Interoperability support between MPEG-7/21 and OWL in DS-MIRF

Chrisa Tsinaraki Panagiotis Polydoros Stavros Christodoulakis
chrisa@ced.tuc.gr panpolyd@ced.tuc.gr stavros@ced.tuc.gr
Technical University of Crete, Laboratory of Distributed Multimedia Information Systems and Applications (TUC/MUSIC), University Campus, 73100 Kounoupidiana, Crete, Greece

Abstract. We focus in this paper in interoperable semantic multimedia services that are offered in open environments such as the Internet. The use of well-accepted standards is of paramount importance for interoperability support in open environments. In addition, the semantic description of multimedia content utilizing domain ontologies is very useful for indexing, query specification, retrieval, filtering, user interfaces and knowledge extraction from audiovisual material. With the MPEG-7 and MPEG-21 standards dominating the multimedia content and service description domain and OWL dominating the ontology description languages, it is important to establish a framework that allows these standards to interoperate. We describe here the DS-MIRF Framework, a software engineering framework that facilitates the development of knowledge-based multimedia applications such as multimedia information retrieval, filtering, browsing, interaction, knowledge extraction, segmentation and content description. DS-MIRF supports interoperability of OWL with the complete MPEG-7 MDS as well as the complete MPEG-21 DIA Architecture so that domain and application ontologies expressed in OWL can be transparently integrated with MPEG-7/21 metadata. This allows applications that recog-nize and use the constructs provided by MPEG-7/21 to make use of domain and application ontologies, resulting in more effective retrieval and user interaction with the audiovisual material. We also present a retrieval evaluation methodology and comparative retrieval results.

1. Introduction

The advent of the Internet demonstrated the paramount importance of standards for the industry. Even though in closed environments bound by the organizational boundaries the needs of the organizations may be well covered by ad-hoc developed software tailored to their needs, in open environments, where contact with remote companies or users via Internet is of great importance, interoperability through the use of industry standards has become crucial.

The MPEG standards have led the industrial efforts in the multimedia domain. MPEG-7 [6] is today a widely accepted standard for describing aspects of the multimedia content related to retrieval and filtering, like content structuring metadata, user filtering metadata, usage metadata, segmentation metadata etc. MPEG-21 defines an “open framework for multimedia delivery and consumption” [21] and standardizes content adaptation and management, privacy and rights protection. Future standardization work in the multimedia domain will have to be based on the existing MPEG-7 and MPEG-21 standards, extending them according to needs.

Audiovisual content retrieval and filtering is a very important but difficult subject of research for the academia and the industry, and has received a lot of attention in scientific publications [8] [10]. It has been shown in many real-world applications that the retrieval effectiveness (as measured, for example, by the precision-recall curves) can be greatly improved when domain and application knowledge encoded in ontologies is used for indexing and retrieval purposes.

The work described in this paper is motivated by the need to provide interoperable mechanisms for extending the MPEG-7/21 metadata with domain and application knowledge.

MPEG-7 models events, but it offers abstract semantics that can be interpreted in a different way in different domains. We had to associate valid events, agents and experiences to achieve better precision. In the framework described here we extend the MPEG-7 semantic content description metadata with domain knowledge in a way that allows the existing MPEG-7 based applications to remain operational while the same or other applications can take advantage of the additional knowledge. Consider as an example a user who wants to retrieve the audiovisual segments where the soccer player Zidane scores against the goalkeeper Buffon. Using only the MPEG-7 constructs, the exact query *"Give me the segments where Zidane scores against Buffon"* cannot be expressed. Instead, the less accurate query *"Give me the segments where an event takes place, having Zidane as agent and Buffon as experiencer"* can be posed. This query will result in false drops, since the query result will include every segment which contains an event with Zidane and Buffon being the agent and the experiencer respectively of the event (e.g. even when Zidane serves a drink to Buffon in a social event). Using domain knowledge captured in a soccer ontology together with the MPEG-7 constructs allows a user to express the query *"Give me the segments where a goal event takes place, where the player Zidane scores against the goalkeeper Buffon"*. The later query has the same meaning with the original user query, thus resulting in enhanced retrieval effectiveness.

Consider now a user interested in receiving only *"the segments containing the goals scored by France"*. Even if domain knowledge is integrated in MPEG-7 semantic content descriptions, it cannot be expressed in this MPEG-7/21 user preference description, because MPEG-7/21 allow keyword-only descriptions of the user preferences. However, if the user relies in the keywords "goal" and "France" it may be the case that a goal is scored against France. The example demonstrates that structured MPEG-7 metadata descriptions and domain knowledge integrated with them cannot be utilized in the MPEG-7/21 user preference descriptions.

