this method has an important restriction: the leaf text in the different documents has to be exactly identical. This is hardly the case when combining different metadata standards. Likewise, Yang et al. [14] propose to integrate XML Schemas. They use a more semantic approach, using the ORA-SS data model to represent the information available in the XML Schemas and to provide mappings between the different documents. The ORA-SS data model allows to define objects and attributes to represent hierarchical data, however more advanced mappings involving semantic relationships cannot be represented.

Cruz et al. [1] introduced an ontology-based framework for XML semantic integration. For each XML source integrated, a local RDFS ontology is created and merged in a global ontology. During this mapping, a table is created that is further used to translate queries over the RDF data of the global ontology to queries over the XML original sources. The authors assume that every concept in the local ontologies is mapped to a concept in the global ontology. This assumption can be hard to maintain when the number and the degree of complexity of the incorporated ontologies increases. Poppe et al. [11] advocates a similar approach to deal with interoperability problems in content management systems. An OWL upper ontology is created and the different XML-based metadata formats are represented as OWL ontologies and mapped to the upper ontology using OWL constructs and rules. However, the upper ontology is dedicated to content management system and, as such, is not as general as the approach proposed in this paper.

The W3C Multimedia Semantics Incubator Group[16] elaborated on the inherent problems of using XML-based metadata standards[17]. The goal of the group was to investigate the usage of Semantic Web Technologies to overcome interoperability issues. The group discussed the advantages and open issues regarding the use of Semantic Web technologies but was not chartered for providing one common ontology for metadata annotation.

---

[16] `http://www.w3.org/2005/Incubator/mmsem/`

[17] Such metadata standards consist generally of an XML schema defining a syntax and a textual description specifying in prose the semantics of the standard

3 The W3C Media Ontology

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.

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 [7], 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 this content is related to other metadata formats.

3.1 The Media Ontology Core Properties

The set of core properties defined in the Media Ontology 1.0 (`ma` namespace) 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 essential when describing a certain instantiation of the content[18].

All properties are defined within the `ma` namespace since we have tried to clarify and disambiguate their definitions in the context of media resources description. However, whenever these properties exist in other standards, we try to explicitly define how they are related. Additionally, for many of the descriptive properties, we have foreseen subtypes 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[19] of the media resource being described. Other properties describe the creation of the content (the creation date, creation location, the different kinds of creators and contributors, etc.), 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 the collections the described resource belongs to, and to express relations to other media resources, e.g. source and derived works, thumbnails or trailers. As we consider digital rights management out of our scope, the set of properties only contains a copyright statement and a reference to a license (e.g. Creative Commons or

[18] This distinction is also present in the FRBR model where a `Work` is distinguished from a `Manifestation`.

[19] The locator is the physical place where the resource can be accessed.

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 rate and frame rate, the format (specified as MIME type), the compression type, the number of tracks and the average bit rate. These were the only properties that were needed for the different use cases listed by the group.

This set of annotation properties is not considered final and properties might be added if it turns out to be useful. However, the aim is to keep the size of the ontology limited. If necessary, profiles can be defined, e.g. to group the properties that apply to a certain media type.

3.2 Expressing Mappings with other Standards

This core set of annotation properties has often correspondences with existing metadata standards. The working group has therefore further specified a mapping table that defines *one-way* mappings between the Media Ontology core properties and the metadata fields from 24 other standards [8].

The mappings that have been taken into account have different semantics, which can be characterized as:

- Exact matches: the semantics of the two properties are equivalent in most of the possible contexts. For example, `ma:title` matches exactly `dc:title`.
- More specific: the property of the vocabulary taken into account has a semantic that covers only a subset of the possibilities expressed by the property defined in the Media Ontology. For example, `ipr_names@description` and `ipr_person@description` defined in in DIG35 are more specific than the property `ma:publisher`.
- More generic: the inverse of the above, the property of the vocabulary taken into account has a semantic that is broader than the property defined in the Media Ontology. For example, `location` defined in the DIG35 is more general than `ma:location`.
- Related: the two properties are related in a way that is relevant for some use cases, but this relation has no defined semantics. For example, `media:credit` defined in MediaRSS[20] is related to `ma:creator`.

We discuss in the next section how these mappings can be represented.

4 Implementation Approaches

The W3C Media Ontology has been designed to be a meaningful subset of common annotation properties defined in standards used on the Web (see Section 2).

