The Landscape of Multimedia Ontologies in the last Decade

Mari Carmen Suárez-Figueroa[1], Ghislain Auguste Atemezing[2], Oscar Corcho[1]

(1)Ontology Engineering Group (OEG). Facultad de Informática. Universidad Politécnica de Madrid (UPM)

Phone: +34-91-336-3672

Fax: +34-91-352-4819

email: {mcsuarez, ocorcho}@fi.upm.es

(2)Multimedia Department, Eurecom. Sophia-Antipolis, France

Phone: +33-(0)4-9300-81 78

Fax: +33-(0)4 9000 82 00

email: auguste.atemezing@eurecom.fr

Abstract. Many efforts have been made in the area of multimedia to bridge the so-called "*semantic-gap*" with the implementation of ontologies from 2001 to the present. In this paper, we provide a comparative study of the most well-known ontologies related to multimedia aspects. This comparative study has been done based on a framework proposed in this paper and called FRAMECOMMON. This framework takes into account process-oriented dimension, such as the methodological one, and outcome-oriented dimensions, like multimedia aspects, understandability, and evaluation criteria. Finally, we derive some conclusions concerning this one decade state-of-art in multimedia ontologies.

Keywords: Ontology, Multimedia, RDF(S), OWL, Comparative Framework.

1. Introduction

Vision and sound are the most used senses to communicate experiences and knowledge. These experiences or knowledge are normally recorded in media objects, which are generally associated to text, image, sound, video and animation. In this regard, a multimedia object can be considered as a composite media object (text, image, sound, video, or animation) that is composed of a combination of different media objects.

Nowadays, a growing amount of multimedia data is being produced, processed, and stored digitally. We are continuously consuming multimedia contents in different formats and from different sources using Google[1], Flickr[2], Picasa[3],

[1] http://www.google.com

YouTube[4], and so on. The availability of huge amounts of multimedia objects implies the need for efficient information retrieval systems that facilitate storage, retrieval, and browsing of not only textual, but also image, audio, and video objects. One potential approach can be based on the semantic annotation of the multimedia content to be semantically described and interpreted both by human agents (users) and machines agents (computers). Hence, there is a strong need of annotating multimedia contents to enhance the agents' interpretation and reasoning for an efficient search.

The annotation of multimedia objects is difficult because of the so-called *semantic gap* [24]; that is, the disparity between low level features (e.g., colour, textures, fragments) that can be derived automatically from the multimedia objects and high level concepts (mainly related to domain content), which are typically derived based on human experience and background. In other words, the semantic gap refers to the lack of coincidence between the information that machines can extract from the visual data and the interpretation that the same data have for a particular person in a given situation. The challenge of unifying both low level elements and high level descriptions of multimedia contents in a unique ontology is one of the ways to contribute to bridge this semantic gap.

The need for a high level representation that captures the true semantics of a multimedia object led at the beginning to the development of the MPEG-7 standard [9] for describing multimedia documents. This standard provides metadata descriptors for structural and low level aspects of multimedia documents, as well as metadata for information about their creators and their format [4]. Thus, MPEG-7 can be used to create complex and comprehensive metadata descriptions of multimedia content. Since MPEG-7 is defined in terms of an XML schema, the semantics of its elements have no formal grounding. Thus, this standard is not enough to provide semantic descriptions of the concepts appearing in multimedia objects. The representation and understanding of such knowledge is only possible through formal languages and ontologies [2]. Expressing multimedia knowledge by means of ontologies increases the precision of multimedia retrieval information systems. In addition, ontologies have the

[2] http://www.flickr.com/
[3] http://picasaweb.google.com
[4] http://www.youtube.com/

potential to improve the interoperability of different applications producing and consuming multimedia annotations.

For this reason, during the last decade, many efforts to build ontologies that can bridge the semantic gap have been done (and even still undergoing) involving sometimes national and international initiatives. The first initiatives were focused on transforming existing standards to ontology-alike formats (e.g., MPEG-7 transformation in [15]). However, as there were many subdomains to cover in the multimedia field (audio, video, news, image, etc.) with different proprietary standards, the need of converging efforts to build multimedia ontologies taking into account existing standards and resources was an imperative. The COMM Ontology [3] was one of the first references in that direction.

However, there is not yet an accepted solution to the problem of how to represent, organize, and manage multimedia data and the related semantics by means of a formal framework [16].

Thus, the aim of this paper is twofold: on the one hand we provide a review of the most well-known and used ontologies in the multimedia domain from 2001 up to now, with special attention to the ones that are free available in RDF(S) or OWL. On the other hand, we propose a comparative framework called FRAMECOMMON to contrast the aforementioned multimedia ontologies, with the purpose of providing some guides to ontology practitioners in the task of reusing ontologies. These guides will be a help to take an adequate decision of which multimedia ontology used either for a new ontology development or for its use in an application in the multimedia domain.

The rest of this paper is organized as follows: Section 2 describe the most well-known ontologies in the multimedia domain as well as the most used standard, that is, MPEG-7. Section 3 puts forward the comparative framework called FRAMECOMMON. Then, Section 4 presents the results of applying FRAMECOMMON to the ontologies described in Section 2. Section 5 presents some relevant related work. Finally, Section 6 draws some conclusions from the comparative analysis.

2. A Catalogue of Multimedia Ontologies

Many multimedia metadata formats, such as ID3[5], EXIF (Exchangeable Image File) or MPEG-7[6], are available to describe what a multimedia asset is about, who has produced it, how it can be decomposed, etc. [14]. For professional content found in archives and digital libraries, a range of in-house or standardized multimedia formats is used. Similar issues arise with the dissemination of user generated content found at social media websites such as Flickr, YouTube, or Facebook[7]. In addition, many efforts to build ontologies that can bridge the semantic gap have been done (and even still ongoing) for diverse applications (annotation areas, multimedia retrieval, etc.), involving sometimes many national or international initiatives.

In this section we summarize a representative set of the most well-known ontologies designed and implemented for describing multimedia aspects, from 2001 up to now, with special attention to the ones that are free available in RDF(S) or OWL. This set cannot be considered as exhaustive, but rather cover as much as possible multimedia ontologies presented in the literature.

It is worth mentioning that we do not deal with controlled vocabularies or standards neither with thesauri. The only exception is the MPEG-7 standard that is presented due to two reasons (1) for its importance in the multimedia domain to describe media contents using low level descriptors and (2) for having being transformed to owl-alike formats in various ontologies presented in the literature. After describing the MPEG-7 standard in Section 2.1, we follow in Section 2.2 by the presentation of the ontologies dedicated to describe multimedia objects in general. With respect to visual aspects, Section 2.3 presents ontologies describing images and shapes, as visual elements for representing images; while Section 2.4 presents ontologies for describing visual objects in general. Regarding audio aspects, we present music ontologies in Section 2.5. To sum up, Fig. 1 shows in a chronological order when the different ontologies presented in this section have been released. Finally, in Section 2.6, we provide a brief summary of the 16 ontologies presented.