An ontology is a logical theory accounting for the intended meaning of a formal vocabulary, i.e. its ontological commitment to a particular conceptualization of the world [22]. Ontologies are often expressed in ontology definition languages based on Description Logics (DL), which allow rich ontology structures. *OWL (Web Ontology Language)* [3] is the dominant standardization effort in ontology description and it is expected that both, many OWL domain ontologies will exist in the future, as well as that many scientists will be familiar with OWL and will be using it for the definition of new ontologies. It is therefore very important for the audiovisual industry to have a methodology for the interoperability of OWL with MPEG-7/21 and for the integration of domain knowledge expressed in OWL within MPEG-7/21. This way, MPEG-7/21 will become first class Semantic Web objects and the methodologies and tools developed in the Semantic Web (such as reasoners) may be used with them.

In this paper we present an approach for interoperability support between OWL and MPEG-7/21 in the context of the *DS-MIRF (Domain-Specific Multimedia Indexing, Retrieval and Filtering) framework*, a software engineering framework that aims to facilitate the development of knowledge-based multimedia applications utilizing and extending the MPEG-7/21 standards. We also describe how application-specific extensions to MPEG-7/21 may be expressed in OWL as application ontologies and how they may guide the production of both standard MPEG-7/21 metadata and metadata complying with models extending MPEG-7/21 appropriately for advanced application support. Then we demonstrate how domain ontologies described in OWL can be integrated transparently in MPEG-7/21. Finally, we show how retrieval and filtering applications are supported, according to the proposed approach, in the context of the DS-MIRF framework. We also present a retrieval evaluation methodology and comparative retrieval results which are encouraging.

This work extends previous work described in [9] [10] to cover all the aspects of the *MPEG-7 MDS (Multimedia Description Schemes)* [5] and the *MPEG-21 DIA (Digital Item Adaptation) Architecture* [19]. Little has been published in the past in this area of research, although the importance of domain ontologies in audiovisual content recognition, indexing and retrieval is widely recognized [1] [7].

The previous research efforts closest to ours are presented in [4], [11] and [8]. In [4] the DAML+OIL [12] ontology definition language is used to partially describe the MPEG-7 metadata structures, but not the complete MPEG-7 MDS. The authors of [4] do not propose a specific methodology and/or software for the integration of domain ontologies in MPEG-7. In [11], the ABC Upper Ontology [13] is used together with an OWL ontology partially capturing MPEG-7 (mainly the low-level visual features) and an OWL ontology for fuels. The main difference of this approach with our work is that it focuses on the low-level visual features and not on the higher-level MPEG-7 semantic concepts. Thus, the domain and application knowledge is not integrated with the MPEG-7 semantics; it is rather extending the concepts of the ABC Upper Ontology. In [8], a two-layered ontology-based approach is taken for the semantic description of audiovisual segments. The first layer consists of an ontology that captures the concepts of genre, theme and technical process and the second layer provides domain knowledge, expressed as a set of domain ontologies, that extends the semantics of the first layer. The domain knowledge in this work extends an ontology capturing a limited set of concepts and therefore it does not allow the full exploitation of the constructs provided in MPEG-7/21 for semantic content description.

The lack of semantic support in the user preference descriptions of MPEG-7/21 led to the development of MPEG-7/21 based systems that either utilize keyword-only metadata in the user preference descriptions and ignore the rich content descriptions that utilize the structured

MPEG-7 semantic metadata [16] [17] [18], or developments that ignore the MPEG-7 user preference model and follow proprietary filtering approaches on top of the structured MPEG-7 semantic metadata [15]. In order to overcome this serious limitation we have developed a semantic user preference model that has as special case the existing MPEG-7/21 user preference model but it allows in its more general forms to fully exploit the semantic structures of MPEG-7 and the encoded domain knowledge in the content metadata descriptions. We have integrated an implementation of the model as an OWL application-specific implementation of the DS-MIRF framework. This way we have a complete semantic multimedia framework that fully supports semantic content description utilizing interoperability mechanisms for MPEG-7/21 and OWL, as well as user preference descriptions which allow to fully exploit the semantic MPEG-7 content descriptions, while they have as a special case the existing MPEG-7/21 user preference descriptions.