[20] `http://search.yahoo.com/mrss/`

The question is therefore how to implement or serialize the mapping relationships between the core set of properties defined by the Media Ontology and the other standards. This section discusses two classes of approaches: expressing a direct mapping using a more or less epxressive semantic web language (Sections 4.1 and 4.2) , or using a pivot upper ontology (Sections 4.3 and 4.4).

We illustrate each approach with a simple and a complex mapping between a property defined in the Media Ontology and its correspondence in another standard. These mappings concern the `ma:title` property which value is a simple string and the `ma:frameSize` property which value is composed of two integers representing the width and height of the video frames. The example 1.1 lists the prefixes we use for representing these mappings though all ontologies are not yet dereferencable.

```
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix ma: <http://www.w3.org/2009/09/mediaont#> .
@prefix ebu: <http://www.ebu.ch/metadata/ontologies/> .
```

Precondition 1.1. Declaration of prefix used in the examples.

4.1 Expressing Mappings in SKOS

SKOS[21] is a W3C Recommendation that defines a vocabulary for representing Knowledge Organization Systems (i.e. vocabularies) and relationships amongst them. SKOS provides constructs to formalize how concepts are related to each other. These constructs include `skos:exactMatch`, to express that two concepts are equivalent in *most* cases, `skos:closeMatch`, to express an equivalence valid in *some* cases, `skos:narrowMatch` and `skos:broaderMatch`, to express hierarchical relationships between concepts, and `skos:relatedMatch`, to express any other type of relatedness.

The first approach consists in applying these constructs to express all mapping relationships considered by the working group[22].

```
ma:title skos:exactMatch dc:title .
```

Example 1.2. A simple mapping represented in SKOS

[21] `http://www.w3.org/TR/skos-reference/`

[22] The mapping tables are available from `http://www.w3.org/2008/WebVideo/Annotations/drafts/ontology10/WD/mapping_table.html`

The use of these properties has a first implication: it entails that the properties `ma:title` and `dc:title` become instances of `skos:Concept` per definition of the `skos:exactMatch` construct. Second, we use the `skos:Collection` construct to group and list items, enabling the representation of a mapping between a simple property on the one hand, and multiple ones on the other hand. `skos:OrderedCollection` represents an ordered list of properties, enabling a more precise matching if necessary, but complex operations cannot be expressed. For example, the creator property defined in the Media Ontology has a simple value, whereas other vocabularies such as MPEG-7 define people with multiple properties: first name, last name, role, etc. SKOS cannot be used to represent that these values must be aggregated and concatenated to be used as value in the Media Ontology.

```
ma:frameSize skos:closeMatch [
  skos:Collection [
        skos:member ebucore:formatHeigth , ebu:formatWidth
    ] ;
 ] .
```

Example 1.3. A complex mapping represented in SKOS

Benefits of this approach:

- Scalability: new properties can be added to the mapping list;
- Fuzziness: mappings are created between properties that are more loosely related than a strict equivalence, which is often the case across schemas designed for specific applications.

Drawbacks:

- Assume that schemas and ontologies to be aligned have been formalized in RDF;
- Inference possibilities are limited;
- No formal complex rule can be attached to this representation.

4.2 Expressing Mappings in OWL and SWRL

Another approach consists in using a more expressive knowledge representation language to express direct mappings between the Media Ontology and other standards. The authors in [11] propose to use OWL and SWRL constructs as shown in the example 1.4 for defining a formal semantic equivalence between the title property defined in EBUCore and Dublin Core and in the Media Ontology.

Additionally, logical rules can be employed to do any type of conversion (including syntactic ones) and transformation of values (e.g., convert bps to kbps). Example 1.5 expresses in SWRL [5] that the value of `ma:frameSize` property

```
ma:title owl:equivalentProperty dc:title .
```

Example 1.4. A simple mapping represented in OWL.

```
[r1: (?res rdf:type ebu:ResourceManifestation)
     (?res ebu:width ?width)     (?res ebu:height ?height)
     (?width ebu:unit "pixels")  (?width  ebu:value ?w1)
     (?height ebu:unit "pixels") (?height ebu:value ?h1)
     -> (?size1 rdf:type ma:Size)
        (?size1 ma:width ?w1) (?size1 ma:height ?h1)
        (?res ma:size ?size1)]
```

Example 1.5. A complex mapping represented in SWRL.

can be filled from the values of the `ebu:width` and `ebu:height` properties. Benefits of this approach:

– Scalability: new properties can be added to the mapping list;
– Formalization: all sort of mappings can be formally represented, including complex ones, allowing inferences to be performed.

