

---

# **Towards a Common Multimedia Ontology Framework**

**(Analysis of the Contributions to Call for a Common multimedia  
Ontology Framework Requirements)**

---

**Hariklia Eleftherohorinou, Vasiliki Zervaki, Anastasios Gounaris,  
Vasileios Papastathis, Yiannis Kompatsiaris (ITI) and Paola Hobson  
(MOTOROLA Labs)**

**with contributions from:  
(List of contributions in Section 2)**

Version:	1.0
Date:	25-04-2006
State:	
Distribution:	

## Table of Contents

<b>Table of Contents</b> .....	<b>2</b>
<b>List of Acronyms and Abbreviations</b> .....	<b>3</b>
<b>1 Introduction</b> .....	<b>4</b>
<b>2 Responses to the call of requirements</b> .....	<b>4</b>
<b>3 Existing Ontologies</b> .....	<b>6</b>
3.1 Ontologies describing Content Structure.....	6
3.2 Specific Domain Ontologies .....	9
3.3 Multimedia Upper Ontologies.....	13
3.4 Multimedia Core Ontology .....	18
<b>4 Requirements</b> .....	<b>20</b>
4.1 Scope and usage .....	20
4.2 Media description - Description of the Information Object .....	20
4.3 Representation of Content structure .....	21
4.4 Representation of Content Description .....	22
4.5 Other Concepts Need to be Covered by the Ontology.....	23
4.6 Knowledge representation formalism .....	23
4.7 Reasoning support.....	24
4.8 Following an established methodology .....	24
4.9 Other requirements .....	24
<b>5 Harmonization approaches</b> .....	<b>26</b>
5.1 Modularization .....	26
5.2 Ontology cross-domain linking .....	26
5.3 Harmonization of core ontologies.....	27
5.4 Development of a common upper ontology .....	27
5.5 Establishment of a common upper model.....	27
5.6 Compliance to standards .....	28
<b>6 Conclusions</b> .....	<b>28</b>

## List of Acronyms and Abbreviations

3GPP	3 <sup>rd</sup> Generation Partnership Project
API	Application Program Interface
CCM	Commercial Content Management
CC/PP	Composite Capabilities/Preference Profiles
CG&V	Computer Graphics & Vision
CIDOC	International Committee for Documentation
CIDOC CRM	CIDOC Conceptual Reference Model
DOLCE	Descriptive Ontology for Linguistic and Cognitive Engineering
FaCT	Inference Engine. Full name: FaCT - Fast Classification of Terminologies
H-ANIM	Humanoid Animation
IPTC	International Press Telecommunications Council
JENA	Java framework for building Semantic Web applications
MPEG-7	Motion Picture Experts Group
MPEG-7 MDS	MPEG-7 Multimedia Description Schemes
MSO	Multimedia Structure Ontology
NCBI	National Centre for Biotechnology Information
OMA	Open Mobile Alliance
OWL	Web Ontology Language
OWL-DL	OWL with Description Logics
OWL-QL	The OWL Query Language
PCS	Personal Content Management
RACER	A Core Inference Engine for the Semantic Web
RDBMS	Relational Database Management System
RDF	Resource Description Framework
RDFS	RDF Schema
SABO	Standard Animal Behaviour Ontology
SEKT	Semantically Enabled Knowledge Technologies.
SQL	Structured Query Language
SUMO	Suggested Upper Merged Ontology
ULAN	Union List of Artist Names
VDO	Visual Descriptor Ontology
VRA	Visual Resources Association
VRML	Virtual Reality Modelling Language
W3C	World Wide Web Consortium
XML	Extensible Markup Language

## 1 Introduction

The increasing number of multimedia ontologies being used in various projects for multimedia processing, raises the need for a common, share-able harmonized multimedia ontology, and the question as to whether such an ontology is feasible. The approach for harmonization that we are following aims at defining a set of core features and specifications, which is a "minimum set" for a multimedia ontology to be used as a basis for all multimedia related applications to build on and extend. To this end, a call for submitting requirements reports was made in December 2005, to which there were 14 responses. Both the call and the contributions can be found at the following address: [http://www.acemedia.org/aceMedia/reference/multimedia\\_ontology/index.html](http://www.acemedia.org/aceMedia/reference/multimedia_ontology/index.html)

The purpose of this report is to summarise and analyse the 14 contributions to the call. The contributions belong to two main categories: firstly, contributions that describe specific, existing ontologies and secondly, contributions that discuss requirements. Note that the contributions of the first category may also contain requirements (in a more indirect manner). The requirements themselves can also be distinguished between the requirements that refer to ontology properties and features, and those that refer to harmonisation approaches.

The structure of the report is as follows. Section 2 enumerates and provides the references for the contributions. Section 3 covers the existing ontologies described within the contributions and presents them in a unified manner with a view to making more explicit which exactly features they implement and what they can be used for. The synopsis of the requirements for ontologies is discussed in Section 4. Section 5 deals with the harmonisation of multimedia ontologies and presents the various proposals towards this objective. Finally, the conclusions are in Section 6.

## 2 Responses to the call of requirements

The responses to the call of requirements are sufficient to acquire a clear view of what has been done and what is needed in the field of multimedia ontologies. The contributions are listed below.

- [1] Yannis Avrithis, National Technical University of Athens, "Requirements for Multimedia Analysis & Annotation"
- [2] Michela Spagnuolo, Istituto per la Matematica Applicata e le Tecnologie Informatiche, Consiglio Nazionale delle Ricerche, AIM@SHAPE Consortium, [AIM@SHAPE Ontology, Deliverable 1.2.2](#)
- [3] Christine Golbreich, M. Bouet, University Rennes, "Multimedia Reasoning with Medical Images: Mammography interpretation and therapeutic decisions"
- [4] Samer Abdallah, Centre for Digital Music, Queen Mary, University of London, "Information management for music analysis systems"

## Towards a Common Multimedia Ontology Framework

- [5] Gerald Jaschke, Matthias Hemmje, FernUniversität in Hagen, Lehrgebiet Multimedia und Internetanwendungen, Informatikzentrum, Semantic "Access, Coordination, and Orchestration of Information Visualization and Visual Data Mining Services ISO/FDIS 21127"
- [6] Patrick Sinclair, University of Southampton, UK, "Proposal for Multimedia Ontologies: ISO/FDIS 21127"
- [7] Vangelis Karkaletsis, Inst. of Informatics & Telecommunications, N.C.S.R. "Demokritos", "Multimedia Information Extraction and Semantic Annotation"
- [8] Georg Thallinger, Institute of Information Systems & Information Management, JOANNEUM RESEARCH Forschungsgesellschaft mbH, "Multimedia Ontologies for annotation and retrieval tasks"
- [9] Carsten Saathoff, Simon Schenk, Steffen Staab, ISWeb, University of Koblenz, Requirements for a Multimedia Ontology Framework
- [10] Dipl.-Inf. Lyndon J B Nixon, Freie Universität, Netzbasierte Informationssysteme, Institut fuer Informatik, "Semantic Web-enabled Multimedia Presentation system"
- [11] Ovidio Salvetti, 'Signals and Images' Laboratory, Institute of Information Science and Technologies "A. Faedo" (ISTI-CNR), Italian National Research Council (CNR), "MUSCLE NoE Ontology"
- [12] David M. Shotton, Image Bioinformatics Research Group, Department of Zoology, University of Oxford, "Requirements for Scientific Images"
- [13] Chrisa Tsinaraki, Technical University of Crete, "MPEG-7 MDS Upper Ontology, Semantic User Preference Ontology and 3D Ontology"
- [14] Raphael Troncy, Jacco van Ossenbruggen, CWI Amsterdam, Christian Halaschek-Wiener, Maryland Information and Network Dynamics Lab, Giorgos Stamou, Vassilis Tzouvaras, National Technical University of Athens, "Multi-Perspective Requirements for a Common Multimedia Ontology Framework"
- [15] Yulia Bacharovka, University of Twente, "Towards Common Multimedia Ontology Framework-Introducing Modality Perspective"

