International Journal of Semantic Computing

A Method for Integrating Multimedia Metadata Standards and Metadata Formats with the Multimedia Metadata Ontology

Ansgar Scherp, Daniel Eißing, and Carsten Saathoff

University of Koblenz-Landau, Koblenz, Germany
scherp,eissing,saathoff@uni-koblenz.de
http://www.uni-koblenz-landau.de/koblenz

Accepted 15 August 2011

Summary: The M3O abstracts from the existing metadata standards and formats and provides generic modeling solutions for annotations, decompositions, and provenance of metadata. Being a generic modeling framework, the M3O aims at integrating the existing metadata standards and metadata formats rather than replacing them. This is in particular useful as today's multimedia applications often need to combine and use more than one existing metadata standard or metadata format at the same time. However, applying and specializing the abstract and powerful M3O modeling framework in concrete application domains and integrating it with existing metadata formats and metadata standards is not always straightforward. Thus, we have developed a step-by-step alignment method that describes how to integrate existing multimedia metadata standards and metadata formats with the M3O in order to use them in a concrete application. We demonstrate our alignment method by integrating seven different existing metadata standards and metadata formats with the M3O and describe the experiences made during the integration process.

Keywords: Ontology alignment; multimedia metadata; metadata standards; metadata formats

1. Introduction

Looking at the existing metadata models like [1–5] and metadata standards such as [6–11], we find it hard to decide which of them to use in a complex multimedia application. They focus on different media types, are very generic or designed for a specific application domain, overlap in the functionality provided, are semantically ambiguous, or provide only limited features for modeling the multimedia metadata. In addition, the existing standards and formats typically cannot be combined with each other as they are not designed for such a combined use. However, building a complex multimedia application often requires using several of these standards and formats *together*, e.g., when different tools have to be integrated along the media production process [12]. Existing approaches to enable metadata interoperability like XMP [7], the Metadata Working Group [13], and the W3C Media Annotations Working Group [14] focus on single media assets. Thus, they do not consider structured multimedia content as it can be created with formats like Flash [15], SMIL [16], and SVG [17]. In addition, the existing approaches do not provide a generic framework suitable for integrating arbitrary metadata formats and arbitrary metadata standards.

To solve this problem, we have developed the pattern-based Multimedia Metadata Ontology (M3O) [18]. The M3O is a sophisticated modeling framework for representing among others the annotation, decomposition, and provenance of multimedia content and multimedia metadata. The goal of the M3O is to provide a framework for the integration of existing metadata formats and metadata standards rather than replacing them. Being an abstract and powerful representation model for multimedia metadata, it is not straightforward how to use the M3O in concrete multimedia applications and integrating and aligning it with existing metadata standards and metadata formats. In this paper, we fill this gap and present a step-by-step alignment method describing how to integrate existing formats and standards for multimedia metadata and the M3O.

As an example for applying the M3O, let us consider an audio object money-abba-1 and annotate it with the date when it has been recorded and the band that actually has produced the recording. For conducting such annotations, we are using the Annotation Pattern provided by the M3O. For representing the recording date and the band, we use a M3O-aligned ontology of the widely adopted ID3 [19] metadata specification for describing audio files.

The audio object we want to annotate with its information when it has been originally recorded and by whom is represented by the individual money-abba-1 as depicted in Figure 1. The individual is an instance of the AudioConcept and is defined in the ID3 ontology. In the context of the M3O Annotation Pattern, the audio object money-abba-1 is classified by the individual ac-1 using the classifies relation as the individual that is being annotated (AnnotatedConcept). The AnnotationPattern id3ap-1 uses the defines relation to determine the ID3 compliant annotations to be used, namely the OriginalRecordingConcept with its two hasPart relations, the ArtistConcept and ReleaseYearConcept. These concepts defined by the AnnotationPattern provide the structure of the annotation. A Quality represents an attribute of an entity and connects the entity with a concrete Region. The Regions are the value spaces of the Quality. In the example, the object money-abba-1 has two qualities that connect the object with the two region concepts ArtistRegion ar-1 and ReleaseYearRegion ryr-1. To associate the regions with the annotation structure defined by the AnnotationPattern, they are parameterized by the corresponding ArtistConcept and ReleaseYearConcept. The actual annotations of our audio object money-abba-1 are attached to the region concepts ArtistRegion and ReleaseYearRegion using the hasRegionDataValue relation. In our example, the audio object money-abba-1 has been originally recorded by the pop group "ABBA" in "1976". The actual annotations of the audio object money-abba-1 by the Regions ar-1 and ryr-1 are embedded by using the hasSetting relation into what is called the concrete situation or application context of the Annotation Pattern, represented by the AnnotationSituation id3as-1. As the AnnotationSituation id3as-1 is a valid annotation with respect to the annotation structure defined by the AnnotationPattern id3ap-1, we say the AnnotationSituation satisfies the AnnotationPattern.

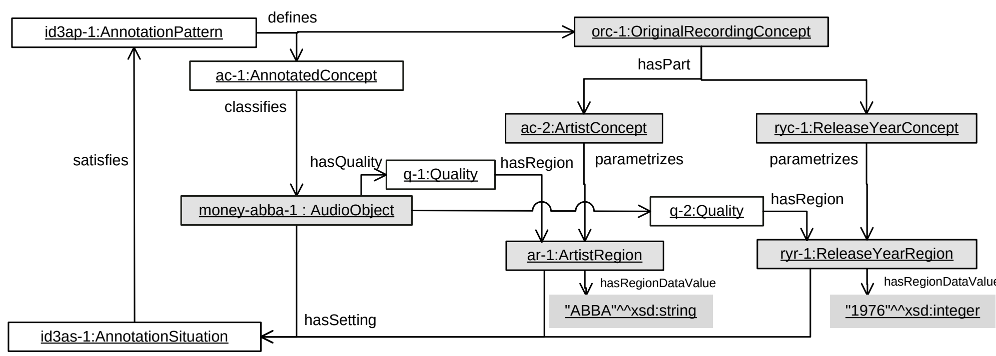


Figure 1: Application of the ID3 metadata format using M3O.

A systematic introduction to the concepts and patterns of the M3O is provided in the subsequent Section 2. In Section 3, we describe the step-by-step alignment method along the tasks that have to be performed for integrating existing metadata standards and metadata formats with the M3O. Subsequently, we demonstrate the application of the alignment method in Sections 4 to 10 at the examples of seven metadata standards and metadata formats. These are the Core Ontology on Multimedia (COMM) [2], EXIF [6], XMP [7], ID3 [19], Dublin Core [10, 11], Yahoo!'s SearchMonkey Media [20], and the Ontology for Media Resource [21]. Section 11 presents two larger examples of the combined use of several metadata standards and metadata formats aligned with the M3O. We present related work in Section 12, before we conclude the paper.

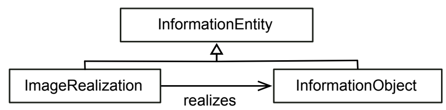
2. Introduction to the Multimedia Metadata Ontology

The Multimedia Metadata Ontology (M3O) [18] provides a generic modeling framework for representing multimedia metadata. It has been designed by abstracting from existing multimedia metadata standards and metadata formats (see related work in Section 12). Such a generic model is not limited to a single media type such as images, video, text, and audio. It also provides support for structured multimedia content, i.e., a combination of several media assets of different media types like video and text coherently arranged in time and space. Such structured multimedia content can be created with today's multimedia presentation formats like SMIL, SVG, and Flash. The M3O is modeled as a highly axiomatized core ontology that is based on the foundational ontology DOLCE+DnS Ultralight (DUL) [22]. The

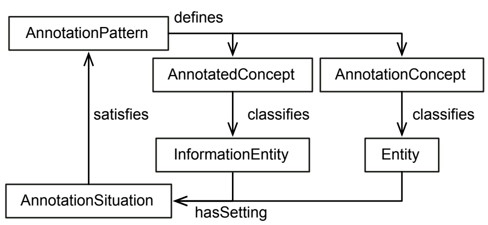
M3O follows a pattern-based approach to ontology design. Ontology design patterns provide generic modeling solutions for recurring modeling problems [23] and thus are similar to design patterns in software engineering [24]. Each ontology design pattern is focused on modeling a specific and clearly identified aspect of the domain such as annotation and decomposition in the M3O (cf. competency question for ontology design patterns in [25]). We make use of the ontology design patterns provided by DUL such as the Descriptions and Situations (DnS) Pattern. The DnS Pattern provides an ontological formalization of context through the introduction of roles [26,27]. By this, it allows to formally represent different views onto entities.