Drawbacks:

– Not all metadata standards have formal representations. Sometimes, there are even multiple formalizations of the same standard (e.g. MPEG-7 [12]);
– Complexity: the use of OWL constructs and complex rules can yield in undecidable reasoning.

4.3 Expressing Mappings Using a Format Independent Ontology

An alternative approach is to mediate the mappings through a pivot ontology. The following proposal extends an approach for mapping metadata elements between different stages of the production process of audiovisual media. Different metadata formats and standards are used in the workflow, containing metadata elements with similar and partly overlapping semantics, though not fully identical. In the context of the 2020 3D Media project[23], it has been attempted to model the metadata elements used throughout the production process in a format independent way by creating an ontology that models these elements and the relationships between them [4]. Modeling is done at a meta level, considering grouping and definition relations between the elements. The work considers three problems: (i) verify whether a given metadata element is defined by another given metadata element, (ii) find all metadata elements that are defined by a given metadata element and (iii) find all metadata elements that define a given metadata element. A demo application that addresses the first of these problems for a small set of production metadata items is available at `http://meon.joanneum.at`.

[23] `http://www.20203dmedia.eu`

OWL-DL is used to formally capture the semantics of the metadata elements and their relations. The ontology is format independent and contains the classes `Concept`, with subclasses `AtomicConcept` and `CompoundConcept`. Specific metadata properties are instances of these concepts. The relation `contains` exists between `CompoundConcept` and a set of concepts, the relation `defines` between concepts (bidirectional `defines` relations express identity of concepts). Additionally logical rules are used to infer implicit knowledge about relations between metadata elements. The existing implementation ignores specific data types of the metadata properties.

This approach can be extended for expressing mappings between multimedia metadata schemas and the Media Ontology. In addition to the schema independent ontology, schema specific ones are created for each standards following the same pattern. A new relation type is introduced, which relates concepts between the two ontologies. The relation is modeled as a class, that has properties for qualifying the relation (similar to the SKOS properties) and mapping instructions for data format conversion. The classes representing concepts in the schema specific ontology can be extended to carry additional information needed for mapping, e.g. XPath or binary key of the metadata element. The same rules can be used in both the generic and schema specific ontology for inference.

Figure 1 shows a schematic example for aligning some properties from EBU Core to the Media Ontology. The generic `meon` ontology represents the set of concepts, in that case `title`, `resolution`, `lines` and `columns`. It also models their relations, i.e. the compound of `lines` and `columns` is equivalent to the `resolution`. Relations are introduced to link concepts from the different ontologies. Hence, both `dc:title` and `ma:title` are completely aligned with `meon:mainTitle`. The value for these three properties being a literal, the mapping instruction is the identity function operating on simple datatypes.

The example of the frame size is more interesting. `ebu:formatWidth` (resp. `ebu:formatHeight`) is identical to `meon:colums` (resp. `meon:lines`) with potentially the help of a conversion of the number format. `ma:frameSize` is also equivalent to `meon:resolution`, again with a possible conversion of the format (which is specified by a function name in the relation). Using rules, we can infer from the relations within the `meon` ontology and between the ontologies that `ma:frameSize` defines both `ebu:formatWidth` and `ebu:formatHeight`, but not vice versa. In addition, because of modeling resolution as a compound concept in `meon`, we can also infer that `ebu:formatWidth` and `ebu:formatHeight` *together* define `ma:frameSize`. From the relations along the path between the elements we can collect the format mapping instructions to obtain a chain of functions that maps data types from EBU Core to the Media Ontology. These instructions are applied to the instances of the concepts encountered in the input document.

Benefits of this approach:

- Clean separation between generic concepts and schema specific concepts;
- Formal representation of the semantics of the properties in one format, which can e.g. also be used for validation;
- Inference is used to generate implicit relations and compound concepts.