### 3 Existing Ontologies

In order to understand better the requirements and whether they have been implemented to a certain degree already by any initiatives, it would be useful to present all ontology contributions that are tangible. Thus, in this section, existing ontologies are divided in three groups, according to their domain or generally to their framework. For each ontology, we examine several dimensions: its scope (in which kind of multimedia processing they can be used), which type of digital artefacts it supports, the structure of multimedia content assumed, the domain in which it is applied, the modularity, the representation language and the use of reasoning, and aspects such as specific user requirements and known tools and projects that have employed the relevant ontology.

#### 3.1 Ontologies describing Content Structure

The ontologies in this category focus on the description of multimedia content structure.

##### ■ aceMedia Ontology Framework [1]

<b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modeling, storage etc.)</b>	<input type="checkbox"/> Annotation <input type="checkbox"/> Analysis <input type="checkbox"/> Use Of multimedia content
<b>Media description - Description of the Information Object Covered by the Ontology</b>	<input type="checkbox"/> Still images & video
<b>Content Structure Covered by the Ontology</b>	<input type="checkbox"/> MPEG – 7 <ul style="list-style-type: none"> <li>• VDO (Visual Descriptor Ontology)</li> <li>• MSO (Multimedia Structure Ontology)</li> </ul> <input type="checkbox"/> Spatio – Temporal extensions
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	<input type="checkbox"/> PCS (Personal Content Management) <ul style="list-style-type: none"> <li>• Holiday</li> <li>• Family</li> </ul> <input type="checkbox"/> CCM (Commercial Content Management) <ul style="list-style-type: none"> <li>• Motorsports</li> <li>• Tennis</li> </ul>
<b>Other Concepts Covered by the Ontology</b>	
<b>Other (upper) ontologies used - modularity</b>	Multimedia Ontologies <ul style="list-style-type: none"> <li>□ VDO – EXT (Visual Annotation Ontology)</li> </ul> Domain Ontologies <ul style="list-style-type: none"> <li>□ PCS (Personal Content Management) Ontologies</li> <li>□ CCM (Commercial Content Management) Ontologies</li> </ul> Core Ontology <ul style="list-style-type: none"> <li>□ DOLCE</li> </ul>
<b>Language - Reasoning Support</b>	<input type="checkbox"/> RDFS <input type="checkbox"/> OWL (DOLCE & Spatio –temporal)

## Towards a Common Multimedia Ontology Framework

	extensions) To be used in other parts <input type="checkbox"/> Appropriate sub – language of OWL
<b>User preferences, specific constraints taken into account, etc.</b>	
<b>Known Tools using the ontology</b>	<input type="checkbox"/> M – OntoMat – Annotizer
<b>Known projects using the ontology</b>	<input type="checkbox"/> aceMedia <a href="http://www.acemedia.org/aceMedi/">http://www.acemedia.org/aceMedi/</a>
<b>Other info</b>	

### ■ AIM@SHAPE Ontology.

*Ontology Description:* AIM@SHAPE Ontology is an ontology developed for representing, modelling and processing knowledge, which derives from digital shapes. The term digital shape involves every visible object that possesses some space, and that is digitally stored [2].

<b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modeling, storage etc.)</b>	<input type="checkbox"/> Retrieval of shape models, structures and tools <input type="checkbox"/> Analysis of shape models <input type="checkbox"/> Annotation of shape models, structures and tools <input type="checkbox"/> Reasoning about shape models, structures and tools
<b>Media description - Description of the Information Object Covered by the Ontology</b>	<input type="checkbox"/> Images (2D and 3D) <input type="checkbox"/> Videos <input type="checkbox"/> 3D models of shapes <input type="checkbox"/> Animation of 3D objects
<b>Content Structure Covered by the Ontology</b>	<input type="checkbox"/> Segmentation of 3D shapes <input type="checkbox"/> Skeletonization of 3D shapes <input type="checkbox"/> H-ANIM, for the description of virtual humans
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	<input type="checkbox"/> Virtual Humans <input type="checkbox"/> Product Design <input type="checkbox"/> Shape acquisition and processing
<b>Other Concepts Covered by the Ontology</b>	<input type="checkbox"/> Tools <input type="checkbox"/> Workflows of interest for researchers working in the Computer Graphics & Vision (CG&V) field <input type="checkbox"/> Acquisition devices for the digitalisation of 3D shapes
<b>Other (upper) ontologies used - modularity</b>	<input type="checkbox"/> AIM@SHAPE common ontology for shapes, tools and workflows
<b>Language - Reasoning Support</b>	<input type="checkbox"/> OWL <input type="checkbox"/> Racer for reasoning <input type="checkbox"/> User interface to ease the building of nRQL queries
<b>User preferences, specific constraints taken into account, etc.</b>	The resources are covered by a licencing scheme that is inspired by the Creative Commons <sup>1</sup> A user policy is adopted to give different user privileges for downloading, uploading etc.

<sup>1</sup> <http://creativecommons.org/licenses/by-nc-sa/2.0/>

## Towards a Common Multimedia Ontology Framework

<b>Known Tools using the ontology</b>	
<b>Known projects using the ontology</b>	AIM@SHAPE is a Network of Excellence project within the EC's IST Sixth Framework Programme
<b>Other info</b>	AIM@SHAPE has already produced an ontology for the geometric representation of shapes, which can be used as a basis for the 3D part of the "minimum set"

### ■ Music Information Ontology.

*Ontology Description:* Music Information Ontology is an ontology developed for music information management [4].