The **Information Realization Pattern** in Figure 2a models the distinction between information objects and information realizations [22]. Consider a digital image that is stored on the hard disk in several formats and resolutions. An information object represents the image as an abstract *concept* or *idea*, namely the information that an image has been taken. This information object might be realized by different information realizations, i.e., different files may exist that *realize* the same information object. As shown in Figure 2a, the pattern consists of the InformationRealization that is connected to the InformationObject by the realizes relation. Both are subconcepts of InformationEntity, which allows treating information in a general sense as we will see in the Annotation Pattern.

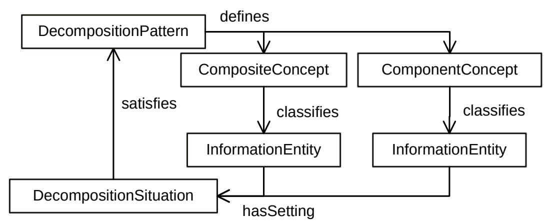
Annotations are understood in the M3O as the attachment of metadata to an information entity. Metadata comes in various forms such as low-level descriptors obtained by automatic methods, non-visual information covering authorship or technical details, or semantic annotation aiming at a formal and machine-understandable representation of the contents. The **Annotation Pattern** models the basic structure that underlies all types of annotation. This allows for assigning arbitrary annotations to information entities while providing the means for modeling provenance and context. In Figure 2b, we see that an annotation is not modeled as a direct relationship between some media item and an annotation. It is defined by a more complex structure, which is inherited by the Descriptions and Situations Pattern of DUL. Basically, a Descriptions and Situations Pattern is two-layered. The *Description* defines the structure, in this case the structure of an annotation, which contains some entity that is annotated and some entity that represents the metadata. The *Situation* contains the concrete entities for which we want to express the annotation. The pattern allows us to add further concepts and entities into the *context* of an annotation, e.g., expressing provenance or confidence information. At the top half of Figure 2b, we see that the AnnotationPattern defines an AnnotatedConcept and an AnnotationConcept. The AnnotatedConcept classifies an InformationEntity and thus expresses that the information entity is the subject of the annotation. The AnnotationConcept classifies some Entity, which identifies this entity as the annotation or metadata. The entity can be some complex data value, e.g., representing some low-level features represented using the Data Value Pattern, but also some concept located in a domain ontology such as DBpedia [28]. All the entities have as setting the AnnotationSituation, which satisfies the AnnotationPattern.



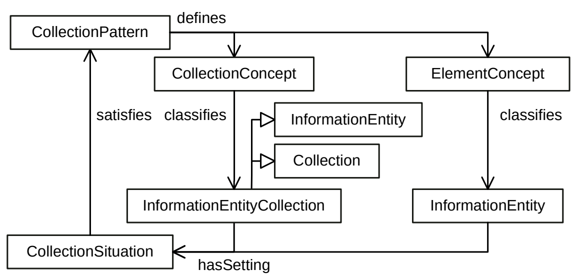
(a) Information Realization Pattern



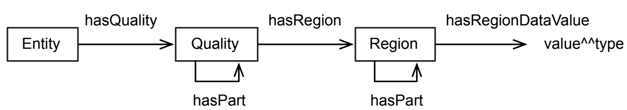
(b) Annotation Pattern



(c) Decomposition Pattern



(d) Collection Pattern



(e) Data Value Pattern

Figure 2: Patterns of the Multimedia Metadata Ontology

The **Decomposition Pattern** shown in Figure 2c is like the Annotation Pattern a specialization of DUL's Descriptions and Situations Pattern. It allows to decompose (multi-)media assets into parts in order to annotate them further using the Annotation Pattern. In the context of the Decomposition Pattern, the (multi-)media asset that is decomposed is called the composite and the parts are the components. The Decomposition Pattern introduces a DecompositionPattern as specialization of DUL's Description concept and a DecompositionSituation as specialization of the Situation concept. It defines exactly one CompositeConcept and at least one ComponentConcept. The CompositeConcept and ComponentConcept classify InformationEntity that have the DecompositionSituation as setting.

A set of information entities that have one or more common properties can be represented as a collection of information entities using the **Collection Pattern** of the M3O. As depicted in Figure 2d, the Collection Pattern is also based on the Descriptions and Situations Pattern of DUL. By using the Descriptions and Situations Pattern, the Collection Pattern supports the collaborative creation of collections. In addition, it is aware of the source of the information entities that are added. An example of a collection is the set of images collected by different people under a common topic like *antique cars* or *cute dogs* as it is possible by, e.g., the use of groups on the photo sharing platform Flickr[a]. The design of the Collection Pattern as shown in Figure 2d introduces a CollectionPattern as specialization of the Description concept and provides a CollectionSituation as specialization of the Situation concept of DUL's Descriptions and Situations pattern. As axiomatization, the concept of the CollectionPattern defines that there has to be exactly one CollectionConcept that classifies an InformationEntityCollection. The InformationEntityCollection is the actual collection of information entities. Thus, it is specialized from DUL's InformationEntity and Collection concepts. In addition, the CollectionPattern contains an axiom that defines that there are some ElementConcepts that classify InformationEntity, i.e., the elements of the collection.

To represent concrete data values such as strings and numerical values in DUL, there exists the concept Quality. A Quality represents an attribute of an Entity, i.e., that only exist together with the Entity. Regions are used to represent the values of an Quality and the data space they come from. The **Data Value Pattern** depicted in Figure 2e assigns a concrete data value to an attribute of that entity. The attribute is represented by the concept Quality and is connected to the Entity by the hasQuality property. The Quality is connected to a Region by the hasRegion relation. The Region models the data space the value comes from. We attach the concrete value to the Region using the relation hasRegionDataValue.

3. Alignment Method

This section introduces our method for aligning existing multimedia metadata standards and multimedia metadata formats with the M3O. The method has been derived from our experiences applying and specializing the M3O for seven existing multimedia formats and standards, namely COMM [2], EXIF [6], XMP [7], ID3 [19], Dublin Core [10, 11], Yahoo!'s SearchMonkey Media [20], and the Ontology for Media Resource [21]. In contrast to automatic, adaptive, or machine learning approaches for ontology alignment [29–31], we conduct a pure manual alignment, as only a manual alignment ensures the high quality of the integration and minimizes ambiguities and imprecise matching. We consider the time and effort for manual alignment manageable, as we assume that each metadata format or standard has to be aligned only once and that updates to the integrated formats or standards will be infrequent and mostly incremental.

We propose an iterative alignment method that helps ontology engineers to integrate existing metadata standards and metadata formats. In each iteration, we

[a] `http://www.flickr.com/`

consecutively evolve the alignment of the standard or format with the M3O. Following an iterative approach, we are able to identify, analyze, and flexibly react to problems and challenges encountered during previous iterations of the alignment. Each iteration consists of four steps: The first step targets at the understanding of the format or standard to be integrated. The second step reorders concepts and properties into coherent groups. The third step maps the concepts and properties and their arrangement in groups with the M3O. The fourth step proves and documents the validity of the alignment. It finalizes the iteration.

To introduce our alignment method, we proceed as follows: For each step, we first outline the goals and provide a brief summarization. Subsequently, we describe the core tasks to be performed within the step and provide concrete examples that show its relevance and application for concrete ontologies.