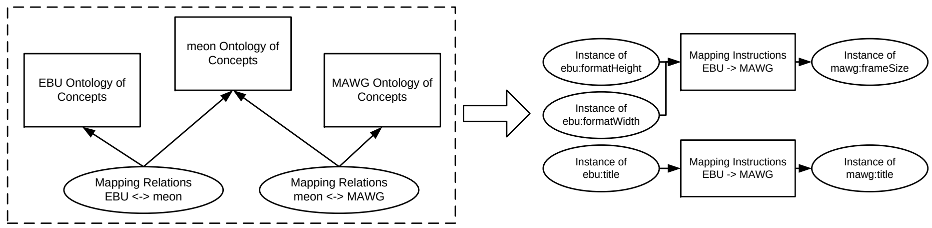

Fig. 1. Mapping using format independent ontology.

Drawbacks:

- Requires building ontology of properties for each schema, which may not be trivial;
- Scalability might be an issue with hundreds of concepts;
- Data type conversions might need built-in functions in the rule engine or external code to be executed.

4.4 Expressing Mappings with Built-in Properties

We present finally an alternative to the approach presented in the Section 4.3. The mappings are still mediated through a pivot ontology, but this ontology is directly related to the Media Ontology. This pivot ontology can be described as followed. Instances of the `MAWGMetadataProperty` class are described by the core set of annotation properties of the Media Ontology, while instances of the `StandardMetadataProperty` class are described by annotation properties of multimedia metadata schemas to be mapped. The `MetadataProperty` class is a superclass of these two classes. The `MetadataPropertyRelation` class characterizes the nature of the mapping relationship. It provides further information such as the transformation rule to operate on the values, the type of the mapping (e.g. exact) or whether it is a compound relationship or not. A priority operator can also be defined, in case various metadata properties from various standards can be aligned to a particular annotation property from the Media Ontology. This operator aims at defining a priority hierarchy for implementing a *SET* functionality in a API built on top of the Media Ontology.

The examples 1.6 and 1.7 illustrate this approach for the `ma:title` and `ma:frameSize` properties. Benefits of this approach:

- No specific representation format (e.g., OWL) of metadata standards is needed.

Drawbacks:

- No distinction between different versions of metadata formats. This issue could produce inconsistencies;
- No inference (e.g. between properties) is possible;

```
:MAWGMetadataProperty_21 a :MAWGMetadataProperty ;
  rdfs:isDefinedBy ma:title ;
  skos:inScheme <http://www.w3.org/2009/09/mediaont#> ;
  :hasMetadataPropertyRelation [
    :isCompositeRelation false ;
    :relationSemantic "exact" ;
    :hasStandardMetadataProperty [
      skos:inScheme <http://purl.org/dc/elements/1.1/>;
      rdfs:isDefinedBy dc:title ] ] .
```

Example 1.6. A simple mapping.

```
:MAWGMetadataProperty_10 a :MAWGMetadataProperty ;
  rdfs:isDefinedBy ma:frameSize ;
  skos:inScheme <http://www.w3.org/2009/09/mediaont#> ;
  :hasMetadataPropertyRelation [
    :isCompositeRelation true ;
    :relationSemantic "exact" ;
    :hasStandardMetadataProperty [
      skos:inScheme <http://www.ebu.ch/metadata/ontologies/>;
      rdfs:isDefinedBy [ owl:unionOf (
        [ a owl:Restriction ; owl:onProperty ebu:formatWidth ;
             owl:allValuesFrom xsd:int]
        [ a owl:Restriction ; owl:onProperty ebu:formatHeigth ;
             owl:allValuesFrom xsd:int] ) ] ; ] ; ] .
```

Example 1.7. A complex mapping.

5 Conclusion and Future Work

This paper addresses the interoperability issue between multimedia metadata formats. The related work described in section 2 and the numerous use cases summarized in [7] show that there is a need for solving this issue. We have presented a core set of annotation properties defined in the Media Ontology developed by the W3C Media Annotations Working Group. Furthermore, we have discussed how mapping relationships between this core set of annotation properties and the multimedia metadata standards can be represented, either directly using semantic web languages (SKOS, OWL, or the forthcoming RIF[24] recommendations) or through a pivot ontology.