<b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modelling, storage etc.)</b>	<input type="checkbox"/> Annotation <input type="checkbox"/> Retrieval <input type="checkbox"/> Reasoning of music content
<b>Media description - Description of the Information Object Covered by the Ontology</b>	<input type="checkbox"/> Sound <input type="checkbox"/> Scores
<b>Content Structure Covered by the Ontology</b>	<input type="checkbox"/> Spatial information <input type="checkbox"/> Temporal information
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	<input type="checkbox"/> Description of temporal phenomena (music is a time-based phenomenon) <input type="checkbox"/> Time <input type="checkbox"/> Location <input type="checkbox"/> Human agency (representation of people and groups) <input type="checkbox"/> Instruments
<b>Other Concepts Covered by the Ontology</b>	<input type="checkbox"/> Evaluation (statistics) <input type="checkbox"/> Mode (functional representation of relations & attributes) <input type="checkbox"/> People <input type="checkbox"/> Physical Events <input type="checkbox"/> Signals and processing (not just music) <input type="checkbox"/> Relationship between information Objects <input type="checkbox"/> Physical manifestations of information objects (printed score/physical sound) <input type="checkbox"/> Probabilistic models
<b>Other (upper) ontologies used - modularity</b>	<input type="checkbox"/> Event Ontology <input type="checkbox"/> Computation Ontology <input type="checkbox"/> MusicBrainz ontology (semantic-web service) Upper Ontology <input type="checkbox"/> Dublin core ontology
<b>Language - Reasoning Support</b>	<input type="checkbox"/> OWL-DL for ontologies and <input type="checkbox"/> XSB-Prolog for DL-reasoning
<b>User preferences, specific constraints taken into account, etc.</b>	
<b>Known Tools using the ontology</b>	
<b>Known projects using the ontology</b>	

<p><b>Other info</b></p>	<ul style="list-style-type: none"> <li>❑ Storage of models in RDBMS accessed via SQL or RDF managed by JENA</li> <li>❑ Computational engines             <ul style="list-style-type: none"> <li>• Matlab as via extension to SWI Prolog</li> <li>• Octave</li> <li>• LISP</li> <li>• Java</li> <li>• Haskell</li> </ul> </li> </ul>
--------------------------	---

### 3.2 Specific Domain Ontologies

The ontologies that belong to this class have been created to serve a particular domain that is described, in each case, in the "Content Description".

#### ■ Medical Image Domain Ontology.

*Ontology Description:* The present Ontology is an ontology developed for multimedia reasoning with medical images in order to interpret mammography images and assist diagnosis [3].

<p><b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modelling, storage etc.)</b></p>	<ul style="list-style-type: none"> <li>❑ Instance Retrieval of images</li> <li>❑ Text retrieval</li> </ul>
<p><b>Media description - Description of the Information Object Covered by the Ontology</b></p>	<ul style="list-style-type: none"> <li>❑ Image</li> <li>❑ Text</li> </ul>
<p><b>Content Structure Covered by the Ontology</b></p>	<ul style="list-style-type: none"> <li>❑ MPEG 7 Descriptors, e.g. ,             <ul style="list-style-type: none"> <li>• Colour</li> <li>• Shape</li> <li>• Texture</li> </ul> </li> <li>Of images or regions             <ul style="list-style-type: none"> <li>❑ Extend current VDO with                 <ul style="list-style-type: none"> <li>• Shape 3D Descriptors</li> <li>• SphericShape3D</li> <li>• EggShaped3D, etc</li> </ul> </li> <li>❑ Spatial information</li> </ul> </li> </ul>
<p><b>Content Description Covered by the Ontology – Linking to specific Domains</b></p>	<p>Breast Imaging and Reporting</p>
<p><b>Other Concepts Covered by the Ontology</b></p>	
<p><b>Other (upper) ontologies used - modularity</b></p>	<p>VDO (aceMedia Visual Descriptor Ontology v9.0)</p>
<p><b>Language - Reasoning Support</b></p>	<ul style="list-style-type: none"> <li>❑ OWL – DL</li> <li>❑ Reasoning using DL inference</li> <li>❑ Fuzzy Logics</li> </ul> <p>Plan to apply</p> <ul style="list-style-type: none"> <li>❑ Case Based Reasoning – Case base is needed – presented as OWL DL instances</li> </ul>

## Towards a Common Multimedia Ontology Framework

User preferences, specific constraints taken into account, etc.	
Known Tools using the ontology	
Known projects using the ontology	
Other info	<ul style="list-style-type: none"> <li><input type="checkbox"/> Domain ontology (ACR) based on BI-BRADS lexicon</li> <li><input type="checkbox"/> Racer is suggested</li> </ul>

### ■ NM2 ontology.

*Ontology Description:* NM2 Ontology is an ontology developed for creating personalised versions of narratives *in conjunction with a rule-based component* [8].

Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modeling, storage etc.)	<ul style="list-style-type: none"> <li><input type="checkbox"/> Automatic and manual annotation</li> <li><input type="checkbox"/> Retrieval</li> <li><input type="checkbox"/> Contextual filtering and matching of multimedia content</li> </ul>
Media description - Description of the Information Object Covered by the Ontology	Multimedia programmes
Content Structure Covered by the Ontology	MPEG-7 (audio, video, images)
Content Description Covered by the Ontology – Linking to specific Domains	Description of intrinsic MM-items features, genre, objects (actors, behaviour, etc.) and wrapper for narrative structures
Other Concepts Covered by the Ontology	<ul style="list-style-type: none"> <li><input type="checkbox"/> The story</li> <li><input type="checkbox"/> The plot and</li> <li><input type="checkbox"/> The single media objects</li> </ul>
Other (upper) ontologies used - modularity	<ul style="list-style-type: none"> <li><input type="checkbox"/> Multimedia ontologies will be utilised</li> <li><input type="checkbox"/> Core Ontology, that will be extended</li> <li><input type="checkbox"/> Production – specific ontologies</li> </ul>
Language - Reasoning Support	<ul style="list-style-type: none"> <li><input type="checkbox"/> OWL-DL</li> <li><input type="checkbox"/> Prolog</li> </ul>
User preferences, specific constraints taken into account, etc.	
Known Tools using the ontology	
Known projects using the ontology	NM2
Other info	<ul style="list-style-type: none"> <li><input type="checkbox"/> OWL Data Store for relevant Ontologies</li> <li><input type="checkbox"/> A description tool (management, manipulating, exploring content)</li> <li><input type="checkbox"/> Content analysis tools (automated content extraction)</li> </ul>

### ■ MEPCO (Media Presence and Campaign Ontology).

*Ontology Description:* MEPCO is an ontology developed for cross relating and linking media campaigns [8].

Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-	Semantic annotation and generally knowledge management
--	--

## Towards a Common Multimedia Ontology Framework

modelling, storage etc.)	
<b>Media description - Description of the Information Object Covered by the Ontology</b>	Multimedia - TV, internet
<b>Content Structure Covered by the Ontology</b>	MPEG-7 (audio, video, images)
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	<input type="checkbox"/> Media TV <input type="checkbox"/> Press and <input type="checkbox"/> Internet
<b>Other Concepts Covered by the Ontology</b>	
<b>Other (upper) ontologies used - modularity</b>	PROTON upper-level ontology ( <a href="http://proton.semanticweb.org/">http://proton.semanticweb.org/</a> )
<b>Language - Reasoning Support</b>	Standards for media-related metadata such as <ul style="list-style-type: none"> <li><input type="checkbox"/> NewsML and</li> <li><input type="checkbox"/> News Codes</li> </ul> from IPTC
<b>User preferences, specific constraints taken into account, etc.</b>	
<b>Known Tools using the ontology</b>	KIM platform
<b>Known projects using the ontology</b>	<input type="checkbox"/> MediaCampaign <input type="checkbox"/> SEKT project
<b>Other info</b>	

### ■ ImageStore Ontology – BioImage Database.

*Ontology Description:* ImageStore Ontology is an ontology developed for BioImage Database, an ontology – driven image database created to provide a free open – access database of digital images that are concerned with biological research themes [12].