[5] http://www.id3.org
[6] http://www.chiariglione.org/mpeg
[7] http://www.facebook.com

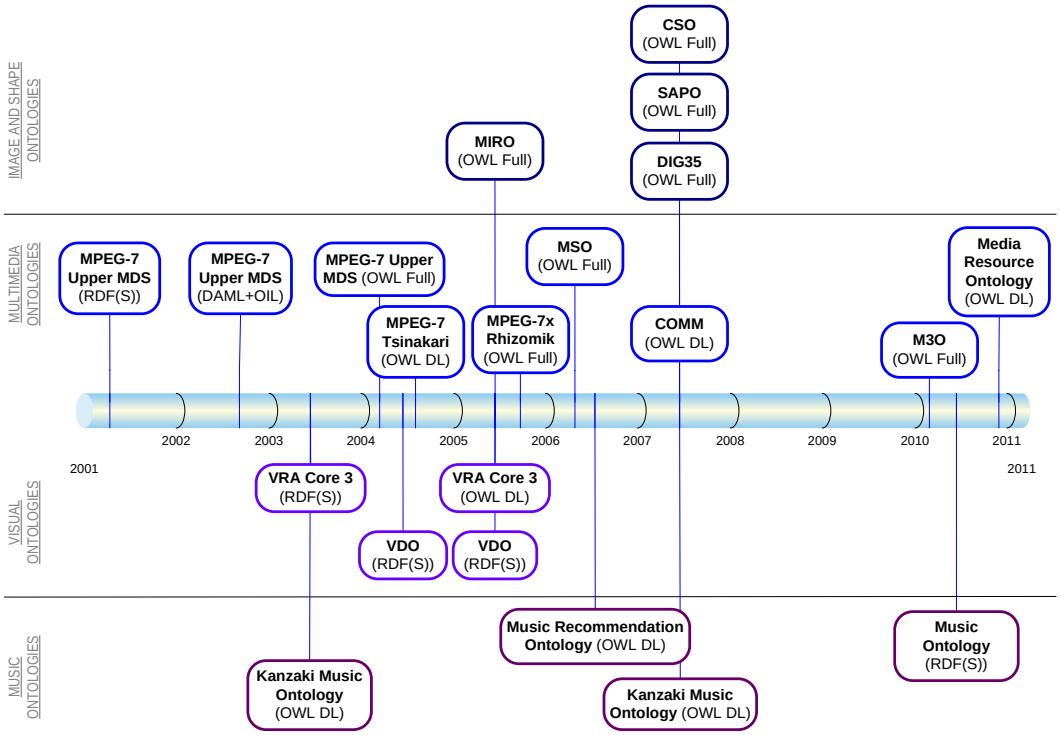

Fig. 1 Time line for the ontologies in the multimedia domain from 2001 to 2011

2.1 MPEG-7

MPEG-7 [17, 18] is an ISO/IEC standard developed by MPEG (Moving Picture Experts Group), formally named “Multimedia Content Description Interface”. It is a standard for describing the multimedia content data that supports some degree of interpretation of the information meaning, which can be passed onto, or accessed by, a device or a computer code. The MPEG-7 standard aims to be a set of descriptors for describing any multimedia content. MPEG-7 standardizes the “description tools” for multimedia content: Descriptors (Ds), Description Schemes (DSs) and the relationships between them. *Descriptors* are used to represent specific features of the content, generally low level features such as visual (e.g., texture, camera motion) or audio (e.g., melody), while *description schemes* are metadata structures for describing and annotating audio-visual content and refer to more abstract description entities (usually a set of related descriptors). These description tools as well as their relationships are represented using the Description Definition Language (DDL).

MPEG-7 defines, in terms of an XML Schema, a set of descriptors where a semantically identically metadata can be represented in multiple ways [27]. For instance, different semantic concepts like frame, shot or video cannot be distinguished based on the provided XML Schema. Thus, ambiguities and

inconsistencies can appear because of the flexibility in structuring the descriptions. For this reason, one of the drawbacks of MPEG-7 is the lack of precise semantics.

2.2 Ontologies for describing Multimedia Objects

In this section, we first present three ontologies (COMM, M3O, and Media Resource Ontology) which can be considered to be generic for the multimedia domain. The way two of these ontologies (COMM and M3O) have been developed is a nice example of what it is nowadays used and recommended in Ontology Engineering, that is, the reuse of knowledge resources[8] in the ontology development. In the second part of this section, we present (a) three initiatives (MPEG-7 Upper MDS, MPEG-7 Tsinaraki, and MPEG-7 Rhizomik) focused on "translating" the MPEG-7 standard to RDF(S) and OWL and (b) one ontology called MSO that combines high level domain concepts and low level multimedia descriptions.

2.2.1 COMM: Core Ontology for MultiMedia

The Core Ontology for MultiMedia (COMM)[9] was proposed by [3] and developed within the X-Media project[10] as a response to the need of having a formal description of a high quality multimedia ontology satisfying a set of requirements such as MPEG-7 standard compliance, semantic interoperability, syntactic interoperability, separation of concerns, modularity and extensibility. Thus, the aim of COMM is to enable and facilitate multimedia annotation. The intended use of COMM is to ease the creation of multimedia annotations by means of a Java API[11] provided for that purpose.

COMM is designed using DOLCE [12] and two ontology design patterns (ODPs): one pattern for contextualization called Descriptions and Situations (DnS) and the second pattern for information objects called Ontology for Information Object (OIO). The ontology is implemented in OWL DL. COMM covers the description schemes and the visual descriptors of MPEG-7. This ontology is composed of 6

[8] Knowledge resources refer to ontologies, non-ontological resources, and ontology design patterns.
[9] http://multimedia.semanticweb.org/COMM/
[10] http://www.x-media-project.org
[11] http://comm.semanticweb.org

modules (visual, text, media, localization, datatype, and core). Just to mention some of the knowledge, `Multimedia-data` is an abstract concept that has to be further specialized for concrete multimedia content types (e.g., `Image-data` that corresponds to the pixel matrix of an image). In addition, according to the OIO pattern, `Multimedia-data` is realized by some physical media (e.g., an image).

2.2.2 M3O: Multimedia Metadata Ontology

The ontology M3O[12] [7], developed within the weKnowIt project[13], aims at providing a pattern that allows accomplishing exactly the assignment of arbitrary metadata to arbitrary media. This ontology is used within the SemanticMM4U Component Framework[14] for the multi-channel generation of semantically-rich multimedia presentations.

M3O is based on requirements extracted from existing standards, models, and ontologies. This ontology provides patterns that satisfy the following five requirements: (1) identification of resource, (2) separation of information objects and realizations, (3) annotation of information objects and realizations, (4) decomposition of information objects and realizations, and (5) representation of provenance information.

To fullfil the five requirements abovementioned, M3O represents data structures in the form of ODPs based on the formal upper-level ontology DOLCE+DnS Ultralight[15] (DUL). Thus, there is a clear alignment with DOLCE+DnS Ultralight as formal basis. The following three patterns specialized from DOLCE and DUL are reused in the M3O: Description and Situation Pattern (DnS), Information and Realization Pattern, and Data Value Pattern.

Besides, M3O provides four patterns[16] that are respectively called annotation pattern, collection pattern, decomposition pattern, and provenance pattern. M3O annotations are in RDF and can be embedded into SMIL (Synchronized Multimedia Integration Language) multimedia presentations. M3O has been

[12] http://www.uni-koblenz-landau.de/koblenz/fb4/AGStaab/Research/ontologies/m3o
[13] http://www.weknowit.eu/
[14] http://sourceforge.net/projects/semanticmm4u/
[15] http://ontologydesignpatterns.org/wiki/Ontology:DOLCE%2BDnS_Ultralite
[16] http://m3o.semantic-multimedia.org/ontology/2010/02/28/

aligned[17] with the following ontologies and vocabularies: COMM, Media Resource Ontology of the W3C, and the image metadata standard EXIF.