3.1. *Step 1: Understanding*

Summary A precise understanding of the metadata standard or metadata format to be integrated is an import prerequisite for aligning it with the M3O. Consequently, the first step of alignment is an in-depth analysis of the structure and core concepts and properties of the model at hand. While this advise may seem obvious, this is a task easily underestimated and problems neglected at an early stage can cause time-consuming problems along the integration process. Additional documentation, if available, will help to (re-)produce the overall structure not explicitly expressed in the formal specification.

Detailed Description and Examples In general, we have found three distinct modeling approaches to be very common for multimedia metadata standards and metadata formats.

Predicate-centric In a predicate-centric approach as followed, e.g., by the Ontology for Media Resource, the ontology is predominantly specified through a set of properties. Such a model offers very little structure in a machine readable format, e.g., in terms of conceptual relations between properties. However, by analyzing the documentation, we infer additional information about intended groupings of properties and the structural composition of the format or standard to be integrated.

Pattern-based Pattern-based ontologies, e.g., COMM, provide a high degree of axiomatization and structure in a formal and precise way. Through our analysis, we understand the patterns used and the functionality they provide. This allows us to compare the patterns of the ontology to be integrated with those provided by the M3O.

Legacy Models Other standards and formats have not yet been semantified. By analyzing the concepts and relations expressed in the specification of the format, we decide how the core concepts can be expressed in a formal and precise way using the M3O.

Ambiguities that are found during the initial analysis are discussed at this point. It is not our intention to revise all modeling decisions made for the ontology to be integrated. However, we consider the alignment a good opportunity to correct some of the *bad smells* [32] discovered. Once we have reached a sufficient understanding of the format or standard to be integrated, we proceed with the grouping step.

3.2. *Step 2: Grouping*

Summary Ontologies should provide structural information on the relations and groupings of concepts it defines. However, although many standards and formats provide this information in their documentation, the information is sometimes lost when the models are transformed to an ontology. By using the original specifications and documentations, we are able to preserve and recreate this information grouping, and provide them through formal specification in the aligned ontology.

Detailed Description and Examples In principle, we distinguish three forms of available grouping information:

Explicit Grouping With pattern-based models, we find an explicit grouping of concepts into coherent patterns. The patterns are often accompanied by a rich axiomatization on how the concepts in these patterns relate. As an example, the definition of a color histogram annotation in COMM specifies a ColorQuantizationComponentDescriptorParameter that groups the concepts ColorComponent and NumberOfBinsPerComponent.

Implicit Grouping For other metadata models grouping information may not be explicitly represented. This is often the case with predicate-centric approaches, e.g., the Ontology for Media Resource. In these cases, we refer to the textual documentation in order to (re-)construct the implicit groupings of the properties or concepts. As an example, the documentation of the Ontology for Media Resource offers a textual description on the grouping of its properties, e.g., in terms of identification or creation. However, this information is not accessible in the Resource Description Framework (RDF) representation [21] as proposed by the W3C. By defining the appropriate axioms, we have appended the implicit grouping information in a formal and explicit way, e.g., by stating that an IdentificationAnnotation hasPart some TitleAnnotation, LanguageAnnotation, and LocatorAnnotation.

Recovery of Groupings In other cases grouping information is lost when transferring metadata formats or standards to RDF. For example the EXIF metadata standard provides textual descriptions about groupings, e.g., in terms of pixel composition and geo location. However, this distinction got lost in the adaption of EXIF to an RDF vocabulary [33].

Once we have provided all relevant grouping relations through a formal specification, we continue with the mapping step.

3.3. *Step 3: Mapping*

Summary This step achieves the mapping of the ontology's concepts and structure to the modeling framework of the M3O. The goal of this step is to create a working ontology, which, after validation, can be published on the Internet or used as basis for further iterations.

Detailed Description and Examples For the alignment, we follow a sequence of the following three steps:

1. **Mapping of Concepts** If some superclass of the concept to be aligned is present in both ontologies, direct mapping of concepts is feasible. This is mainly the case for ontologies that share the same foundation, e.g., COMM and the M3O, which both base on the DUL foundational ontology. All axioms of the aligned concepts are preserved as long as they are applicable through the M3O. If a concept is not applicable in the M3O, we align all dependent subclasses and references to the nearest matching M3O concept. As an example, the COMM DigitalData concept, which is a subclass of the DUL InformationObject, has been removed during the alignment. The dangling dependencies and references have been resolved by subclassing or referencing the InformationObject instead. In other cases, an alignment could be directly conducted by introducing new subclasses based on the description of the metadata format. For example, the Yahoo!'s SearchMonkey Media vocabulary provides media types like article, audio, image, photo, text, and video. They have been aligned with DUL by introducing subclasses of InformationRealization for each media type, namely ArticleRealization, AudioRealization, ImageRealization, PhotoRealization, TextRealization, and VideoRealization.
2. **Structural Mapping** For structural mapping, we consider the functionality of the concept and property groupings as obtained from the second alignment step. The concept and property groupings can be provided, e.g., in form of ontology design patterns like in the case of COMM, as set of re-organized concepts and properties as in the example of the Ontology for Media Resource, or as set of newly created concepts and properties like for EXIF and ID3. If a grouping of concepts and properties, e.g., an ontology design pattern, offers the same or an equal functionality than a pattern of the M3O, we replace it with the M3O pattern. By adapting the M3O pattern, we are often able to express the same functionality using a more generic approach. As an example, COMM proposes the Digital Data Pattern to express data values in a digital domain. A similar functionality is provided by the M3O Data Value Pattern, which expresses data values through the generic concepts of Quality and Region. The COMM Digital Data Pattern can be considered as special case of expressing data values and therefor has been replaced by using the M3O Data Value Pattern instead.

In the same way, we simplify the structural composition of the existing model by merging multiple concepts and patterns that offer the same or an equal functionality. As an example, COMM defines three annotation patterns. Each deals with a different aspect of multimedia annotation, although they vary only slightly in their structural composition. We have aligned those patterns by adapting the M3O Annotation Pattern. The domain specific concepts that result from the separation into three distinct patterns have been preserved by subclassing the corresponding concepts of the M3O Annotation Pattern. This simplifies the structure of the model, while also preserving the original functionality.

3. **Removing Unnecessary Concepts** We finalize the mapping step by cleaning up unused dependencies from the ontology. Concepts that either have no further relevance for the target context or are sufficiently covered by the M3O are removed at this point. An example, the COMM AnnotatedMediaRole offers an equal functionality as the M3O AnnotatedInformationRealizationConcept. We therefore have removed COMM's AnnotatedMediaRole and replaced any formal relation that involves the concept.

3.4. *Step 4: Validation and Documentation*

In each iteration of the alignment process, we need to check the consistency of the resulting ontology. This can be done by using a reasoner like Fact++[b] or Pellet[c]. Any problem encountered during the alignment can be resolved by reiterating the four steps of the alignment method. After proving the consistency of the resulting ontology, we finalize the process by documenting all major decisions and adjustments made during the alignment.

3.5. *Summary*

In this section, we have presented a four-step alignment method for integrating existing multimedia metadata standards and multimedia metadata formats with the M3O. In the following Sections 4 to 10, we demonstrate the application of our alignment method at the examples of seven existing standards and formats for multimedia metadata.

4. Example 1: Core Ontology on Multimedia (COMM)

The Core Ontology on Multimedia (COMM) [2] is a formal specification of the MPEG-7 metadata standard [9]. In contrast to other approaches to modeling MPEG-7 as an ontology [1, 34], COMM is not designed as a one-to-one mapping, but provides a set of patterns that cover the core and repetitive building blocks

[b]`http://owl.man.ac.uk/factplusplus/`
[c]`http://clarkparsia.com/pellet/`

of MPEG-7. The central challenge of the alignment of COMM and M3O is understanding the patterns of COMM and mapping them to the modeling framework provided by the M3O.