<b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modelling, storage etc.)</b>	<input type="checkbox"/> Annotation of the images <input type="checkbox"/> Physical storage in BioImage Database <input type="checkbox"/> Search and content retrieval in BioImage Database
<b>Media description - Description of the Information Object Covered by the Ontology</b>	<input type="checkbox"/> Multidimensional images, 2D, a 3D, a video, or a 4D (x, y, z, time) <input type="checkbox"/> Details of image formation, capture and subsequent digital processing <input type="checkbox"/> Details of the image itself (dimensions, zoom factor etc)
<b>Content Structure Covered by the Ontology</b>	<input type="checkbox"/> Uses a subset of the class model of the Advanced Authoring Format ( <a href="http://sourceforge.net/projects/aaf">sourceforge.net/projects/aaf</a> and <a href="http://www.aafassociation.org">www.aafassociation.org</a> ) to describe media objects <input type="checkbox"/> Uses a subset of MPEG-7 to describe multimedia content
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	<input type="checkbox"/> Biological specimen <input type="checkbox"/> Description and interpretation of what the image represents
<b>Other Concepts Covered by the Ontology</b>	<input type="checkbox"/> Experimental procedure <input type="checkbox"/> People, institutions and funding agencies

## Towards a Common Multimedia Ontology Framework

	<ul style="list-style-type: none"> <li>❑ Curation and provenance of the image</li> </ul>
<b>Other (upper) ontologies used - modularity</b>	<p>Third party domain-specific ontologies imported to describe biological aspects of the subjects on the images</p> <ul style="list-style-type: none"> <li>❑ NCBI taxonomy of biological species</li> <li>❑ Gene Ontology – provides controlled vocabulary to describe gene and gene product attributes in any organism</li> </ul> <p>An upper ontology</p> <ul style="list-style-type: none"> <li>❑ DOLCE</li> </ul>
<b>Language - Reasoning Support</b>	<ul style="list-style-type: none"> <li>❑ Ontology language – OWL-DL</li> <li>❑ In the BioImage Database metadata stored in RDF</li> <li>❑ DL reasoner - Racer</li> </ul>
<b>User preferences, specific constraints taken into account, etc.</b>	Semantic Web Standards supported by W3C
<b>Known Tools using the ontology</b>	
<b>Known projects using the ontology</b>	Developed within EC- funded ORIEL Project (Online Research Information Environment for the Life Sciences).
<b>Other Info</b>	Ontology development environment: Protégé with the OWL plug-in

### ■ Soccer Domain Ontology

*Ontology Description:* An ontology developed for assisting the (manual or automatic) annotation, retrieval and filtering of audiovisual content containing (parts of) soccer games [13].

<b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, metamodeling, storage etc.)</b>	<p>Currently</p> <ul style="list-style-type: none"> <li>❑ Annotation</li> <li>❑ Retrieval</li> <li>❑ Filtering</li> </ul> <p>Plan to be used for</p> <ul style="list-style-type: none"> <li>❑ Analysis</li> </ul>
<b>Media description - Description of the Information Object Covered by the Ontology</b>	
<b>Content Structure Covered by the Ontology</b>	Extends the MPEG-7 Upper Ontology with knowledge from the soccer domain.
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	Soccer games
<b>Other Concepts Covered by the Ontology</b>	
<b>Other (upper) ontologies used - Modularity</b>	MPEG-7 Upper Ontology
<b>Language - Reasoning Support</b>	OWL-DL
<b>User preferences, specific constraints taken into account, etc.</b>	
<b>Known Tools using the ontology</b>	GraphOnto component,
<b>Known projects using the ontology</b>	DELOS II Network of Excellence

Other info	
------------	--

### ■ Formula 1 Domain Ontology

*Ontology Description:* An ontology developed for assisting the (manual or automatic) annotation, retrieval and filtering of audiovisual content containing (parts of) Formula 1 races [13].

Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, metamodeling, storage etc.)	Currently <input type="checkbox"/> Annotation <input type="checkbox"/> Retrieval <input type="checkbox"/> Filtering Plan to be used for <input type="checkbox"/> Analysis
Media description - Description of the Information Object Covered by the Ontology	
Content Structure Covered by the Ontology	Extends the MPEG-7 Upper Ontology with knowledge from the Formula 1 domain.
Content Description Covered by the Ontology – Linking to specific Domains	Formula 1
Other Concepts Covered by the Ontology	
Other (upper) ontologies used - Modularity	MPEG-7 Upper Ontology
Language - Reasoning Support	OWL-DL
User preferences, specific constraints taken into account, etc.	
Known Tools using the ontology	GraphOnto component,
Known projects using the ontology	DELOS II Network of Excellence
Other info	

### 3.3 Multimedia Upper Ontologies

The seven (7) ontologies that follow have been created for a general use, thus they are characterized as upper, in order to make multimedia-handling procedures more homogeneous. In each case, this is described in the "Ontology Description".

#### ■ SWeMPs ontology.

*Ontology Description:* SWeMPs ontology is seen as an upper ontology which exists at a higher level than the individual subject, resource and service metadata ontologies [10].

Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-	<input type="checkbox"/> Meta – modelling <input type="checkbox"/> Reasoning for Multimedia processes (decision making)
--	--

## Towards a Common Multimedia Ontology Framework

modelling, storage etc.)	
<b>Media description - Description of the Information Object Covered by the Ontology</b>	N/A
<b>Content Structure Covered by the Ontology</b>	
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	
<b>Other Concepts Covered by the Ontology</b>	Multimedia generation process
<b>Other (upper) ontologies used - modularity</b>	
<b>Language - Reasoning Support</b>	OWL-DL
<b>User preferences, specific constraints taken into account, etc.</b>	
<b>Known Tools using the ontology</b>	
<b>Known projects using the ontology</b>	
<b>Other info</b>	Protégé

### ■ ZyX ontology.

*Ontology Description:* ZyX Ontology is an upper ontology developed for multimedia metamodeling [10].

Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modelling, storage etc.)	<ul style="list-style-type: none"> <li><input type="checkbox"/> Meta – modelling</li> <li><input type="checkbox"/> Reasoning for Multimedia processes (recommending)</li> </ul>
<b>Media description - Description of the Information Object Covered by the Ontology</b>	N/A
<b>Content Structure Covered by the Ontology</b>	
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	Ontological description of an abstract multimedia presentation model
<b>Other Concepts Covered by the Ontology</b>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Intends to model an abstract multimedia model</li> </ul>
<b>Other (upper) ontologies used - modularity</b>	
<b>Language - Reasoning Support</b>	OWL full
<b>User preferences, specific constraints taken into account, etc.</b>	
<b>Known Tools using the ontology</b>	
<b>Known projects using the ontology</b>	
<b>Other info</b>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Protégé</li> <li><input type="checkbox"/> It is based on the ZyX model specification by Suzanne Boll and Wolfgang Klas</li> </ul>

### ■ Salero ontologies.

*Ontology Description:* Salero Ontology is an upper ontology developed for Media Production. It aims to develop multimedia objects with context – aware

## Towards a Common Multimedia Ontology Framework

behaviour in order to be re-used by artists who want to create complex and compelling media [8].

<b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modelling, storage etc.)</b>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Annotation (basis for creation and manipulation of multimedia object)</li> <li><input type="checkbox"/> Retrieval</li> <li><input type="checkbox"/> Reasoning</li> <li><input type="checkbox"/> Cross-platform delivery</li> </ul>
<b>Media description - Description of the Information Object Covered by the Ontology</b>	Images, sounds, videos, characters, movements, behaviours
<b>Content Structure Covered by the Ontology</b>	Uses essence internal formats to describe low-level 'physical' features, MPEG-7 to describe low-level semantic features, WSML or OWL to describe high-level semantic features and WSML or SWRL to describe context-information
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	Multilingual media content
<b>Other Concepts Covered by the Ontology</b>	
<b>Other (upper) ontologies used - modularity</b>	Ontology API's
<b>Language - Reasoning Support</b>	WSML family or OWL family, depending on the exact requirements
<b>User preferences, specific constraints taken into account, etc.</b>	Base low-level description on MPEG-7
<b>Known Tools using the ontology</b>	
<b>Known projects using the ontology</b>	Salero
<b>Other info</b>	Suite of tools: <ul style="list-style-type: none"> <li><input type="checkbox"/> Ontology workbench (editor, versioning, alignment, etc.)</li> <li><input type="checkbox"/> Ontology management system (includes a scalable repository for ontology data)</li> </ul>

### ■ MPEG – 7 Upper Ontology.

*Ontology Description:* The MPEG – 7 Ontology is an ontology developed for supporting knowledge-based multimedia applications so as to capture all the MPEG – 7 MDS [13].

<b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modelling, storage etc.)</b>	Currently <ul style="list-style-type: none"> <li><input type="checkbox"/> Annotation</li> <li><input type="checkbox"/> Retrieval</li> <li><input type="checkbox"/> Filtering</li> </ul> Plan to be used for <ul style="list-style-type: none"> <li><input type="checkbox"/> Analysis</li> </ul>
<b>Media description - Description of the Information Object Covered by the Ontology</b>	
<b>Content Structure Covered by the</b>	<input type="checkbox"/> MPEG – 7

## Towards a Common Multimedia Ontology Framework

<b>Ontology</b>	<ul style="list-style-type: none"> <li>• Visual parts</li> <li>• Audio parts</li> </ul> <input type="checkbox"/> MPEG – 7 MDS
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	
<b>Other Concepts Covered by the Ontology</b>	
<b>Other (upper) ontologies used - modularity</b>	<input type="checkbox"/> User Preference Ontology Plan: <ul style="list-style-type: none"> <li><input type="checkbox"/> The use of MPEG 21 DIA Ontology</li> <li><input type="checkbox"/> Harmonization with SUMO core ontology</li> <li><input type="checkbox"/> Harmonization with DOLCE core ontology</li> </ul>
<b>Language - Reasoning Support</b>	OWL DL
<b>User preferences, specific constraints taken into account, etc.</b>	
<b>Known Tools using the ontology</b>	GraphOnto
<b>Known projects using the ontology</b>	DELOS II Network of Excellence
<b>Other info</b>	

### ■ Semantic User Preference Ontology.

*Ontology Description:* Semantic User Preference Ontology is an ontology developed to be used in conjunction the MPEG-7 MDS Ontology and with domain ontologies, in order to interoperate with MPEG-7 and allow domain knowledge utilization [13].

<b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modelling, storage etc.)</b>	Currently <ul style="list-style-type: none"> <li><input type="checkbox"/> Retrieval</li> <li><input type="checkbox"/> Filtering</li> </ul> Plan to be used for <ul style="list-style-type: none"> <li><input type="checkbox"/> Summarization</li> </ul>
<b>Media description - Description of the Information Object Covered by the Ontology</b>	
<b>Content Structure Covered by the Ontology</b>	<input type="checkbox"/> MPEG – 7 MDS with <input type="checkbox"/> Semantic User Preference Description
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	Semantic User Preference Model
<b>Other Concepts Covered y the Ontology</b>	
<b>Other (upper) ontologies used - modularity</b>	<input type="checkbox"/> MPEG – 7 Upper Ontology
<b>Language - Reasoning Support</b>	OWL DL
<b>User preferences, specific constraints taken into account, etc.</b>	
<b>Known Tools using the ontology</b>	GraphOnto
<b>Known projects using the ontology</b>	DELOS II Network of Excellence
<b>Other info</b>	

### ■ Ontology X3D.

## Towards a Common Multimedia Ontology Framework

*Ontology Description:* Ontology X3D is an ontology developed for semantic – driven 3D visualization [13].

<b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modelling, storage etc.)</b>	Semantic content visualization
<b>Media description - Description of the Information Object Covered by the Ontology</b>	
<b>Content Structure Covered by the Ontology</b>	
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	
<b>Other Concepts Covered by the Ontology</b>	<ul style="list-style-type: none"> <li><input type="checkbox"/> X3D Standard (VRML successor)</li> <li><input type="checkbox"/> Graphics and</li> <li><input type="checkbox"/> Virtual reality concepts</li> </ul>
<b>Other (upper) ontologies used – modularity</b>	Plan: <ul style="list-style-type: none"> <li><input type="checkbox"/> Harmonization with SUMO core ontology</li> <li><input type="checkbox"/> Harmonization with DOLCE core ontology</li> </ul>
<b>Language - Reasoning Support</b>	OWL
<b>User preferences, specific constraints taken into account, etc.</b>	
<b>Known Tools using the ontology</b>	
<b>Known projects using the ontology</b>	
<b>Other info</b>	Interoperation with domain ontologies

### ■ Modality Ontology

*Ontology Description:* It has been developed in order to represent all existing modalities, their capabilities to represent specific information features or types of information and their optimal combination that can best convey the intended message [15].

<b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modelling, storage etc.)</b>	Support the processes of selecting the most appropriate modality combinations that can represent particular types of information
<b>Media description - Description of the Information Object Covered by the Ontology</b>	All existing modalities hierarchically represented according to Bernsen's taxonomy.
<b>Content Structure Covered by the Ontology</b>	Compositional dependencies between modalities.
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	Does not model the content but the form in terms of suitability to represent certain types of content.