2.2.3 Media Resource Ontology

The Media Resource Ontology[18] of the W3C Media Annotation Working Group[19], which is still in development, 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 on the Web at the moment. The Media Resource Ontology defines mapping with the following 23 general multimedia metadata: CableLabs 1.1, CableLabs 2.0, DIG35, Dublin Core, EBUCore, EBU P-Meta, EXIF 2.2, FRBR, ID3, IPTC, iTunes, LOM 2.1, Core properties of MA WG, Media RDF, Media RSS, MPEG-7, METS, NISO MIX, Quicktime, SearchMonkey, Media, DMS-1, TV-Anytime, TXFeed, XMP, and YouTube Data API Protocol. This ontology aims to unify the properties used in such formats. The basic properties include elements to describe: the identification, creation, content description, relational, copyright, distribution, fragments and technical properties. The core set of properties and mappings provides the basic information needed by targeted applications for supporting interoperability among the various kinds of metadata formats related to media resources that are available on the Web. The properties defined in the ontology are used to describe media resources that are available on the web.

Regarding some important classes, it is worth mentioning that a `MediaResource` can be one or more images and/or one or more `Audio Visual (AV) MediaFragment`. By definition, in the model, an `AV MediaResource` is made of at least one `MediaFragment`. A `MediaFragment` is the equivalent of a segment or a part in some standards like NewsML-g2 or EBUCore. At the same time, a `MediaFragment` is composed of one or more media components organized in tracks (separate tracks for captioning/subtitling or signing if provided in a separate file): audio, video, captioning/subtitling, and signing.

[17] http://semantic-multimedia.org/index.php/M3O:Main#Mappings
[18] http://dev.w3.org/2008/video/mediaann/mediaont-1.0/ma-ont.owl
[19] http://www.w3.org/TR/mediaont-10/

2.2.4 MPEG-7 Upper MDS

The MPEG-7 Upper MDS ontology[20] [15] was developed within the Harmony Project[21] with the aim of building an ontology that can be exploited and reused by other communities on the Semantic Web to enable the inclusion and exchange of multimedia content through a common understanding of the associated MPEG-7 multimedia content descriptions. The ontology was firstly developed in RDF(S), then converted into DAML + OIL, and is now available in OWL-Full. The ontology covers the upper part of the Multimedia Description Scheme (MDS) of the MPEG-7 standard.

2.2.5 MPEG-7 Tsinaraki

This MPEG-7 ontology[22] [28] was developed in the context of the DS-MIRF Framework, partially funded by the DELOS II Network of Excellence in Digital Libraries[23]. The ontology was used for annotation, retrieval, and personalized filtering for the Digital Library-related areas (the later in conjunction with the Semantic User Preference Ontology described in [28]). Some other intended use was for summarization and content adaptation.

The ontology is implemented in OWL DL and covers the full MPEG-7 MDS (including all the classification schemes) and partially the MPEG-7 Visual and Audio Parts. MPEG-7 complex types correspond to OWL classes, which represent groups of individuals interconnected because they share some properties. The simple attributes of the complex type of the MPEG-7 MDS are represented as OWL datatype properties. Complex attributes are represented as OWL object properties, which relate class instances. Relationships between the OWL classes correspond to the complex MDS types and are represented by instances of `RelationBaseType` [28].

2.2.6 MPEG-7 Rhizomik

This MPEG-7 ontology [13] has been produced fully automatically from the MPEG-7 standard using XSD2OWL[24], which transforms an XML Schema into an

[20] http://metadata.net/mpeg7/mpeg7.owl
[21] http://itee.uq.edu.au/~jane/
[22] http://elikonas.ced.tuc.gr/ontologies/av_semantics.zip
[23] http://www.delos.info/
[24] http://rhizomik.net/html/redefer/#XSD2OWL

OWL ontology. The ontology aims to cover the whole standard and it is thus the most complete one (with respect to the ontologies presented in Sections 2.2.4 and 2.2.5). The definitions of the XML Schema types and elements of the ISO standard have been converted into OWL ones according to the set of rules given in [9]. The ontology can easily be used as an upper-level multimedia ontology for other domain ontologies (e.g., music ontology).

2.2.7 MSO

The Multimedia Structure Ontology (MSO) [6] was developed within the context of the aceMedia[25] project based on MPEG-7 MDS, along with three other ontologies: Visual Descriptors Ontology, Spatio-Temporal Ontology, and Middle Level Ontology. The main aims of the ontologies developed were (a) to support audiovisual content analysis and object/event recognition, (b) to create knowledge beyond object and scene recognition through reasoning processes, and (c) to enable a user-friendly and intelligent search and retrieval. MSO combines high level domain concepts and low level multimedia descriptions, enabling for new media content analysis. MSO covers the complete set of structural description tools from MPEG-7 MDS. The ontology has been aligned to DOLCE.

MSO played a principal role in the automatic semantic multimedia analysis process, through tools developed in aceMedia projects (M-OntoMat-Annotizer, Visual Descriptors Extraction (VDE) plugin, VDE Visual Editor and Media Viewer). The purpose of these tools is to automatically analyze content, generate metadata/annotation, and support intelligent content search and retrieval services.

2.3 Ontologies for describing Images and Shapes

In this section, we make a brief description of ontologies that were developed with special emphasis in images and shapes, as visual elements for representing images. We first describe the DIG35 ontology, which aims at describing digital images. Then, we follow by presenting SAPO, CSO, and MIRO that respectively treat about shape acquisition, commonly shapes description, and specific regions of images.

[25] http://www.acemedia.org/aceMedia

2.3.1 DIG35

DIG35 specification [11] is a set of public metadata for digital images. This specification promotes interoperability and extensibility, as well as a uniform underlying construct to support interoperability of metadata between various digital imaging devices. The metadata properties are encoded within an XML Schema and cover the following aspects: *Basic Image Parameter* (a general-purpose metadata standard); *Image Creation* (e.g., the camera and lens information); *Content Description* (who, what, when, and where aspects of an image); *History* (partial information about how the image got to the present state); *Intellectual Property Rights* (metadata to either protect the rights of the owner of the image or provide further information to request permission to use it); and *Fundamental Metadata Types and Fields* (to define the format of the field described in the metadata block).

The DIG35 ontology[26] is an OWL Full ontology developed by the IBBT Multimedia Lab[27] (University of Ghent) in the context of the W3C Multimedia Semantics Incubator Group. This ontology provides an OWL Schema covering the entire DIG35 specification.

2.3.2 SAPO

The Shape Acquisition and Processing Ontology (SAPO)[28] [1] was intended to provide a starting point for the formalization of the knowledge involved in the creation and processing of digital shapes. The ontology was developed within the AIM@SHAPE project[29].

SAPO is an OWL Full ontology that covers the development, usage and sharing of hardware tools, software tools, and shape data in the field of acquisition and reconstruction of shapes. Examples of classes are `Acquisition Condition`, materialized by two conditions used to acquire data: environmental and logistic; `Acquisition Device`, being a system of sensors connected to a storage device designed for acquiring data; `Shape Type`, to describe categories of

[26] http://multimedialab.elis.ugent.be/users/chpoppe/Ontologies/DIG35.owl
[27] http://www.mmlab.be
[28] http://www.aimatshape.net/resources/aas-ontologies/shapecommonontology.owl/download
[29] http://www.aimatshape.net/

shapes; `Shape Data`, is the concrete data associated to a shape; and `Processing System` and `Processing Session`.

2.3.3 CSO

The purpose of the Common Shape Ontology (CSO)[30] [29], developed also within the AIM@SHAPE project, is to integrate some shared concepts and properties from the domain ontologies and the metadata information from the Shape Repository[31] (a shared repository populated with a collection of digital shapes) that can be associated with any shape model.