The rest of the paper is organized as follows: Brief descriptions of the MPEG-7 MDS, the MPEG-21 DIA Architecture and OWL are provided in section 2. An overview of the DS-MIRF framework is presented in section 3, while the ontological infrastructure provided in the DS-MIRF framework is described in section 4. The support provided in the DS-MIRF framework for multimedia content retrieval and filtering services and the retrieval evaluation results are presented in section 5. Conclusions and future work are discussed in section 6.

2. Multimedia and Ontology Description Standards

In this section we present an overview of the multimedia and ontology description standards on which the DS-MIRF framework relies: The MPEG-7 MDS in subsection 2.1, the MPEG-21 DIA Architecture in subsection 2.2 and the OWL ontology definition language in subsection 2.3.

2.1. The MPEG-7 MDS

The *MPEG-7 MDS (Multimedia Description Schemes)* provides constructs for the definition of metadata for multimedia content and service description. The MPEG-7 MDS has been defined using the MPEG-7 DDL, which is essentially the XML Schema Language [2], extended with the basic datatypes necessary for the definition of the *Description Schemes (DSs)* of the MPEG-7 MDS. The DSs are essentially complex datatypes used for the description of concepts in their scope. The MPEG-7 MDS is comprised of the following major components [5]: (a) The *Basic Elements*, which include schema tools, basic datatypes, mathematical structures, linking and media localization tools as well as basic DSs, which are used as elementary components of more complex DSs; (b) The *Content Description and Management Elements*, which provide the *Creation and Production*, *Media* and *Usage* tools in order to capture content management information and the *Structural Aspects* and *Semantic Aspects* tools in order to capture content description information; (c) The *Navigation and Access Elements*, where browsing is supported through multimedia content summary descriptions and multimedia content variations; (d) The *User Interaction Elements*, which are used to describe user preferences regarding multimedia content, as well as audiovisual material consumption aspects; and (e) The *Content Organization Elements*, where the organization of the multimedia content is addressed by classification, modelling and multimedia document collection specification.

In addition to the DSs, the MPEG-7 MDS provides the *Classification Schemes (CSs)*, which essentially are thesauri comprised of term hierarchies. The 40 standardized CSs provided in the MPEG-7 MDS contain the allowed values of specific attributes of the MPEG-7 MDS DSs. The CSs are XML documents compliant to the XML schema of the MPEG-7 MDS.

2.2. The MPEG-21 DIA Architecture

The major objective of the *MPEG-21 DIA (Digital Item Adaptation) Architecture* [19] is to support, in the current usage environment, the adaptation of the MPEG-21 Digital Items (DIs), which are the fundamental units of distribution and transaction inside the MPEG-21 framework. The MPEG-21 DIA tools are based on the *Schema Tools*, which provide uniform root elements for all the DIA descriptions, and the *Low-Level Data Types*, which provide some low-level and basic datatypes which can be used by several DIA tools independently. The DIA tools are classified into eight categories: The *Usage Environment Description Tools*, the *BSDLink*, the *Bitstream Syntax Description (BSD)*, the *Terminal and Network*

Quality of Service (QoS) Tools, the *Universal Constraints Description Tools*, the *Metadata Adaptability Tools*, the *Session Mobility Tools* and the *DIA Configuration Tools*.

The most important for us are the *Usage Environment Description Tools*, where the *Terminal Capabilities*, the *Natural Environment Characteristics*, the *Network Characteristics* and the *User Characteristics* are specified. In the user characteristics, the user features are captured, including information about the user, user preference descriptions, usage history, presentation and media conversion preferences, accessibility characteristics and information that allows to better support context-based services (like mobility characteristics, destination and focus of attention). The user preferences and the usage history descriptions are formed according to the homonym MPEG-7 MDS types. A user preference description comprises of: (a) a set of *FilteringAndSearchPreferences (FASP)* elements, which describe the user preferences regarding multimedia content filtering and searching; and (b) a set of *BrowsingPreferences* elements, which describe the user preferences regarding multimedia content summaries.

2.3. The Web Ontology Language (OWL)

The *Web Ontology Language (OWL)* [3] is the dominant standard in ontology definition. OWL has been developed according to the description logics paradigm and uses RDF(S) [23], [24] syntax. Three OWL species of increasing descriptive power have been specified: OWL Lite, OWL DL and OWL Full. The basic functionality provided by OWL is following:

- Import of XML Schema Datatypes, through the `rdfs:Datatype` construct, for the representation of simple types that extend or restrict the basic datatypes (e.g. ranges etc.).
- Definition of OWL Classes, using the `owl:class` construct, for the representation of sets of individuals sharing some properties. Class hierarchies may be defined using the `rdfs:subClassOf` construct.
- Definition of OWL properties, for the representation of the features of the OWL class individuals. Two kinds of properties are provided by OWL: (a) *Object Properties*, which relate individuals of one OWL class (the property domain) with individuals of another OWL class (the property range). Object properties are defined using the

`owl:objectProperty` construct; and (b) *Datatype Properties*, which relate individuals belonging to one OWL class (the domain of the property) with values of a given datatype (the range of the property). Datatype properties are defined using the `owl:datatypeProperty` construct.