## Towards a Common Multimedia Ontology Framework

<b>Other Concepts Covered by the Ontology</b>	
<b>Other (upper) ontologies used – modularity</b>	Can be aligned with existing ontologies that represent content thus providing possibilities to model the content – form relationships.
<b>Language - Reasoning Support</b>	
<b>User preferences, specific constraints taken into account, etc.</b>	User preferences, characteristics of the situation, goals of the presentation, discourse structure of the presentation can be taken into consideration
<b>Known Tools using the ontology</b>	
<b>Known projects using the ontology</b>	
<b>Other info</b>	

### 3.4 Multimedia Core Ontology

The ontology that follows has been created for a general use, in order to provide the grounds, concepts and relations, so that a domain ontology can efficiently be founded and extended. This is described in the "Ontology Description".

#### ■ CIDOC CRM Ontology.

*Ontology Description:* CIDOC CRM is proposed to be considered as part of a core ontology for Multimedia Objects supporting all multimedia objects, especially concerning cultural heritage items and events [6].

<b>Scope and usage (annotation, analysis, retrieval, reasoning, personalized filtering, meta-modeling, storage etc.)</b>	<input type="checkbox"/> Annotation <input type="checkbox"/> Retrieval
<b>Media description - Description of the Information Object Covered by the Ontology</b>	<input type="checkbox"/> 2DImages <input type="checkbox"/> 3D scanning & photogrammetry of objects <input type="checkbox"/> 3D objects <input type="checkbox"/> Recordings <input type="checkbox"/> Multispectral images <input type="checkbox"/> Physical objects (QT movie, VRML)
<b>Content Structure Covered by the Ontology</b>	<input type="checkbox"/> MPEG 7 <input type="checkbox"/> Spatial information <input type="checkbox"/> Temporal information
<b>Content Description Covered by the Ontology – Linking to specific Domains</b>	
<b>Other Concepts Covered by the Ontology</b>	<input type="checkbox"/> Creation data <input type="checkbox"/> Publication data <input type="checkbox"/> Use data
<b>Other (upper) ontologies used - modularity</b>	
<b>Language - Reasoning Support</b>	<input type="checkbox"/> OWL <input type="checkbox"/> RDFS <input type="checkbox"/> XML
<b>User preferences, specific</b>	<input type="checkbox"/> Performing arts recording

## Towards a Common Multimedia Ontology Framework

<b>constraints taken into account, etc.</b>	<ul style="list-style-type: none"><li><input type="checkbox"/> Ethnographic &amp; historical scenes</li><li><input type="checkbox"/> Aerial &amp; satellite images</li></ul>
<b>Known Tools using the ontology</b>	
<b>Known projects using the ontology</b>	
<b>Other info</b>	<ul style="list-style-type: none"><li><input type="checkbox"/> Dublin Core Compatible</li><li><input type="checkbox"/> Existing domain vocabularies</li><li><input type="checkbox"/> Linking domain vocabularies with cultural domain, e.g. Skos upper Ontology</li><li><input type="checkbox"/> Use of standard controlled lists, e.g. ULAN</li></ul>

### 4 Requirements

The complexity of content and structure that characterizes multimedia objects in combination with the need for high-level semantic analysis have turned the ontology driven representation of information related to and concerning multimedia content into a rather demanding process. The research activities, approaches and applications concerning multimedia ontology development, though, increase rapidly. The need for the establishment of a common multimedia ontology framework that would lead to their sharing and harmonization has become essential.

All research initiatives agree on the need for development of highly focused, easy to use, comprehensible, non-overlapping ontologies. They should be kept simple and small, addressing the substantial needs of the applications/systems that are going to use them. Modularity, clarity and unambiguousness should characterize them. More specifically, multimedia ontologies should address the requirements listed below:

#### 4.1 Scope and usage

The right design of generic multimedia ontologies is a rather challenging task, since it is concerned with the representation of common knowledge that can be accessed and utilized by arbitrary domain ontologies. The research activities regarding multimedia knowledge representation should focus on the right development of ontologies covering the whole range of scopes/usages listed below. Multimedia ontologies should be designed to serve one or more of these purposes, accurately and adequately.

- **Annotation** – e.g. summarization of multimedia content
- **Analysis** – e.g. ontology driven semantic analysis of multimedia content, etc
- **Retrieval** – e.g. context-based retrieval and recommendations
- **Reasoning** – e.g. application of reasoning techniques to multimedia content
- **Personalized filtering** – e.g. delivery of multimedia content according to user preferences
- **Meta-Modeling** – e.g. ontologies used to model multimedia processes, procedures, etc

#### 4.2 Media description - Description of the Information Object

The ontologies need to describe and represent knowledge for either one or even more of the following top-level hierarchical types of multimedia documents:

- Image
- Video
- 3D Graphics
- Audio
- AudioVisual
- Multimedia presentation

A further requirement is the clear distinction between annotations describing the information object and these concerning the multimedia document's content (e.g. semantic concepts depicted in an image) ([3], [9], [14]).

### 4.3 Representation of Content structure

- **Multimedia structure representation.** Multimedia ontologies should be able to represent the structure of a multimedia document itself, depending on the type of document ([1], [5], [7], [14]) and the relations between different structural elements. Structure in terms of segmentations - that is, still regions for images, shots for video clips, further decomposed into moving regions and so on - should be represented in order to provide domain knowledge for reasoning. A hierarchical structure of multimedia segments is thus needed in order to capture all possible types of media decompositions and relations. Additionally, representation may follow the structure of the multimedia artifact. For example, a set of annotations of a video fragment may be structured along the same scene/sequence/shot/ hierarchy as the video itself [14].

MPEG-7 Multimedia Description Schemes (MDS) supports hierarchical structures for multimedia segment decomposition. Thus, in the design of an associated ontology, a large part of MPEG-7 could be appropriately captured, aligned and used ([1], [7], [9], [13]).

- **Spatiotemporal Relation Representation.** The sole description of audiovisual characteristics seems to be an inadequate representation of a multimedia document. Spatial relations are needed to describe how segments are placed and relate to each other in 2D space (e.g. right and above). Topological spatial relations are needed to describe how the spatial boundaries of the segments relate (e.g. touches and overlaps). Temporal relationships among segments or events, providing information about the sequence in time, need to be represented, especially when the multimedia object is of type video. Besides, spatiotemporal relations, defined for a particular object, are closely related to the content structure of that object. Spatiotemporal and partonomic relations, thus, should be accommodated in a separate module of the framework ([1], [2], [7], [9], [12], [14]). Existing parts of ontologies describing spatiotemporal relations (e.g. as in the DOLCE ontology) can be enriched to accommodate topological and directional relations between regions of different types [7].

MPEG-7 Multimedia Description Schemes (MDS), which supports spatial, temporal and spatiotemporal multimedia segment relations, could thus be adopted and used ([1], [13]) without implying that an MPEG-7 MDS based ontology is the most favorable [9]. The forthcoming TemporalURI RFC standard has also been suggested [14].