CSO is an OWL Full ontology that represents for example the following knowledge: types of *geometrical representations* such as contour set, points sets or mesh, and *structural descriptors* for shapes such as centre line graph, multidimensional structural descriptor. These two metadata information (geometrical representations and structural descriptors) are considered common to any kind of shape regardless of the domain.

CSO has been used in (a) the Digital Shape Workbench (DSW)[32], a common infrastructure for integrating, combining, adapting, and enhancing existing and new software tools and shape databases; and (b) the Geometric Search Engine (GSE)[33], for simple search of digital resources.

2.3.4 MIRO

The main purpose of the Mindswap Image Region Ontology (MIRO)[34] is to provide the expressiveness to assert what is depicted within various types of digital media, including image and videos [14]. MIRO has been applied in the annotation tool PhotoStuff[35], which aims at providing annotation of an image and its regions with respect to concepts from any number of ontologies specified in RDF(S) or OWL [14].

MIRO is an OWL Full ontology that models concepts and relations covering various aspects of the digital media domain (Image, Segment, Video, Video Frame, etc). The ontology defines concepts including: `digital media` to

[30] http://www.aimatshape.net/resources/aas-ontologies/shapecommonontology.owl/
[31] http://shapes.aimatshape.net/
[32] http://dsw.aimatshape.net/
[33] http://dsw.aimatshape.net/sse/Search.jsp?ontology=shapes
[34] http://www.mindswap.org/2005/owl/digital-media
[35] http://www.mindswap.org/2003/PhotoStuff/

model digital media data; `segment`, class for fragments such as video segment of digital media content; and `video text` to model spatio-temporal regions of video data that correspond to text and caption. The ontology also defines relations such as `depicts`, `segmentOf`, `hasRegion`, and `regionOf`.

2.4 Ontologies for describing Visual Objects

In this section, we present two ontologies, VDO and VRA Core 3; describing respectively visual descriptors and collection of cultural works.

2.4.1 VDO

The Visual Descriptor Ontology (VDO)[36] [23] deals with semantic multimedia content, analysis, and reasoning. VDO was developed within the aceMedia project. VDO was used in the automatic semantic multimedia analysis process, through tools developed in aceMedia.

VDO, available in RDF(S), contains representations of MPEG-7 visual descriptors and models concepts and properties that describe visual characteristics of objects. Examples are `basic descriptors` containing spatial coordinates and temporal interpolation; `colour descriptor` with descriptors for colour layout, colour structure or colour dominant descriptor; `meta concepts` such as colour space type, motion model type; and `motion descriptor`, `shape descriptor` and t`exture descriptor`. VDO has been aligned to the DOLCE ontology.

2.4.2 VRA Core 3

The Visual Resource Association (VRA)[37] is an organization consisting of many American universities, galleries, and art institutes. These often maintain large collections of (annotated) slides, images, and other representations of works of art. This association has defined the VRA Core Categories [30] to describe such collections. The last release version is VRA Core 4.0[38], which consists of 19 descriptors for 3 types of objects: work (`vra:Work`), collection of works and/or images (`vra:Collection`) and image (`vra:Image`). This version includes

[36] http://www.acemedia.org/aceMedia/files/software/m-ontomat/acemedia-visual-descriptor-ontology-v09.rdfs
[37] http://www.vraweb.org/

one more type of object (`vra:Collection`) with respect to VRA Core 3.0. The VRA Core 3.0 elements were designed to facilitate the sharing of information among visual resources collections about works and images. A *work* is a physical entity that exists, has existed at some time in the past, or that could exist in the future (e.g., painting, composition, an object of material culture). An *image* is a visual representation of a work (it can exist in photomechanical, photographic and digital formats). A *visual resources collection* may own several images of a given work.

Two versions of VRA 3.0 were developed in RDF(S)[39] and OWL[40]. In both ontologies, a VisualResource can be an image or a work, insert in a Period and supported in a Material.

2.5 Ontologies for describing Music

In this section, ontologies for describing the audio media type, particularly those objects related to music are described. The ontologies concerned are the following: Music ontology, Kanzaki Music vocabulary, and Music Recommendation ontology.

2.5.1 Music Ontology

The Music Ontology[41] [22] is an attempt to provide (a) a vocabulary for linking wide range music-related information, and (b) a democratic mechanism for doing so. The parts of the Music Ontology related to the production process of a particular piece of music (composition, performance, arrangement, etc.) as well as the parts dealing with time-related information are based on three external ontologies: Time, TimeLine (a timeline being a coherent backbone for temporal things) and Event (to express knowledge about the production process of a piece of music). Likewise, in order to describe music-related events, they consider describing the workflow beginning with the creation of a musical work to its release on a particular record. Apart from the three ontologies cited before, the Music Ontology is mainly influenced by the FRBF Final report[42], the ABC

[38] http://www.vraweb.org/projects/vracore4/index.html
[39] http://simile.mit.edu/2003/10/ontologies/vraCore3
[40] http://www.w3.org/2001/sw/BestPractices/MM/vracore3.owl
[41] http://motools.sourceforge.net/doc/musicontology.rdfs
[42] http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records

ontology from the Harmony Project[43] and the FOAF project[44]. In addition, the Music Ontology reuses the WGS84 Geo Positioning vocabulary[45].

Some relevant concepts implemented in the Music Ontology are the following: `event` related to the process of releasing a musical work such as arrangement, composition, recording, show, etc.; `musical item` containing different types of mediums such as vinyl, CD, stream, magnetic tape; and `release type` of a particular manifestation, such as album, review, or remix.

2.5.2 Kanzaki's Music Vocabulary

Kanzaki[46] is an OWL DL ontology to describe classical music and performances. Classes for musical works, events, instruments and performers, as well as related properties are defined. In Kanzaki ontology, it is important to distinguish musical works (e.g., Ballet) from performance events (Ballet Event), or works (e.g., Choral Music) from performer (Chorus) whose natural language terms are used interchangeably. Some relevant classes modelled are the following ones: `musical work` which contains among other classes opera, religious music, orchestral work or choral music; `musical representation`, representation of a musical work, such as a score, sheet music, performance, recoding, etc.; `musical instruments` such as string instrument, woodwind, brass, percussion and keyboards instruments; and `artist`, `musical groups` and `singer` that are specialization of the FOAF ontology.

2.5.3 Music Recommendation Ontology

The Music Recommendation Ontology[47] is an ontology implemented in OWL DL that describes basic properties of the artists and the music titles, as well as some descriptors extracted from the audio (e.g., tonality -key and mode-, rhythm -tempo and measure-, intensity). The ontology is part of a music recommender system (foafing the music) [8] which aims at recommending music to users depending on (a) personalized profiles (FOAF profile and listening habits) and (b) RDF Site Summary (RSS) vocabularies. Therefore, music information (new album releases,

[43] http://www.metadata.net/harmony/ABCV2.htm
[44] http://www.foaf-project.org/
[45] http://www.w3.org/2003/01/geo/wgs84_pos#
[46] http://www.kanzaki.com/ns/music
[47] http://foafing-the-music.iua.upf.edu/music-ontology#

podcast sessions, audio from MP3 blogs, related artists' news. and upcoming gigs) is gathered from thousands of RSS feeds. In addition, a way to align this ontology with the MusicBrainz[48] ontology and the MPEG-7 standard is proposed in [13].

2.6. Summary

In this section we provide a short summary of the 16 ontologies briefly described in this paper.