Understanding COMM follows a pattern-based approach to ontology design and builds on the DUL foundational ontology. Some of the core patterns, i.e., the Descriptions and Situations Pattern, are shared between COMM and the M3O. Others, e.g., the Digital Data Pattern, form major structural differences.

COMM defines five structural patterns, namely the Content Annotation Pattern, Media Annotation Pattern, and Semantic Annotation Pattern for media annotation, the Decomposition Pattern for media (de-)composition, and the Digital Data Pattern, which expresses annotations in a digital domain. Domain specific knowledge is separated from the core concepts and defined in separate ontologies, e.g., concepts concerning annotation of visual entities are defined in the *visual ontology*.

Some ambiguities that were found in the initial analysis have been resolved at this point. As an example, COMM specifies concepts such as NumberOfBinsPerComponent that are specialization of both Parameter and Region. While this may not be syntactically incorrect, it violates the DnS pattern of DUL. In the DnS pattern, a Parameter parametrizes a Region. Thus, these two concepts should not have common sub-concepts. To solve this problem, we have removed the superclass relations to the Parameter concept and introduced a parametrizes relation. For example, COMM specified a ColorComponent and NumberOfBinsPerComponent, which are subclasses of both the ColorQuantizationComponentDescriptorParameter and the Region concept. We have removed the superclass relation from the ColorComponent and NumberOfBinsPerComponent to the ColorQuantizationComponentDescriptorParameter, which instead now parametrizes these concepts.

Grouping Following a pattern-based design, COMM already provides a rich degree of conceptual groupings and their axiomatization in a machine readable format. However, reusability can be improved by redistributing concepts among the six ontologies of COMM, *core*, *datatype*, *localization*, *media*, *visual*, and *textual*, respectively. As an example, the concept RootSegmentRole, located in the COMM *core ontology*, is not used in any pattern definition and has therefore been relocated to the *localization ontology*.

Mapping The main challenge of aligning COMM and the M3O concern the differences of the patterns used and how to relate them. Although some principles are shared between the ontologies, there are also major differences, e.g., the Digital Data Pattern of COMM and the Information Realization Pattern of the M3O.

Often COMM patterns have been replaced using a more generic pattern of the M3O. As an example, Figure 3 displays the adaption of the COMM Digital Data Pattern through the M3O. For the alignment, we have decided that the functionality of the Digital Data Pattern, i.e., expressing data values, can be maintained by adopting the M3O Data Value Pattern instead. All related concepts have either been removed or mapped to the next matching M3O concept. As an example, the StructuredDataDescription concept has been removed as it held no further relevance in the context of the Data Value Pattern. The StructuredDataParameter concept on the other hand has been preserved as specialization of the M3O Annotation Pattern. To accommodate StructuredDataParameters with the M3O, we consider StructuredDataParameters as subclass of the AnnotationConcept. Through parameterizing the appropriate Region, we can constrain the range applicable for a specific StructuredDataParameter. The value itself is expressed using the hasRegionDataValue relation. In a similar manner, the three annotation patterns of COMM have been replaced through the M3O Annotation Pattern and all dependent concepts have been mapped to the M3O Annotation Pattern instead.

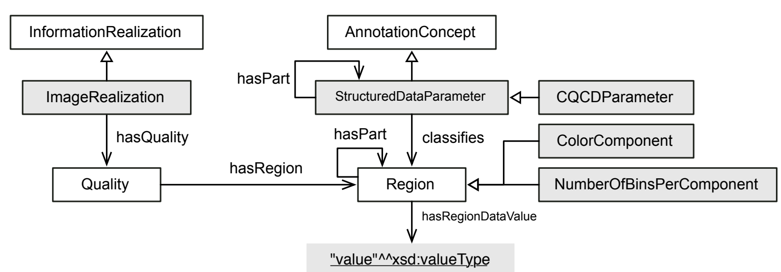


Figure 3: Excerpt of COMM subsequently to the integration with the M3O. White boxes show the concepts of the M3O or DUL, whereas grey boxes represent concepts of COMM aligned to the M3O.

Validation and Documentation The alignment of COMM and the M3O has been validated using Fact++ and Pellet reasoner. Please note that the validation and documentation step is the same for all metadata formats and metadata models. Thus, it is not mentioned again in the remainder of the paper.

Application of the Aligned Ontology Figure 4 demonstrates the application of StructuredDataParameters using COMM aligned with the M3O. We specify a ColorQuantizationComponentDescriptorParameter (CQCDParameter) as part of the RGBHistogramAnnotationConcept. The CQCDParameter parametrizes the ColorComponents and NumberOfBinsPerComponent, which are considered part of the RGBHistogramRegion. The hasRegionDataValue relation expresses the primitive value for this annotation, e.g., an unsigned integer for the NumberOfBinsPerComponent concept. Staying in line with the specification of the M3O Data Value Pattern, we consider the use of StructuredDataParameters optional. Thus, we do not specify that an AnnotationConcept must specify any StructuredDataParameters in a hasPart relation but recommend using them as they add an additional layer of formal expressiveness.

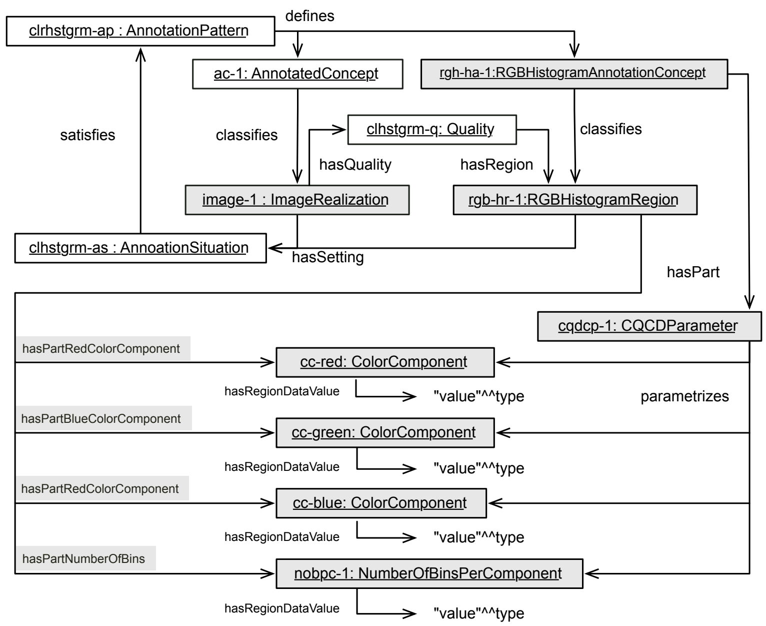


Figure 4: Application of COMM after its integration with the M3O

5. Example 2: EXIF

EXIF is a common metadata standard for images and supports mainly technical metadata [6]. It is embedded directly into media assets such as JPEG files. The following section presents the alignment of EXIF and the M3O.

Understanding The key-value based metadata specified in EXIF is binary encoded into the header of, e.g., JPEG files. Consequently the mapping of the non-semantified concepts onto the modeling framework of the M3O posed the major challenge for this particular alignment. Thus, for aligning EXIF and the M3O, we first needed to semantify the key-value pairs of EXIF.

Grouping The EXIF metadata standard has been translated to a RDF vocabulary [33] by the W3C through an one-to-one mapping. Here, each key of the EXIF specification has been directly mapped to a corresponding property. This approach ignores the groupings of metadata keys that is provided in the original EXIF specification such as pixel composition and geo location. For the alignment, we have reconstructed this grouping information from the EXIF specification.

Mapping When mapping EXIF to the M3O, special consideration has been provided on how to map EXIF properties to information objects and information realizations. For example, locations have been constrained to information objects, as they convey information on where the original picture has been taken. Image resolutions describe a quality of a concrete realization, e.g., a JPEG file, and are therefore associated with the information realization instead. Specific properties can be referred to by using preexisting vocabularies, e.g., the WGS84 vocabulary [35] for GPS information. As EXIF restrains itself to describing qualities of images, all keys have been mapped as specialization of the M3O Annotation Pattern.