- **Localization.** The regions that correspond to semantic objects need to be described in terms of their location within the multimedia content ([1], [3], [14]).

### 4.4 Representation of Content Description

- **Low-Level Description Representation.** The ontologies should model concepts and properties that describe audiovisual characteristics of objects, especially audio segments and visualizations in still images and videos, in terms of low-level features. In order to semantically express this kind of characteristics associated with a concept, several different audiovisual attributes (e.g. color, shape, texture, motion, localization etc) need to be represented, depending on the concept ([1], [2], [3], [5], [7], [8], [9], [12]).
- **Support of Multiple Audiovisual Descriptions.** Audiovisual characteristics of domain concepts cannot be described using one single instance of the audiovisual descriptors in question. For example, while the net of a tennis racket might be described in terms of its texture only once, its shape heavily depends on the viewing angle and occlusions (e.g. by the player in front of the net). The required conceptualisation thus has to provide means for multiple prototypical descriptions of a domain concept ([1], [7]).
- **Alignment to existing metadata standards.** For reasons of interoperability and sharing, the need for alignment to existing metadata standards grows higher.  
The MPEG-7 multimedia content description standard already provides tools for representing fragments such as colour, texture, shape and motion descriptors and since its ontological representation has already been implemented, an approach aligned to MPEG-7 Standard could be followed ([1], [7], [8], [9], [11], [14]). At this stage, linking to MPEG-7 Standard seems to be convenient for multimedia ontology development, but since the limitations of MPEG-7 have been identified [14] and need to be overcome, an MPEG-7-based ontology development need not be considered a dominant requirement ([9]). Ongoing efforts are oriented towards the extension of MPEG-7 by adding semantics, particularly for annotation, (intelligent) retrieval and, possibly, reasoning ([3], [11]).  
The use of Semantic Web standards ([1], [2], [14]), W3C standards ([1], [3]), VRA [14], TV Anytime [14] or other IPTC standards has been mentioned, while another issue is the establishment of appropriate structures that would make feasible the cross-reference between metadata standards and ontologies [8].
- **Support for basic and complex data types representation.** Ontologies aligned to standards built on XML Schema, such as MPEG-7 or others, would have to embody the representation of basic data types, such as numeric types, dates, vectors arrays, e.g., numerical vectors for the visual content ontology and so on ([1], [3], [7]). The use of feature matching algorithms on such data, as part of the reasoning process during knowledge-assisted analysis makes this need quite requisite [1].

### 4.5 Other Concepts Need to be Covered by the Ontology

- **Provenance details - Multi-layer annotations.** Details regarding the description of the multimedia object itself that is the creation day the creator, the purpose it was created for or even its subsequent history, etc. should be represented in an ontology [12], since provenance information provides rather important metadata of the multimedia document. More generally, a distinction needs to be drawn between the description of the content of a multimedia object and the description of different aspects of knowledge about that object [15]. Another important feature would be to record a trace of how a specific result has been obtained starting from a particular input [11]. Also, different analysis algorithms come up with different annotations of multimedia objects. Moreover, annotations produced by different annotators (human or not) might be inconsistent and mixing them might result in an invalid and useless annotation [9]. Multi-modality is important, especially for videos. A multimedia ontology framework should support annotations that distinguish between different modalities, annotations of different sources, annotations produced by different annotators (humans or not) [9], thus enabling their linking and further processing.
- **Semantics in the description of the delivery context and usage.** Due to the large number and wide variety of Web access devices, quite essential has become for a web service to know the required device capabilities that playback of a media object and other characteristics of the delivery context that influences their presentation [14]. Examples include network bandwidth, screen resolution, update frequency, colour depth etc. [14] It should be noted that applications need to be able to describe the delivery context of the client itself [14]. Context information is provided by existing standards such as Composite Capabilities/Preference Profiles (CC/PP). In addition, device profiles are defined in Open Mobile Alliance (OMA), and there are also W3C device and service profiles and 3GPP Generic User Profiles.

### 4.6 Knowledge representation formalism

There are many ontology languages (like OWL or WSMO family) available with different expressiveness and reasoning capabilities. The main criteria for the selection of an ontology language are its knowledge representation mechanism and the inferencing / reasoning support needed by an application [8].

The high complexity of multimedia modelling requires a representation language with high semantic expressiveness. This fact in combination with the compliance to W3C standards makes OWL the most appropriate language for multimedia knowledge representation. OWL is the best-suited ontology language, since it has large expressive power, while it preserves low average-case computational complexity.

In case that DL-reasoning should be supported, OWL DL or its future decidable extensions is the right choice for ontology development, due to its high inferencing capability [3]. Also, efficient reasoning engines (FaCT, RACER) and

APIs that support OWL-DL are available. Explicit fuzzy extensions of OWL DL could be derived [7], while such an approach could be extended for OWL-QL to support uncertain rules and queries, as well ([4], [7]).

### 4.7 Reasoning support

Despite many advances in description logic inference technology, for representation languages with high expressivity (e.g. OWL-DL), practical application scenarios with large ontologies require the development of more scalable reasoning techniques.

- **DL-based reasoning.** A lot of groups/projects aim at the proposal of developing multimedia decision-support systems.
- **Rules.** Rule extensions are often required for reasoning for example in medical imaging [3].
- **Uncertainty, fuzziness.** Uncertainty is involved in a multimedia object's recognition process, annotation, analysis and plays a great role in reasoning. For example, in an image, if an area is recognized as blue, it is placed above an area identified as brown and if they are adjacent, then the blue area may be annotated as sky with some degree of confidence, the brown area as sand with another degree of confidence and the multimedia object (image) may be classified to the domain summer vacation (again with a certain) level of confidence. This type of uncertain information needs to be handled and explicitly represented ([3], [14]).

One also should pay respect to the fact that most complex DL reasoners do not scale well which gets increasingly important if you want to handle huge amounts of data. There are mainly two primary reasoning tasks on the Semantic Web: subsumption and query answering. For the subsumption problem efficient description logics algorithms can be used but for query answering only few optimizations exist that do this task efficient. Efficient query answering could be done for example with OWL Light, which however has less expressive power than OWL DL [8].

Also, reasoning may be required at a meta modeling level, to decide on the analysis tools/algorithms that are more appropriate.

### 4.8 Following an established methodology

There is not a single correct methodology for designing and building an ontology. There are, however, a few methodologies that are well established and seem to perform better than others. Such a methodology could be adopted for the development of an ontology, while suggested are the OntoKnowledge methodology [2] and the OntoClean methodology [4].

### 4.9 Other requirements

- **Easy retrieval.** Annotations of multimedia content can become rather complex, especially if they are used for reasoning purposes. Metadata

## Towards a Common Multimedia Ontology Framework

need to be structured in order to allow easy retrieval ([3], [4], [9]). Efforts are being made to address sub-parts of the metadata directly, without the need to traverse the graph or use complex algorithms with a dynamic version of named-graphs [9]. This approach can be useful for rich multimedia annotations, especially if several dimensions of annotation have to be covered, such as low-level features, structure, reasoning-oriented annotations, retrieval-oriented annotations and administrative annotations, provenance tracking, etc [9].