With respect to multimedia ontologies, it is worth mentioning that (a) COMM is an ontology with a modular design, which facilitates its extensibility and integration with other ontologies, (b) M3O is based on ontology design patterns and is targeted to multimedia presentations on the web, (c) the Media Resource Ontology provides a set of mappings with a great range of multimedia metadata, (d) the four ontologies (MPEG-7 Upper MDS, MPEG-7 Tsinakari, MPEG-7 Rhizomik, and MSO) that are the result of transforming the MPEG-7 standard to ontology languages are based on a monolithic design.

Regarding ontologies for describing images and shapes, we can mention that (a) DIG 35 covers the standard DIG 35, (b) SAPO mainly covers shape data and how to process it, (c) CSO implements geometric representations, and (d) MIRO models diverse aspects of the digital media domain.

With respect to visual resource ontologies, it is worth mentioning that VDO covers the MPEG-7 standard and VRA Core 3 is suitable to describe collection of arts work in galleries.

Regarding music ontologies, we can mention that (a) Music Ontology does not cover the low level audio descriptors, (b) Kanzaki Music Ontology distinguishes among musical works, events, and performance, and (c) Music Recommendation Ontology provides descriptors for audio features together with properties for describing artists and music works.

Finally, Table 1 shows an overview of these 16 ontologies with respect to the initiative in which they were developed, entity metrics, and ontology usage.

[48] http://musicbrainz.org/

Ontology Name	*Initiative*	*Metrics*	*Usage*
Multimedia Ontologies			
COMM	X-Media Project	Modules: 6 Classes[49]: 40 Objects Properties: 10	Annotation
M3O	weKnowIt Project	Classes[50]: 126 Objects Properties: 129	Generation (SemanticMM4U Component Framework)
Media Resource Ontology	W3C Media Annotation Working Group	Classes: 14 Objects Properties: 55	Annotation Analysis
MPEG-7 Upper MDS	Harmony Project	Classes: 69 Objects Properties: 38	Annotation Analysis
MPEG-7 Tsinakari	DELOS II Network of Excellence	Classes: 420 Objects Properties: 175	Annotation Personalized filtering (DS-MIRF Framework)
MPEG-7 Rhizomik	Rhizomik	Classes: 814 Objects Properties: 580	Annotation (MusicBrainz intiative)
MSO	aceMedia Project	Classes: 23 Objects Properties: 9	Analysis Retrieval (M-ontoMat-Annotizer, Media Viewer, VDE plugin and VDE Visual Editor)
Image and Shape Ontologies			
DIG 35	W3C Multimedia Semantics Incubator Group	Classes: 149 Objects Properties: 203	Annotation Analysis
SAPO	AIM@SHAPE project	Classes: 51 Objects Properties: 41	Annotation Analysis
CSO	AIM@SHAPE project	Classes: 38 Objects Properties: 14	Annotation Search (Digital Shape Workbench (DSW) and Geometric Search Engine (GSE))
MIRO	DARPA, the Air Force Research Laboratory, and the Navy Warfare Development Command	Classes: 14 Objects Properties: 12	Annotation (PhotoStuff)
Visual Ontologies			
VRA Core 3	SIMILE Project	*RDF(S) Version* Classes: 10 Object Properties: 50 *OWL Version* Classes: 7 Object Properties: 66	Annotation
VDO	aceMedia Project	Classes: 61 Objects Properties: 237	Analysis
Music Ontologies			
Music Ontology	Centre for Digital Music, Queen Mary, University of London[51]	Classes: 138 Objects Properties: 267	Annotation
Kanzaki Music Ontology	--	Classes: 112 Objects Properties: 34	Analysis
Music Recommendation Ontology	Universitat Pompeu Fabra SALERO	Classes: - Objects Properties: -	Annotation (Recommender system 'foafing the music')

Table 1 Overview based on initiative, ontology metrics, and ontology usage

[49] Metrics concern only the "core ontology".
[50] Metrics concern only the "annotation pattern".
[51] http://www.elec.qmul.ac.uk/

3. Comparative Framework for Ontologies in the Multimedia Domain

In this paper we argue that a comprehensive analysis of the most well-known ontologies in the multimedia domain will lead to a more complete understanding of the semantic status in such a domain.

To perform a systematic comparison of the ontologies presented in Section 2, we have designed a comparative framework called FRAMEwork for COntrasting MultiMedia ONtologies (FRAMECOMMON), which is presented in Fig. 2. It is worth mentioning that the objective of FRAMECOMMON is not to make any judgment about the different ontologies in the multimedia domain. Instead, we aim to provide insights and guides on different features that may help practitioners to select the most suitable multimedia ontology both (a) for reusing it in another ontology development or (b) for using it in a semantic application.

FRAMECOMMON is divided into 4 dimensions: the methodological one that is oriented to the process used during the ontology building, and other 3 dimensions (multimedia dimension, usability profiling dimension, and reliability dimension) oriented to the outcome, that is, the ontology.

Since the main aim of our work is to provide help to ontology practitioners in the task of selecting available multimedia ontologies for their reuse, we argue that the process followed during the ontology development is an important dimension to be taken into account. The way in which an ontology has been developed can provide interesting clues about the confidence such an ontology inspires. The modelling choices when the ontology has been developed affect different aspects like (a) the integration and link with other ontologies and (b) the interoperability and scalability of the applications using these ontologies. On the other hand, we also claim that analysing an ontology with respect to the other 3 dimensions proposed helps in the selection task. That is, the rest of dimensions have been proposed for measuring, respectively, the suitability of an ontology with respect to a set of requirements related to multimedia features, the easiness of understanding and using the ontology, and the quality of the ontology.

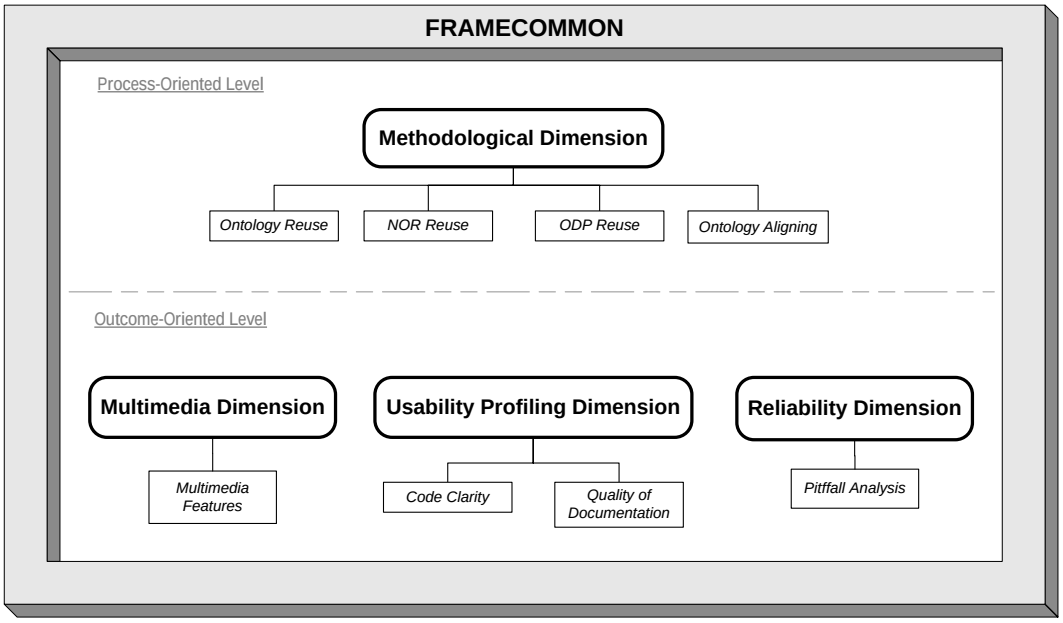

Fig. 2 FRAMECOMMON Dimensions

FRAMECOMMON dimensions are described as follows:

- *Methodological dimension*: it refers to whether the ontology was developed by reusing any knowledge resource (ontological resources, non-ontological resources (NORs), and ODPs), as proposed by the NeOn Methodology [25]. In addition, in this dimension we also analyze whether any alignment has been established with other ontologies and/or NORs.
- *Multimedia dimension*: it refers to which particular multimedia features as within MPEG-7 multimedia content classification [5] (multimedia, audio, video, image, visual, and audiovisual) are covered by the ontology.
- *Usability profiling dimension*: it refers to the communication context of an ontology. In this sense we want to find out if the ontology provides information that facilitates its understanding. In this case, the following criteria should be analyzed:
  - *Code clarity*. It refers to whether the code is easy to understand and modify, that is, if the knowledge entities follow unified patterns and are clear [19, 25]. This would improve the clarity of the ontology and its monotonic extendibility. This criterion also refers to whether the code is documented, that is, if it includes clear and coherent definitions and comments for the knowledge entities represented in the ontology.

- *Quality of the documentation*. It refers to whether there is any communicable material used to describe or explain different aspects of the ontology (e.g., modelling decisions). The documentation should explain the domain and the knowledge pieces represented in the ontology so that a non-expert could learn enough about the domain and be able to understand the knowledge represented in the ontology [19, 25].

- *Reliability dimension*: it refers to analyzing whether we can trust in the ontology, that is, whether the ontology is free of anomalies or worst practices [20, 21]. In this regard, we suggest that soundly developed ontologies are better candidate for reuse.

4. Applying FRAMECOMMON

We have applied FRAMECOMMON to the 16 ontologies described in Section 2. In this section we aim to explain how each dimension of FRAMECOMMON has been analyzed as well as to present the results obtained for each dimension.

In the case of the *methodological dimension*, we have reviewed the available documentation about how the ontology development was performed. We have focused on two key activities in the ontology development that are the reuse of knowledge resources and the aligning with available resources. After this revision we have obtained the results shown in Table 2.

With respect to the *multimedia dimension*, we have manually inspected the ontologies to determine which multimedia features are covered (multimedia, audio, video, image, visual, and audiovisual). The results obtained from this inspection are shown in Table 3.

Regarding the *usability profiling dimension*, we have first focused on the quality of the documentation criterion. In this case, we have analyzed whether the ontology has documentation, and if such documentation really explains the domain and the ontology itself, as well as modelling criteria using during the ontology development. We have considered a high level quality if there is a wiki, an article or even a web page explaining and/or describing the ontology. Secondly, we have focused on the code clarity criterion. In that case we have inspected the ontology code by analyzing the complexity of the definitions (and axioms) implemented the ontology. We have also analyzed whether the code is easy to understand and modify by means of inspecting the following aspects in the

code: (i) if the concepts names are clear, (ii) if the definitions are coherent, and (iii) if the ontology provides comments and metadata. In general, we have considered a low clarity when the concepts are not clear and a high clarity when the ontology in general is intuitively understandable. The results of analysing this dimension are presented in Table 4.

Finally, in the case of the *reliability dimension*, we have manually inspected the ontologies with respect to the catalogue of pitfalls described in [20, 21]. The results of this inspection are shown in Table 5.

	Methodological Dimension			
Ontology Name	***Ontological resources reused***	***Non-ontological resources reused***	***ODPs reused***	***Aligned***
Multimedia Ontologies				
COMM	DOLCE	MPEG-7	DnS, OIO	--
M3O	DOLCE & DnS Ultralight (DUL)	--	DnS, Information and Realization Pattern, Data Value Pattern	COMM, Media Resource Ontology, EXIF
Media Resource Ontology	--	--	--	CableLabs 1.1, CableLabs 2.0, DIG35, Dublin Core, EBUCore, EBU P-Meta, Exif 2.2, FRBR, ID3, IPTC, iTunes; LOM 2.1, Core properties of MAWG, Media RDF, Media RSS, MPEG-7, METS, NISO MIX, Quicktime, SearchMonkey, Media, DMS-1, TV-Anytime, TXFeed, XMP, YouTube Data API Protocol
MPEG-7 Upper MDS	--	MPEG-7 (MDS)	--	--
MPEG-7 Tsinakari	--	MPEG-7	--	--
MPEG-7 Rhizomik	--	MPEG-7	--	--
MSO	--	MPEG-7 (MDS)	--	DOLCE
Image and Shape Ontologies				
DIG 35	--	DIG 35	--	--
SAPO	--	--	--	--
CSO	--	--	--	--
MIRO	--	--	--	--
Visual Ontologies				
VRA Core 3	--	VRA Element Set	--	--
VDO	--	MPEG-7	--	DOLCE
Music Ontologies				
Music Ontology	Time, TimeLine, Event, FOAF, ABC	WGS84 Geo Positioning Vocabulary	--	--
Kanzaki Music Ontology	FOAF	--	--	--
Music Recommendation Ontology	FOAF	RDF Site Summary (RSS)	--	MusicBrainz ontology and the MPEG-7 standard (Proposal)

Table 2 Comparison of ontologies with respect to the methodological dimension

	Multimedia Dimension					
Ontology Name	***Multimedia***	***Audio***	***Video***	***Image***	***Visual***	***Audiovisual***
Multimedia Ontologies						
COMM	Yes	Yes	No	Yes	No	No
M3O	Yes	Yes	Yes	Yes	No	(*)[52]
Media Resource Ontology	Yes	Yes	Yes	(*)	No	(*)
MPEG-7 Upper MDS	Yes	Yes	Yes	Yes	No	Yes
MPEG-7 Tsinakari	Yes	Yes	No	Yes	Yes	No
MPEG-7 Rhizomik	Yes	Yes	No	Yes	Yes	No
MSO	Yes	No	Yes	Yes	(*)	Yes
Image and Shape Ontologies						
DIG 35	No	(*)	No	Yes	No	No
SAPO	No	No	No	Yes	Yes	No
CSO	No	No	(*)	Yes	Yes	No
MIRO	No	No	Yes	Yes	No	No
Visual Ontologies						
VRA Core 3	No	No	No	Yes	Yes	No
VDO	No	No	Yes	Yes	(*)	No
Music Ontologies						
Music Ontology	No	Yes	No	No	No	No
Kanzaki Music Ontology	No	Yes	No	No	No	No
Music Recommendation Ontology	No	Yes	No	No	No	No

Table 3 Comparison of ontologies with respect to the multimedia dimension

	Usability Profiling Dimension	
Ontology Name	***Quality of the documentation***	***Code Clarity***
Multimedia Ontologies		
COMM	High	High
M3O	Medium	High
Media Resource Ontology	High	High
MPEG-7 Upper MDS	Low	Low
MPEG-7 Tsinakari	Low	Medium
MPEG-7 Rhizomik	Low	Low
MSO	Medium	High
Image and Shape Ontologies		
DIG 35	High	High
SAPO	Medium	High
CSO	Medium	High
MIRO	Medium	High
Visual Ontologies		
VRA Core 3	High	Medium
VDO	High	Medium
Music Ontologies		
Music Ontology	High	Medium
Kanzaki Music Ontology	Medium	High
Music Recommendation Ontology	Low	Medium

Table 4 Comparison of ontologies with respect to the usability profiling dimension

[52] (*) stands for "cover more or less the domain".