Each approach presents benefits and drawbacks that can be grouped in the following criteria: *complexity*, *scalability* and *reasoning capabilities*. The listing of these benefits and drawbacks is currently done ad-hoc. As such, future work consists of an in-depth evaluation in which each of the criteria is measured for the different approaches. Expressing direct mappings is intuitive and provide scalability. However, it requires that the metadata formats to be aligned have been formally represented in SKOS, RDFS or OWL. The use of a pivot ontology tends to be a more generic solution which has the price of complexity in terms of the number of triples generated.

[24] http://www.w3.org/TR/rif-core/

Future work deals primarily with the recommendation of the Media Ontology. Its coverage is still evolving and profiles might be introduced, in particular, for offering a degree of variability in the way mappings with other standards is formalized. Another important milestone planned is the design of an API on top of the Media Ontology. The main purpose of this API will be the implementation of appropriate *GET* and *SET* functionalities. One of the open issues concerns the implementation procedure to follow in case of collision between various semantic mappings. The priority operator introduced in the Section 4.4 is a useful contribution with this respect.

Acknowledgments

The authors would like to thank all the participants of the *W3C Media Annotations Working Group* for their willingness to discuss the core ontology, the mappings between metadata standards, and their implementation. This work has been partially supported by the THESEUS Program, which is funded by the German Federal Ministry of Economics and Technology, and under the 7th Framework Programme of the European Union within the ICT project "Presto-PRIME" (FP7 231161).

References

1. Isabel Cruz, Huiyong Xiao, and Feihong Hsu. An Ontology-Based Framework for XML Semantic Integration. In *International Database Engineering and Applications Symposium*, pages 217–226, Coimbra, Portugal, 2004.
2. Mario Döller, Florian Stegmaier, Harald Kosch, Ruben Tous, and Jaime Delgado. Standardized Interoperable Image Retrieval. In *ACM Symposium on Applied Computing (SAC), Track on Advances in Spatial and Image-based Information Systems (ASIIS)*, 2010, to appear.
3. Mario Döller, Ruben Tous, Matthias Gruhne, Kyoungro Yoon, Masanori Sano, and Ian Burnett. The MPEG Query Format: On the way to unify the access to Multimedia Retrieval Systems. *IEEE Multimedia*, 15(4):82–95, 2008.
4. Martin Höffernig and Werner Bailer. Formal Metadata Semantics for Interoperability in the Audiovisual Media Production Process. In *Workshop on Semantic Multimedia Database Technologies (SeMuDaTe 2009), co-located with the 4th International Conference on Semantic and Digital Media Technologies (SAMT2009)*, Graz, Austria, 2009, to appear.
5. Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, and Mike Dean. SWRL: A Semantic Web Rule Language Combining OWL and RuleML. W3C Member Submission, 2004. `http://www.w3.org/Submission/SWRL/`.
6. Dublin Core Metadata Initiative. Dublin core metadata element set - version 1.1: Reference description. `http://dublincore.org/documents/dces/`, 2008.
7. WonSuk Lee, Tobias Bürger, Felix Sasaki, and Véronique 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/`.

8. WonSuk Lee, Tobias Bürger, Felix Sasaki, Véronique Malaisé, Florian Stegmaier, and Joakim Söderberg. Ontology for Media Resource 1.0. W3C Working Draft, 2009. `http://www.w3.org/TR/mediaont-10/`.
9. MPEG-7. Multimedia Content Description Interface. ISO/IEC 15938, 2001.
10. National Electrical Manufacturers Association. Digital Imaging and Communications in Medicine (DICOM). `ftp://medical.nema.org/medical/dicom/2008/`, 2008.
11. Chris Poppe, Gaëtan 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.
12. Raphaël Troncy, Óscar Celma, Suzanne Little, Roberto García, and Chrisa Tsinaraki. MPEG-7 based Multimedia Ontologies: Interoperability Support or Interoperability Issue? In *1st International Workshop on Multimedia Annotation and Retrieval enabled by Shared Ontologies*, 2007.
13. Guangming Xing, Zhonghang Xia, and Andrew Ernest. Building automatic mapping between XML documents using approximate tree matching. In *ACM Symposium on Applied Computing (SAC)*, pages 525–526, Seoul, Korea, 2007.
14. Xia Yang, MongLi Lee, Tok Wang Ling, Lee Tok, and Wang Ling. Resolving Structural Conflicts in the Integration of XML Schemas: A Semantic Approach. In *International Conference on Conceptual Modeling (ER)*, pages 520–533, Chicago, USA, 2003.