- **Openness:** Ontologies need to be open in order to be widely-adopted by multiple and heterogeneous communities ([3], [14]).

## 5 Harmonization approaches

Development of ontologies and especially ontologies regarding multimedia content is quite an empirical process and due to that reason research communities involved with ontology-driven analysis, while sharing the same knowledge do not share a mutual consensus. Harmonization approaches, therefore, need to be followed among all these ontology-driven based systems and applications, in order to pave the way towards the development of an integrated knowledge infrastructure. The requirements for suitable, complete, easy to use and to share multimedia ontologies have been explicitly declared in the previous section. Basically, while research work advances and improves, new requirements arise, which need to be taken into account in the knowledge infrastructure. Nonetheless, based on the requirements stated by groups working on ontology development, the grounds for the establishment of a common multimedia ontology framework can be set, while the maintenance of such a framework could be driven by demand. Suggestions for such an attempt are listed below.

### 5.1 Modularization

While modularity is the requirement, modularisation is the partition of the knowledge infrastructure into accurate ontological modules, modelling different aspects of multimedia content. Multimedia ontologies, in an attempt to meet all the needs of a certain domain, end up being rather large, tangled, difficult to use and to maintain. A good suggestion would be to separate all these existing ontologies into a number of clear, independent ontology modules that would thus allow their re-use, revision and updating, without changes applied to one affecting the other ([1], [2], [4], [9], [11], [12], [14]). Such a task could serve as the first step to the establishment of a common ontology framework. A good example would be the separation into a Visual Descriptors Ontology, a Multimedia Structures Ontology, a Spatiotemporal Ontology ([1], [7], [9]). Maintaining such a division of functionalities would enable the independent use of each ontology, and furthermore, allow for their linking, combination and extension<sup>2</sup>.

### 5.2 Ontology cross-domain linking

Multimedia ontologies represent knowledge that needs to be exploited by a wide range of domains. Thus, after ensuring modularity, alignment/mapping techniques need to be developed, as well, for the linking of multimedia ontologies with domain-specific ontologies ([3], [7], [14]). Merging techniques need to be established in order to accomplish multimedia ontologies coordination ([3], [7]). Machine learning techniques could be used in order to detect similarities between concepts exploiting features in syntactic, structural and semantic level [7].

---

<sup>2</sup> A good approach for modularisation is described in A Rector. **Modularisation of Domain Ontologies Implemented in Description Logics and related formalisms including OWL.** in *Knowledge Capture 2003*, (Sanibel Island, FL, 2003), ACM, 121-128.

### **5.3 Harmonization of core ontologies**

Core ontologies are defined through abstract concepts, which are generic enough to be exploited by a wide range of domains, not necessarily from multimedia. They are developed in order to accomplish interoperability among metadata and determine structures and concepts upon which single domain ontologies could be implemented [10]. Multiple core ontologies have been introduced, such as SUMO (Suggested Upper Merged Ontology), DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering), CIDOC CRM [6], ABC, the Dublin Core Ontology, each of them recommended and used accordingly to the domain it provides general representation for. Furthermore, core ontologies are supposed to be defined in such a way that allows for extension from other ontologies. In fact, when an ontology is not sufficient to supply the requirements, the procedure should be to extend an existing ontology with other identifiers and definitions, and not to build a new one [11]. For these reasons, multiple core ontologies cannot be simply mapped. Rather, they must be "harmonized", so that the transition from one model to the other can be made without compromising the functionality [6]. A series of ontological modelling principles need to be applied, during harmonization, which is quite an iterative procedure, and the result would be real, precise, semantic interoperability [6].

### **5.4 Development of a common upper ontology**

Multiple core ontologies, as mentioned above, exist and are being developed in order to address the need of describing abstract and generic concepts of different domains. Their complete merge or even a successful combination of their sub-parts (classes and sub-classes) could lead to the establishment of one common multimedia core ontology. Moreover, an operational core concept reflecting the aspects of knowledge about multimedia that ontologies aim to systematize needs to be precisely defined. Around this fundamental concept, the common upper ontology shall be developed. The role of that core ontology would be to serve as a starting point for the construction of new ontologies, to provide a reference point for comparisons among different ontological approaches and to serve as a bridge between existing ontologies. It should also support the description of all existing modalities at different levels of granularity and it should be able to capture knowledge about the suitability of different modalities to represent particular types of information [15].

The fulfilment of this ambitious requirement will provide the basis for real semantic interoperability and lead to the representation of multimedia objects in a common, unified way. The need for such an upper ontology has been stressed in the majority, if not in all, of the research initiatives that responded to our call.

### **5.5 Establishment of a common upper model**

As already mentioned, the development of a common upper ontology is favourable and the establishment of a common concise and consistent multimedia ontology framework is the main goal. Taking into account the present situation, after modularisation is applied, a set of specific, modular and

unambiguous ontologies could evolve and form a first upper model. Soon ready for adoption would lead to a larger number of early adopters and could thus deliver useful results on the deployment of the framework [9]. Such a model would back up the common upper ontology and would act as the ground basis for the establishment of an integrated ontology framework ([9], [11]).

### 5.6 Compliance to standards

Due to reasons of interoperability, sharing and re-usability, as already mentioned in the previous section, the multimedia ontology framework should support Semantic Web standards and best practices ([2], [12], [14]), languages and technologies supported by W3C ([3], [12]), MPEG-7 ([1], [7], [8], [9], [11], [13], [14]), without focusing on just one specific standard [14].

## 6 Conclusions

This report provides a synopsis of the contributions submitted to the call for requirements for a common multimedia ontology framework. A first observation is that many interesting approaches exist; nevertheless they typically consider only a narrow range of multimedia usages and applications. Thus, the next step is to try to integrate different proposals with a view to developing a unified framework that will be capable of supporting multimedia annotation, analysis, retrieval, reasoning, personalisation and meta-modeling.

The requirements for multimedia ontologies cover a wide spectrum of dimensions. In brief, multimedia ontologies should be rich enough to describe the multimedia document and its content, the content structure and the spatio-temporal relationships between the entities depicted. They should also be capable of capturing the low-level descriptor information and to allow for basic and complex data types. Moreover, there is a clear request for provenance information and the conceptualisation of the inherent uncertainty to be part of the multimedia ontology. Alignment with standard is a safe approach in order of the ontology to be of practical interest.

Achieving ontology harmonization maybe difficult and complex in practice, but the proposals to this end are simple conceptually. The most significant factors are the efficient modularisation of ontologies, the easy linking to other ontologies, and the specification of minimum set of ontologies to be used for arbitrary applications. We believe that the specification of requirements can be of great help for the definition of such a minimum set, as it makes it more explicit which multimedia concepts are essential for any application.