	Reliability Dimension
Ontology Name	***Pitfalls***
Multimedia Ontologies	
COMM	Missing disjointness Missing domain or range in properties
M3O	Missing annotations
Media Resource Ontology	Missing annotations
MPEG-7 Upper MDS	Missing inverse relationships
MPEG-7 Tsinakari	Using different naming criteria along the ontology
MPEG-7 Rhizomik	Missing annotations Missing domain or range in properties Using different naming criteria along the ontology Using the same URI for different ontology elements
MSO	Merging different concepts in the same class Missing disjointness
Image and Shape Ontologies	
DIG 35	Missing annotations
SAPO	Missing annotations Using different naming criteria along the ontology
CSO	Merging different concepts in the same class Missing annotations
MIRO	Creating unconnected ontology elements Merging different concepts in the same class Using the same URI for different ontology elements
Visual Ontologies	
VRA Core 3	Using different naming criteria along the ontology Using in a non correct way ontology elements
VDO	Merging different concepts in the same class Missing annotations
Music Ontologies	
Music Ontology	Missing domain or range in properties
Kanzaki Music Ontology	Missing inverse relationships
Music Recommendation Ontology	Using different naming criteria along the ontology

Table 5 Comparison of ontologies with respect to the reliability dimension

5. Related Work

There are other comparative analyses of multimedia ontologies in the literature. One of these studies [10] presents a systematic survey of seven ontologies based on the MPEG-7 standard. In such a research work the ontologies were compared across two annotation dimensions that are (1) content structure descriptions and (2) linking with domain ontologies. These two dimensions are related at some point with the methodological and multimedia dimensions of FRAMECOMMON. Another important related work is the survey presented in [26]. This study compares four multimedia ontologies (Hunter's MPEG-7, DS-MIRF, Rhizomik, and COMM) with respect to the following three criteria: (1) how the ontologies are linked with the domain semantics, (2) the MPEG-7 coverage of the

multimedia ontology, and (3) the scalability and the modelling rationale of the conceptualization. In this case, the criteria used are partially related with the methodological, multimedia, and reliability dimensions of FRAMECOMMON.

To our knowledge there is no comparative study broader than the one presented in this paper, since we cover a wide range on multimedia ontologies developed during the last decade. In addition, other comparative studies do not take into account together the four dimensions of FRAMECOMMON. Finally, the main aim of our study is different from the aforementioned ones, because our purpose is to use the analysis for helping ontology practitioners in the selection of the most suitable multimedia ontologies to be reused.

6. Conclusions

In this paper we have described relevant ontologies developed in the last decade that aim to bridge the semantic gap in the multimedia field. We have presented important issues addressed by each multimedia ontology. We have first noticed the existence of many standards in multimedia and that the most used for implementing ontologies is MPEG-7.

It is worth stating that COMM proposal marked “a new vision” of developing multimedia ontologies by means of creating a modular design, using un upper ontology (DOLCE), and using ontology design patterns. Thus, COMM is an extensible ontology and allows an easy integration with domain ontologies. Hence, COMM marks an inflection point in multimedia ontology development.

It is important to realize that many works that came after COMM were focused on audio or music aspects; quite different from those works focused on image, audio or video developed before COMM. Moreover, recent efforts to have a generic multimedia ontology reusing existing multimedia standards and knowledge resources (including ODPs) and establishing mappings with multimedia formats are reflected in the M3O and the Media Resource Ontology, respectively.

We have also proposed a comparative framework, FRAMECOMMON, for contrasting ontologies in the multimedia domain. The main aim of this framework is to provide insights and guides on different features that may help ontology practitioners to select the most suitable multimedia ontology to be reused. FRAMECOMMON is divided into 4 dimensions: the methodological one that is oriented to the process used during the ontology building, and other 3 dimensions

(multimedia dimension, usability profiling dimension, and reliability dimension) oriented to the outcome, that is, the ontology itself.

Using this framework we have performed a comparative analysis of the 16 multimedia ontologies presented in this paper. We provide here the most interesting conclusions we have extracted from the comparative analysis.

With respect to the methodological dimension, we can mention that MPEG-7 is the most reused standard since it allows describing multimedia content at any level of granularity and using different levels of abstraction. In addition, in recent years the idea of reusing knowledge resources and performing mappings when developing multimedia ontologies is taking great importance. Ontologies that have being developed reusing well-developed ontological resources as well as those ontologies in which mappings have been established with available resources should be selected in the first place during the reuse task. The reason of this recommendation is that such ontologies allow spreading good practices and increasing the overall quality of ontological models.

Regarding multimedia aspects, we have classified the set of 16 ontologies into 4 categories having in mind the different multimedia types (audio, audiovisual, image, multimedia, and video). The categories are multimedia ontologies, image and shape ontologies, visual ontologies, and music ontologies. This classification can help practitioners to have an overview of the different aspects covered by the ontologies in the multimedia domain. To select the most suitable ontology to be reused for a particular purpose, human intervention is needed. The study performed with the 16 ontologies regarding the multimedia aspects coverage can help during such a human intervention.

Another important point to take into account when a practitioner needs to select an ontology for using them in an ontology building or in a semantic application is the understanding of such an ontology. This refers to the usability profiling dimension we have analyzed in the 16 ontologies presented in this paper. In this regard, ontologies obtained from an automatic transformation of MPEG-7 are less understandable than those developed reusing knowledge resources (such as COMM or the Media Resource Ontology).

Finally, it is well accepted that the evaluation of ontologies is a crucial activity to be performed before using or reusing ontologies in other ontology developments and/or in semantic applications. For this reason we performed the evaluation of

the multimedia ontologies with respect to a set of identified pitfalls. We suggest that soundly developed ontologies are better candidate for reuse. In this regard, it is important to mention that almost half of the ontologies used different naming criteria along the ontology and missed annotations, which makes difficult the understanding of the ontologies.

After applying FRAMECOMMON to the 16 ontologies in the multimedia domain presented in this paper, we can provide several advices to ontology practitioners in the task of selecting the most suitable ontology. This guidance is based on general representation requirements the practitioners have when developing multimedia ontologies. In those cases in which ontology practitioners need to describe in general multimedia objects, we recommend to reuse the Media Resource Ontology because (a) it is being developed within a W3C working group by consensus among its members; (b) it provides mappings with a variety of multimedia formats, which facilitates the interoperability; and (c) it is well documented, which benefits the ontology understanding. In addition, this ontology covers all the multimedia aspects except for the visual one. If ontology practitioners need to represent images and shapes, our suggestion is to reuse the DIG35 ontology that represents knowledge about digital images and is also well documented. In the case ontology practitioners are seeking for an ontology for describing visual resources, we suggest the use of VDO, because it reuses the standard MPEG-7 and it is aligned with DOLCE, which facilitates the integration with domain ontologies. In addition, VDO covers all the visible features (video, image, and visual). Finally, if ontology practitioners are interested in reusing an ontology about music, our advice is to use the Music Ontology, which has a good documentation and is reusing available knowledge resources.

As a final conclusion of our survey, we can mention that during this last decade a lot of efforts have been done in the development of multimedia ontologies. The trend in the present is to build ontologies in the multimedia domain by means of reusing and mapping available knowledge resources (ontologies, NORs, and ODPs) with the aim of (a) reducing the time and costs associated to the ontology development, (b) spreading good practices (from well-developed ontologies), and (c) increasing the overall quality of ontological models.

Acknowledgements. This work has been developed in the framework of the Spanish project BUSCAMEDIA (www.cenitbuscamedia.es), a CENIT-E project with reference number CEN-2009-1026 and funded by the Centre for the Development of Industrial Technology (CDTI). We would like to thank our partners in the project for their help.