A concrete example of applying the EXIF ontology aligned to the M3O can be found in an extended version of this paper published as technical report [36].

6. Example 3: ID3

ID3 is the de-facto metadata standard for describing audio assets [19]. It allows sophisticated descriptions of various features of the audio assets like composer, album, lyrics, copyright, beat-per-minute, volume adjustment, and equalization.

Understanding Like the EXIF standard, metadata in ID3 format is binary encoded and embedded with the audio file it describes. ID3 has a tag header and tag body with one or more frames. Each frame contains one or more groups which actually represent the audio metadata in form of typed key-value pairs. In this paper, we have analyzed and aligned ID3 in version 2.3.

Grouping The ID3 metadata format has been translated into a RDF vocabulary, called the NEPOMUK ID3 Ontology [37], within the context of the the NEPOMUK project. Like the translation of EXIF to a RDF vocabulary (see above), the NEPOMUK ID3 Ontology has been created by a one-to-one mapping of the metadata keys specified in the ID3 format to corresponding ontology properties. This mapping ignores the grouping information provided in the original ID3 documentation like the category concept in ID3 that consists of a *content group description* (e.g., classical music is often sorted in different musical sections [19]), a *title* or *songname*, and *subtitle* or *description refinement*. Other groups of key-value pairs defined in the ID3 format are not mapped to the NEPOMUK ID3 Ontology at all such as the volume adjustment and equalization. For aligning the ID3 metadata format to the M3O, we have reconstructed the grouping information in the resulting ontology and have added the groups of key-value pairs that where not mapped at all.

Mapping Specifically the grouping information that is provided in the ID3 documentation has been considered and appended to the resulting ontology. Apart from that, the mapping of ID3 properties to information objects and information realizations has been considered as with EXIF (see Section 5) and all metadata keys have been mapped as specializations of the M3O Annotation Pattern.

An example of applying the M3O-aligned ID3 ontology is already provided in the introduction Section 1. An additional example is shown in Section 11.2.

7. Example 4: Dublin Core

The Dublin Core Metadata Initiative [10] aims at developing interoperable metadata standards which are applicable in various domains. In this work, we consider the Dublin Core Metadata Element Set [11], which is a set of 15 metadata elements defined in the "classical" Dublin Core.

Understanding The Dublin Core Metadata Initiative provides a RDF version of the simple Dublin Core metadata set [38]. We took the textual description of the Dublin Core Metadata Element Set and the RDF version as input to map the 15 properties such as *creator*, *date*, or *format* to the M3O.

Grouping The textual description of the metadata elements defined in the Dublin Core Metadata Element Set does not contain any grouping information. Given the small number of metadata elements and the heterogeneity of the metadata elements, no additional grouping has been introduced.

Mapping Like with EXIF and ID3, we have considered for each metadata element whether it is applicable to information objects or information realizations. For example, the metadata element *format* is applicable to information realizations only whereas creation *date* and *creator* is in principle applicable to both information objects and information realizations.

The example in Section 11.2 shows the use of the very common creator property of Dublin Core together with properties of ID3 to annotate an audio file.

8. Example 5: XMP

XMP is a property-centric metadata format defined by Adobe to describe image assets [7]. Like EXIF and ID3 metadata, XMP metadata is also embedded within the binary media data. XMP aims integrating different standards for image metadata and thus enabling interoperability along the media production process [12]. Compared to the M3O, the industry-driven XMP metadata format is limited with respect to the functionality provided. In addition, XMP focuses on a single media type only, namely images.

Understanding The XMP metadata format is well described in several documents. Regarding this work, we are referring to Part 1 of the XMP metadata format specification. The additional properties of the XMP Specification Part 2 remain as future work. Part 1 of the XML metadata format specification defines several groups of metadata properties, called *namespaces*. These namespaces of XMP are *basic*, *rights management*, *media management*, and *dublin core*.

The *basic* namespace defines properties that provide basic descriptive information like creation date, label, and rating. In the *rights management* namespace, properties are defined for the legal access of a resource like owner, certificate, and terms of usage. In the *media management* namespace, one finds properties such as for the identification and history of a resource. For example, the XMP *media management* property *derivedFrom* can be used to model for a media asset from which other media asset it is derived from (during the media production process). Finally, the *dublin core* namespace of XMP contains the 15 elements specified in the Dublin Core Metadata Element Set. As the Dublin Core Metadata Element Set has already been aligned with the M3O in Section 7, we are able to reuse the previously aligned ontology for Dublin Core.

Grouping Due to the good documentation of XMP, grouping of the concepts and properties for the M3O-aligned ontology of XMP was straightforward. A group has been created for each of the namespaces as introducted above, except for the *dublin core* namespace that did already exist.

Mapping As the three XMP groupings *basic*, *rights management*, and *media management* contain descriptive information about image assets, they have been mapped as specializations of the Annotation Pattern to the M3O. When mapping the individual metadata elements to the M3O, the usual considerations whether to map to information realizations or information objects have been applied. This was of particular importance for the *media management* namespace. Most XMP metadata elements defined in the *media management* namespace are identifiers that could be replaced by using the concepts InformationRealization and InformationObject from DUL instead. For example, the XMP *media management* property *InstanceID* represents a specific incarnation of a resource and is modeled as InformationRealization.

An example of applying the M3O-aligned XMP ontology is shown in Section 11.1.

9. Example 6: SearchMonkey Media

The SearchMonkey Media vocabulary [20] is part of the SearchMonkey vocabulary collection developed by Yahoo! to annotate pages with semantic metadata. Each vocabulary defines terms and classes relevant in the considered domain. In addition, existing vocabularies are reused such as Dublin Core and domain specific knowledge sources like DBpedia [28].

Understanding The Media vocabulary of SearchMonkey defines ten classes and 15 properties. It is focused on single media assets only and introduces media types such as article, audio, image, photo, text, and video. In addition it defines sets of media assets such as a set of photos and a set of videos. The properties cover technical details of the media assets like *bitrate* and *channels* and also the number of times a media asset has been viewed (*views*). The vocabulary uses the domain and range restrictions of RDF Schema [39].

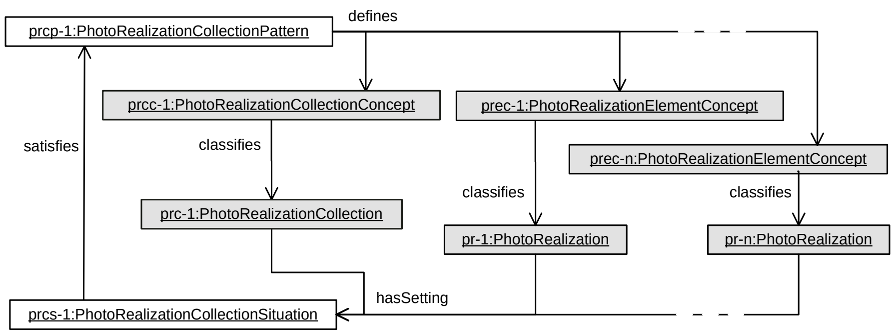


Figure 5: Application of the SearchMonkey Media vocabulary using M3O.

Grouping The SearchMonkey Media vocabulary does not provide any grouping information on the classes and properties it defines. However, some informal subsumptions are documented in the textual description such as the introduction of a superconcept media for the media types. As the SearchMonkey Media vocabulary contains only a few concepts and properties and as they all are serving the same purpose of annotating media assets, we have decided to form only one group.

Mapping The concepts defining the media types in the SearchMonkey Media vocabulary have been mapped to the M3O as subconcepts of InformationRealization, i.e., ArticleRealization, AudioRealization, ImageRealization, TextRealization, and VideoRealization. The concept ImageRealization has two further subconcepts for PhotoRealization and ThumbnailRealization. The concept InformationRealization has been chosen as superconcept for defining the media types instead of InformationEntity, as the properties describing the media assets and associated with SearchMonkey's media concept are of mere technical nature such as *bitrate*, *channels*, *duration*, *fileSize*, *framerate*, and *samplingrate*. For the concepts representing sets of photos and sets of videos, two specializations of M3O's Collection Pattern have been introduced. These are called PhotoRealizationCollection and VideoRealizationCollection.