References

1. Albertoni A, Papapleo L, Robbiano R, Spagnuolo M (2006) Towards a conceptualization for Shape Acquisition and Processing. In: Proceedings of the 1st International Workshop on Shapes and Semantics, Matsushima.
2. Arndt R, Troncy R, Staab S, Hardman L (2009) COMM: A Core Ontology for Multimedia Annotation. In Handbook on Ontologies, 2nd ed., Series: International Handbooks on Information Systems, Steffen Staab, Rudi Studer (Eds.), pages 403-421, 2009, Springer Verlag.
3. Arndt R, Troncy R, Staab S, Hardman L, and Vacura M (2007) COMM: Designing a Well-Founded Multimedia Ontology for the Web. 6th International Semantic Web Conference ISWC'2007), Busan, Korea, November 11-15, 2007.
4. Benitez AB, Rising H, Jörgensen C, Leonardi R, Bugatti A, Hasida K, Mehrotra R, Tekalp AM, Ekin A, Walker T (2002) Semantics of multimedia in MPEG-7. Proceedings of the IEEE International Conference on Image Processing, Rochester, New York, USA (2002).
5. Bloehdorn S, Petridis K, Saathoff C, Simou N, Tzouvaras V, Avrithis Y, Handschuh S, Kompatsiaris Y, Staab S and Strintzis MG (2005) Semantic Annotation of Images and Videos for Multimedia Analysis. 2nd European Semantic Web Conference, ESWC 2005, Heraklion, Greece, May 2005.
6. Bloehdorn S, Simou N, Tzouvaras V, Petridis K, Handschuh S, Avrithis Y, Kompatsiaris I, Staab S, Strintzis M (2004) Knowledge Representation for Semantic Multimedia Content analysis and Reasoning. Proceedings of European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology (EWIMT), London, U.K. 25-26 November 2004.
7. Carsten S, Ansgar S (2010) Unlocking the semantics of multimedia presentations in the web with the multimedia metadata ontology. Proceedings of the 19th International conference on World Wide Web, WWW '10, pages 831–840, New York, NY, USA, 2010. ACM.
8. Celma O (2006) Foafing the Music: Bridging the Semantic Gap in Music Recommendation. Semantic Web Challenge 2006.
9. Chang SF, Sikora T, Puri A (2001) Overview of the MPEG-7 standard. IEEE Transactions on Circuits and Systems for Video Technology 11(6) (2001) 688-695.
10. Dasiopoulou S, Tzouvaras V, Kompatsiaris I, Strintzis M (2010) Enquiring MPEG-7 based multimedia ontologies. Special Issue on Data Semantics for Multimedia Systems; Guest Editors: Mei-Ling Shyu, Yu Cao, Jun Kong, Ming Li, Mathias Lux and Jie Bao. In Journal Multimedia Tools and Applications. Volume 46, Numbers 2-3, 331-370. January 2010.

11. Digital Imaging Group (DIG), DIG35 Specification - Metadata for Digital Images - Version 1.0 August 30, 2000 (http://xml.coverpages.org/FU-Berlin-DIG35-v10-Sept00.pdf).
12. Gangemi A, Guarino N, Masolo C, Oltramari A, Schneider L (2002) Sweetening Ontologies with DOLCE. Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2002). Ontologies and the Semantic Web. Springer-Verlag London. ISBN: 3-540-44268-5.
13. García R and Celma O (2005) Semantic Integration and Retrieval of Multimedia Metadata. 5th Knowledge Markup and Semantic Annotation Workshop, SemAnnot 2005. CEUR Workshop Proceedings, Vol. 185, pp. 69-80, 2006 ISSN 1613-0073.
14. Halaschek-Wiener C, Golbeck J, Schain A, Grove M, Parsia B and Hendler J (2006) Annotation and provenance tracking in semantic web photo libraries. International provenance and annotation workshop (IPAW 2006).
15. Hunter J (2001) Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology. International Semantic Web Working Symposium (SWWS), Stanford, July 30 - August 1, 2001.
16. Moscato V, Penta A, Persia F, Picariello A (2010) MOWIS: A system for building Multimedia Ontologies from Web Information Sources. Proceedings of the 1st Italian Information Retrieval Workshop (IIR'10), pp 89-93, Padova, Italy, January 27-28, 2010. http://ims.dei.unipd.it/websites/iir10/index.html.
17. MPEG-7. Multimedia Content Description Interface. Standard No. ISO/IEC 15938, 2001.
18. Nack F and Lindsay AT (1999) Everything you wanted to know about MPEG-7 (Parts I & II). IEEE Multimedia, 6(3-4), 1999.
19. Pinto HS, Martins JP (2001) A methodology for ontology integration. Proceedings of the 1st International Conference On Knowledge Capture. Victoria, British Columbia, Canada. Pages: 131-138. 2001. ISBN: 1-58113-380-4.
20. Poveda-Villalón M, Suárez-Figueroa MC, Gómez-Pérez A (2010) A Double Classification of Common Pitfalls in Ontologies. Workshop on Ontology Quality (OntoQual 2010), Co-located with EKAW 2010, October 15, 2010, Lisbon, Portugal.
21. Poveda-Villalón M, Suárez-Figueroa MC, Gomez-Perez A (2010). Common Pitfalls in Ontology Development. Current Topics in Artificial Intelligence, CAEPIA 2009 Selected Papers.
22. Raimond Y, Jacobson F, Fazekas G, Gängler T, Reinhardt S (2010) Music Ontology Specification. Specification Document (14 February 2010). http://musicontology.com.
23. Simou N, Tzouvaras V, Avrithis Y, Stamou G, Kollias S (2005) A Visual Descriptor Ontology for Multimedia Reasoning. In Proceedings of Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS '05). Montreux, Switzerland, April 13-15, 2005.
24. Smeulders A, Worring M, Santini S, Gupta A, and Jain R (2000) Content-based image retrieval at the end of the early years. IEEE Transactions Pattern Analysis. Machine Intelligence, 22 (12):1349-1380, 2000.

25. Suárez-Figueroa MC (2010) NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse. PhD thesis. Universidad Politécnica de Madrid, 2010. http://oa.upm.es/3879/.
26. Troncy R, Celma O, Little S, Garcia R, Tsinaraki C (2007) MPEG-7 based Multimedia Ontologies: Interoperability Support or Interoperability Issue? In International Workshop on Multimedia Annotation and Retrieval enabled by Shared Ontologies (MAReSO), p. 2–15.
27. Troncy R, Bailer W, Hausenblas M, Hofmair P, Schlatte R (2006) Enabling Multimedia Metadata Interoperability by Defining Formal Semantics of MPEG-7 Profiles. In 1st International Conference on Semantics And digital Media Technology (SAMT'06), pages 41–55, Athens, Greece, 2006.
28. Tsinaraki C, Polydoros P and Christodoulakis S (2004) Interoperability support for Ontology-based Video Retrieval Applications. Proceedings of the 3rd International Conference on Image and Video Retrieval (CIVR 2004), pp. 582-591, Dublin, Ireland, 21-23 July 2004.
29. Vasilakis G, Garcia-Rojas A, Papaleo L, Catalano C, Spagnuolo M, Robbiano F, Vavalis M, Pitikakis M (2007) A common ontology for multi-dimensional shapes. SAMT 2007 2nd International Conference on Semantic and Digital Media, MAReSO Workshop Proceedings, pp. 31– 43.
30. Visual Resources Association Data Standards Committee. VRA Core Categories, Version 3.0. 20/2/2002. http://www.vraweb.org/vracore3.htm.