Application of the Aligned Ontology The application of the M3O-aligned SearchMonkey Media vocabulary is shown in Figure 5 at the example of a PhotoRealizationCollection prc-1, i.e., a collection of PhotoRealizations. It is defined in the context of a PhotoRealizationCollectionPattern that defines a PhotoRealizationCollection classifying the prc-1. The PhotoRealizationCollectionPattern further defines n PhotoRealizationElementConcepts prec-1 to prec-n, which classify the actual elements of the PhotoRealizationCollection. The elements of the PhotoRealizationCollection are the PhotoRealizations pr-1 to pr-n.

10. Example 7: Ontology for Media Resource

The Ontology for Media Resource [21] developed by the W3C defines a core vocabulary for multimedia annotation. The ontology targets at an unifying mapping of common media formats like EXIF [6] or Dublin Core [10, 11]. The core challenge for this alignment is the mapping of properties to either information object or information realization as provided by M3O's Information Realization Pattern.

Understanding The Ontology for Media Resource follows a property-centric approach to ontology modeling and consists of 28 predicates including properties like *title* and *language*. Some properties are specified in further detail, e.g., through *role* or *type* properties. Only entities such as media assets and persons are represented as concepts. Any other information such as roles or types are represented using primitive values, e.g., strings. The Ontology for Media Resource defines no structural patterns and uses the domain and range restrictions of RDF Schema to specify the properties. Unlike the M3O, there is no distinction between information object and information realization.

Grouping The Ontology for Media Resource's documentation on the web provides a number of conceptual groupings for certain aspects of multimedia description, e.g., identification or fragmentation. However, this information is not manifested in the RDF vocabulary. With the alignment of the Ontology for Media Resource and the M3O, we provide grouping information by defining the appropriate axioms. For example an IdentificationAnnotation concept hasPart some TitleAnnotationConcept, LanguageAnnotationConcept, and LocatorAnnotationConcept.

Mapping For mapping the Ontology for Media Resource to the M3O, we define a subconcept of the AnnotationConcept for each property of the ontology. For example, we define a LocatorAnnotationConcept to match the *locator* property. Concrete values are expressed using the Data Value Pattern of the M3O. To this end, we define appropriate Region concepts. In the case of the LocatorAnnotationConcept, we define a LocatorRegion with the property hasRegionDataValue and an URI specifying a concrete location on the web.

By taking into account the difference between information objects and information realizations, we can improve semantic precision of the aligned ontology. To this end, we have examined each attribute of the Ontology for Media Resource for its inherent meaning and constrain it to the appropriate concept of the Information Realization Pattern of the M3O. As an example, the *locator* property of the Ontology for Media Resource annotates media files that are locatable on the web. This is a quality only applicable for information realizations and is expressed in the definition of the LocatorAnnotationConcept.

We express the *type* property of the Ontology for Media Resource through specialization, e.g., by specifying an ImageRealization as subclass of the Information-

Realization. Finally the *fragments* facet of the Ontology for Media Resource has been modeled using the Decomposition Pattern of the M3O. The functionality indicated by the *namedFragments* property can be obtained by decomposing multimedia assets using the Decomposition Pattern and by using the Annotation Pattern to annotate the resulting fragment with a FragmentLabelAnnotationConcept.

A concrete example of applying the M3O-aligned Ontology for Media Resource is described in our technical report [36] available online.

11. Combined Use of Integrated Metadata Models

In Section 11.1, we show the combined application of XMP, COMM, and Linked Data [40]. Linked Data is a community effort to publish and link semantic data on the web. In Section 11.2, we present the combination of ID3 with Dublin Core and XMPm. In both examples, the qualities connecting the information entities with the regions are not shown for reasons of brevity.

11.1. *Example: Combining Dublin Core, COMM, and DBpedia*

We consider the scenario of the Austrian photographer Ferdinand Schmutzer who took a picture of Albert Einstein during a lecture in Vienna in 1921.[d] The example demonstrates the combined use of the metadata standard Dublin Core, the Core Ontology on Multimedia (COMM), and annotations using Linked Data resources from DBpedia to describe the picture.

The fact that such a picture has been taken during the lecture is represented by the InformationObject aeio-1 as shown in Figure 6. The ImageObject is a specialization of DUL's InformationObject and specified in COMM. The AnnotationPattern exp-1 specifies the annotation of the ImageObject. It defines a CreatorConcept and DateConcept from Dublin Core to represent the creator `"F. Schmutzer"` and date of capturing the picture, namely 1921. The capturing date of the picture is modeled using xsd:string as it cannot be ensured that a concrete date proper to the xsd:date type can be provided, like in our example. The AnnotationPattern further defines an annotation concept DepictedPersonConcept that is specifically introduced for the example and classifies the individual dbpedia-einstein-1. It is a resource from DBpedia representing the person Albert Einstein, namely the URI `http://dbpedia.org/resource/Albert_Einstein`. The individual dbpedia-einstein-1 is of rdf:type NaturalPerson from DUL. The metadata is attached to the ImageObject aeio-1 and not to a concrete realization of this object, as the information about the creator, date, and the person (from DBpedia) is independent of a concrete realization of that object.

A second AnnotationPattern exp-2 is used to describe a concrete realization einstein-jpg-1 of the ImageObject aeio-1. This AnnotationPattern defines a RGBHistogramAnnotationConcept using COMM in order to annotate the ImageRealization

[d] `http://de.wikipedia.org/w/index.php?title=Datei:Einstein1921_by_F_Schmutzer_2.jpg`

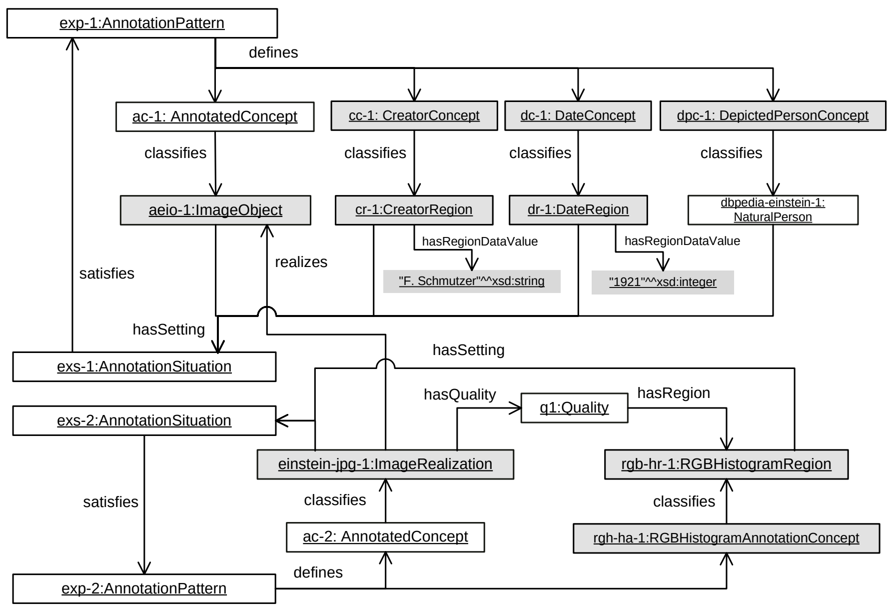


Figure 6: Combination of COMM color histogram with technical details from XMP *basic* namespace and background knowledge in form of Linked Data from DBpedia

einstein-jpg-1 with a histogram. The different parts of the histogram as shown in the COMM example in Section 4 are not shown for reasons on brevity. To connect the InformationRealization einstein-jpg-1 with the InformationObject, the Information Realization pattern of DUL is used as shown in Figure 6. Thus, the ImageRealization einstein-jpg-1 has a property realizes to aeio-1.

11.2. *Example: Combining ID3, Dublin Core, and XMP*

The example shown in Figure 7 demonstrates the combined use of the ID3 metadata format and the Dublin Core and XMP metadata standards. Using the Annotation Pattern and the Decomposition Pattern of the M3O, the song *Mamma Mia* of the pop group ABBA is annotated with information about its musical category, title, and how the media asset has been created. ID3 can embed a picture with the audio asset. This picture is annotated in the example with its mime type, a description, and where it has been derived from.

The DecompositionPattern abba-single-dp-1 defines a CompositeConcept composite-1, which classifies the AudioObject of the ABBA song *Mamma Mia*, represented by the individual mammamia-1. As parts of the AudioObject, the DecompositionPattern defines the two composites composite-1 and composite-2, which classify the AudioRealization track-1 of the ABBA song and the embedded PictureRealization pic-1. This corresponds to a MP3-file of the ABBA song *Mamma Mia* with a picture embedded.

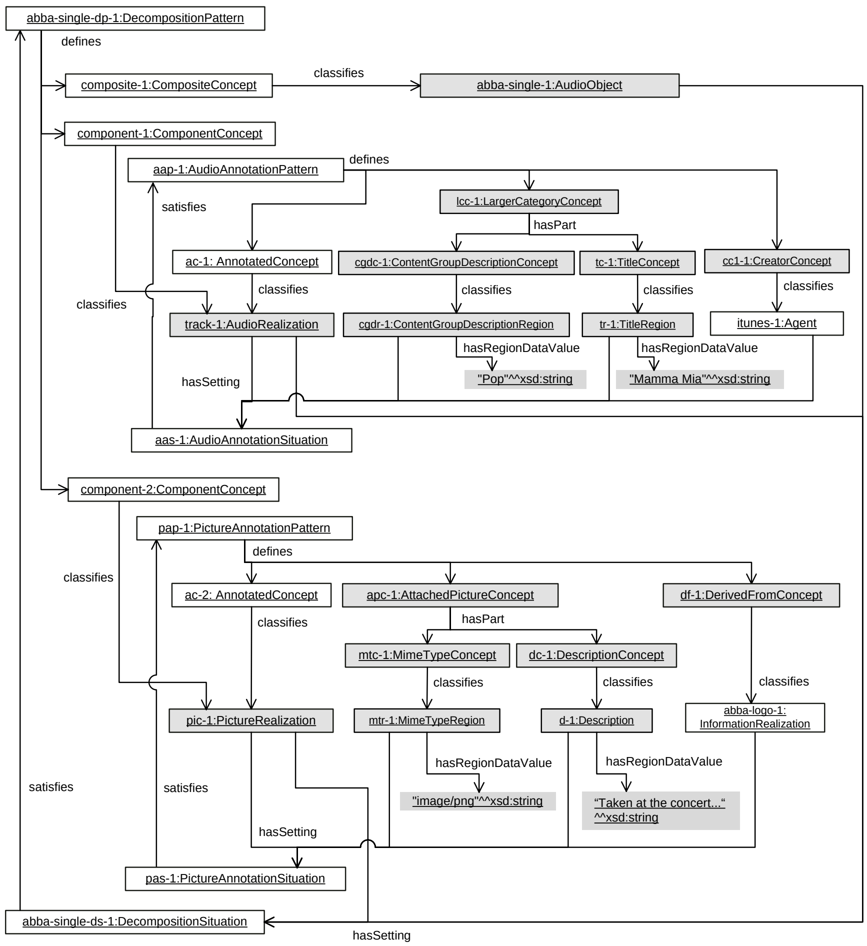


Figure 7: Combination of ID3, DublinCore, and Linked Data from DBpedia

Having introduced the components defined by the Decompositon Pattern, we can start annotating the AudioRealization track-1 and PictureRealization pic-1. The AudioRealization track-1 plays the role of an AnnotatedConcept in the context of the AudioAnnotationPattern aap-1. The annotations of the audio track-1 are defined by the AudioAnnotationPattern, namely the LargerCategoryConcept originating from ID3 and the CreatorConcept from Dublin Core. The LargerCategoryConcept describes by its ContentGroupDescriptionConcept the broader category "Pop" of the ABBA song and its title "Mamma Mia" using the TitleConcept. The CreatorConcept classifies the concrete Agent itunes-1 that has created track-1, namely Apple's iTunes[e] application.

[e] http://www.apple.com/itunes/

The PictureAnnotationPattern pap-1 is used to describe the PictureRealization pic-1. It is annotated using the AttachedPictureConcept from ID3 with its mime type "image/png" and a description. Finally, the origin of the attached picture is modeled using XMP's DerivedFromConcept. The DerivedFromConcept classifies the abba-logo-1, which is a resource from Wikipedia, namely http://upload.wikimedia.org/wikipedia/de/thumb/f/f4/ABBA-Logo.svg/690px-ABBA-Logo.svg.png.

12. Related Work

Numerous metadata standards and metadata formats with different goals and backgrounds have been proposed in research and industry. Most focus on a single media type such as image, text, or video. They differ in the complexity of the data structures they provide and have partly overlapping and partly complementary functionality. With standards like EXIF [6], XMP [7], and IPTC [8], we find metadata models that provide (typed) key-value pairs to represent metadata of the media type image. ID3 provides typed key-value pairs for audio assets [19]. The SearchMonkey Media [20] is a vocabulary developed by Yahoo! to annotate media types such as article, audio, image, photo, text, and video. Like the previous standards and formats, it focuses on single media assets only. However, it introduces sets of media assets such as a set of photos and a set of videos. The Dublin Core Metadata Element Set [11] is a set of 15 metadata elements aimed to describe arbitrary documents. It has been widely adopted and standardized by the International Organization for Standardization. Although the Dublin Core Metadata Element Set is applicable to any kind of documents and thus to different media types, it does not consider the structure of multimedia content like the M3O.

A more complex metadata standard is MPEG-7 [9]. It provides a rich set of complex descriptors that mainly focus on expressing low-level features of images, audio, and video. Several approaches have been published providing a formalization of MPEG-7 as an ontology [34], e.g., by Hunter [1] or the Core Ontology on Multimedia [2]. Although these ontologies provide clear semantics for the multimedia annotations, they still focus on MPEG-7 as the underlying metadata standard. More importantly, these ontologies basically provide a formalization of MPEG-7, but do not provide for the integration of different standards. For further discussions, we refer to the analyses conducted in [18, 41], the report of the W3C Multimedia Semantics Incubator Group [42], and the overview of the W3C Media Annotations Working Group [14].

The drawbacks of these standards are the lacking interoperability and the missing alignment between them. Each standard and format is designed for a specific purpose and application context. A combined use of the different metadata standards and metadata formats, however, is not foreseen. Harmonization efforts like the Metadata Working Group [13] or the Media Annotations Working Group [14] try to tackle these issues and develop a common vocabulary. However, they remain on the same technological level and do not extend their effort beyond single media types, i.e., image, audio, text, and video. In addition, they do not provide a generic framework suitable for integrating arbitrary metadata standards and metadata formats. XMP aims at an integrated standard for image metadata. However, it tackles the problem from a different point of view. While XMP also aims at providing a framework for multimedia metadata, it focuses on images only and does not consider other media types or structured multimedia content. Another major difference is that XMP stays on the level of standards such as EXIF or IPTC and does not take into account requirements such as provenance of metadata, decomposition, or information realization. The Ontology for Media Resource [21] of the W3C aims at providing a common mapping for media formats like EXIF [6] or Dublin Core [10,11]. To this end, it defines a set of 28 properties and specifies them using domain and range restrictions of RDF Schema. However, the RDF Schema of the Ontology for Media Resource defines no structural patterns, i.e., groupings of concepts and properties, and also considers single media types only.

The alignment method presented in this paper is fully manual. There are several publications about (semi-)automatic alignment and matching methods like [29–31]. However, these methods typically do not provide the high accuracy we require for aligning the different metadata standards and metadata formats and are usually applied to problems such as ontology learning or the alignment of domain models.

The M3O is a core ontology for multimedia metadata. As such, it provides some generic modeling structure for specific aspects in a particular field [26,43]. It aims at integrating the existing metadata standards and metadata formats rather than replacing them. In addition, it provides support for representing metadata of structured multimedia content, i.e., media assets of different type coherently organized in time and space. The method presented in this paper shows how to align existing metadata standards and metadata formats with such a core ontology. The goal of this work is producing a specialization of the M3O that inherits the same level of formal precision and conciseness. Achieving this goal with an automatic method currently seems not realistic.

13. Conclusions

In this paper, we have shown how the generic modeling framework provided by the Multimedia Metadata Ontology (M3O) can be specialized to integrate existing multimedia metadata standards and metadata formats. To this end, we have developed a four-step alignment method that describes the tasks to be performed. We have demonstrated the applicability of our approach at the example of seven existing metadata models. The M3O and the alignments of the seven metadata standards and metadata formats are available from the following website: http://west.uni-koblenz.de/m3o. The website also provides an extended version of this paper published as technical report [36].

Bibliography

[1] Hunter, J.: Enhancing the semantic interoperability of multimedia through a core ontology. mTrans. on Circuits and Systems for Video Technology **13**(1) (2003)

[2] Arndt, R., Troncy, R., Staab, S., Hardman, L., Vacura, M.: COMM: designing a well-founded multimedia ontology for the web. In: ISWC+ASWC. (2007) 30–43

[3] Markkula, M., Sormunen, E.: End-user searching challenges indexing practices in the digital newspaper photo archive. Information Retrieval **1**(4) (January 2000) 259–285

[4] Hollink, L., Schreiber, A.T., Wielinga, B.J., Worring, M.: Classification of user image descriptions. Int. J. of Human-Computer Studies **61**(5) (Nov. 2004) 601–626

[5] Hollink, L., Schreiber, G., Wielinga, B.: Patterns of semantic relations to improve image content search. Web Semantics: Science, Services and Agents on the World Wide Web **5**(3) (2007) 195–203

[6] JEITA: Exchangeable image file format for digital still cameras: EXIF version 2.2 (April 2002) `http://www.exif.org/Exif2-2.PDF`.

[7] Adobe Systems, Inc.: XMP (2008) `http://www.adobe.com/devnet/xmp/`.

[8] International Press Telecommunications Council: "IPTC Core" Schema for XMP Version 1.0 Specification document (2005) `http://www.iptc.org/`.

[9] MPEG-7: Multimedia content description interface. Technical report, Standard No. ISO/IEC n15938 (2001)

[10] Dublin Core Metadata Initiative: DCMI Metadata Terms (2008) `http://dublincore.org/documents/dcmi-terms/`.

[11] Dublin Core Metadata Initiative: Dublin Core Metadata Element Set, Version 1.1 (2010) `http://dublincore.org/documents/dces/`.

[12] Hardman, L., Obrenovic, Z., Nack, F., Kerhervé, B., Piersol, K.W.: Canonical processes of semantically annotated media production. Multimedia Syst. **14**(6) (2008) 327–340

[13] MWG: Metadata working group (2010) `http://www.metadataworkinggroup.org/`.

[14] W3C: Media annotations working group: Video, audio, images (2011) `http://www.w3.org/2008/WebVideo/Annotations/`.

[15] Adobe Systems, Inc.: Flash file format (2008) `http://www.adobe.com/licensing/developer/`.

[16] W3C: SMIL 3.0 (2008) `http://www.w3.org/TR/SMIL/`.

[17] W3C: SVG (2009) `http://www.w3.org/TR/SVG/`.

[18] Saathoff, C., Scherp, A.: Unlocking the semantics of multimedia presentations in the web with the multimedia metadata ontology. In: WWW, ACM (2010) 831–840

[19] Nilsson, M., Mutschler, M.: ID3 (2009) `http://www.id3.org/`.

[20] Yahoo!: SearchMonkey Media (2009) `http://developer.yahoo.com/searchmonkey/smguide/searchmonkey-media.html`.

[21] W3C: Ontology for media resource 1.0 (2010) http://www.w3.org/TR/mediaont-10/.

[22] Borgo, S., Masolo, C.: Foundational choices in DOLCE. In: Handbook on Ontologies. 2nd edn. Springer (2009)

[23] Gangemi, A., Presutti, V.: Ontology Design Patterns. In: Handbook on Ontologies. 2nd edn. Springer (2009)

[24] Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design patterns: elements of reusable object-oriented software. Programmer's Choice. Addison-Wesley (2004)

[25] Presutti, V., Gangemi, A.: Content ontology design patterns as practical building blocks for web ontologies. In: Conceptual Modeling, Springer (2008) 128–141

[26] Oberle, D.: Semantic Management of Middleware. Springer (2006)

[27] Gangemi, A., Mika, P.: Understanding the semantic web through descriptions and situations. In: CoopIS/DOA/ODBASE, Springer (2003) 689–706

[28] Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., Hellmann, S.: DBpedia - a crystallization point for the web of data. J. Web Sem. **7**(3) (2009) 154–165
[29] Euzenat, J., Shvaiko, P.: Ontology matching. Springer (2007)
[30] Ehrig, M.: Ontology Alignment: Bridging the Semantic Gap. Volume 4 of Semantic Web and Beyond. Springer (2007)
[31] Blomqvist, E.: OntoCase-automatic ontology enrichment based on ontology design patterns. In: International Semantic Web Conference, Springer (2009) 65–80
[32] Fowler, M., Beck, K., Brant, J., Opdyke, W., Roberts, D.: Refactoring: Improving the Design of Existing Code. Addison Wesley (1999)
[33] Kanzaki, M.: EXIF vocabulary workspace - RDF schema (2003) `http://www.w3.org/2003/12/exif/`, last update in 2007.
[34] Dasiopoulou, S., Tzouvaras, V., Kompatsiaris, I., Strintzis, M.G.: Enquiring MPEG-7 based multimedia ontologies. MTAP (October 2009) 331–370
[35] Brickley, D.: Basic Geo (WGS84 lat/long) Vocabulary (2006) `http://www.w3.org/2003/01/geo/`.
[36] Scherp, A., Eißing, D., Saathoff, C.: Integrating multimedia metadata standards and metadata formats with the multimedia metadata ontology: Method and examples. Technical report, University of Koblenz-Landau, Koblenz, Germany (2011) Available from `http://west.uni-koblenz.de/m3o`.
[37] Mylka, A., Sauermann, L., Sintek, M., van Elst, L.: NEPOMUK ID3 ontology: Task-force ontologies (2007) `http://www.semanticdesktop.org/ontologies/nid3/`.
[38] Dublin Core Metadata Initiative: Expressing Dublin Core metadata using the resource description framework (2008) `http://dublincore.org/documents/dc-rdf/`.
[39] W3C: RDF Vocabulary Description Language 1.0: RDF Schema (February 2004) `http://www.w3.org/TR/rdf-schema/`.
[40] Bizer, C., Heath, T., Berners-Lee, T.: Linked data - the story so far. Int. J. Semantic Web Inf. Syst. **5**(3) (2009) 1–22
[41] Haslhofer, B., Klas, W.: A survey of techniques for achieving metadata interoperability. ACM Comput. Surv. **42**(2) (2010)
[42] Boll, S., Bürger, T., Celma, O., Halaschek-Wiener, C., Mannens, E., Troncy, R.: Multimedia Vocabularies on the Semantic Web (July 2007) `http://www.w3.org/2005/Incubator/mmsem/XGR-vocabularies/`.
[43] Borgo, S., Masolo, C.: Foundational choices in DOLCE. In: Handbook on Ontologies. 2nd edn. Springer (2009)