Property hierarchies may be defined using the `rdfs:subPropertyOf` construct.

- Definition of class individuals and their relationships.
- Definition of restrictions, using the `owl:Restriction` construct, including type restrictions, cardinality restrictions and value restrictions.

3. Overview of the DS-MIRF Framework

In this section we present a brief overview of the DS-MIRF framework. In software engineering a framework is an extendible subsystem that supports a set of related services. The DS-MIRF Framework aims to facilitate the development of knowledge-based multimedia applications utilizing and extending the MPEG-7/21 standards.

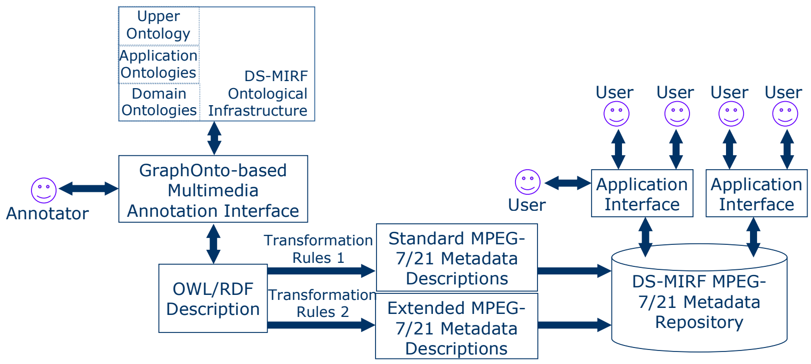

Figure 1: Information Flow in the DS-MIRF Framework

The architecture of the DS-MIRF framework as well as the information flow between its components and the interaction with the end-users are illustrated in Figure 1. In this figure, the multimedia content *annotator* is a special type of user that is responsible for the semantic annotation of multimedia documents. He uses an *annotation interface* that is based on the GraphOnto component [25] that allows ontology-based semantic annotation and utilizes the ontological infrastructure of the DS-MIRF framework. Since all the ontologies in the DS-MIRF framework are expressed in OWL, the result of the annotation process is an OWL description of the multimedia content. The OWL descriptions are then transformed, using the appropriate set of *transformation rules* (implemented in the GraphOnto component), either to *standard MPEG-7/21 metadata descriptions* or to *extended MPEG-7/21 metadata descriptions*, which capture application-specific needs. The extensions are usually specified in other standards or sound models that are expressed in (some of) the application ontologies. The MPEG-7/21 metadata – standard or extended – are stored in the *DS-MIRF MPEG-7/21 Metadata Repository*, which is accessed by the end-users through appropriate *application interfaces*. The application interfaces may provide the end-users with multimedia content services like multimedia content retrieval, filtering and delivery. An OWL reasoner, Pellet[1], has been integrated in the software of the DS-MIRF framework. The reasoner will be useful in activities like complex constraint checking, logic-based knowledge acquisition from multimedia data, and complex query processing. Currently, the validation of both ontologies and ontology-based metadata takes place during the annotation process.

The DS-MIRF ontological infrastructure is shown in Figure 2 and includes:

- An *OWL Upper Ontology* that fully captures the MPEG-7 MDS and the MPEG-21 DIA Architecture, which is the cornerstone of the ontological infrastructure of the DS-MIRF framework and the basis for interoperability between OWL and MPEG-7/21.
- A set of *OWL Application Ontologies* that provide additional functionality in OWL that either makes easier for the user the use of the MPEG-7/21 (usually constructs implied in the MPEG-7/21 text like the typed relationships) or supports advanced multimedia content services (like, for example, semantic user preferences). The Application Ontologies provide general-purpose constructs that are either implied in the text of MPEG-7/21 (but missing in the syntax) or not available in MPEG-7/21.

  We have developed an application ontology that contains a set of extensions for the MPEG-7 MDS, which allows the full representation of typed relationships that are literally described in the MPEG-7 MDS text but their features are not fully captured in the MPEG-

[1] The Pellet OWL Reasoner is available at: http://www.mindswap.org/2003/pellet/index.shtml.

7 MDS syntax. We have also developed an application ontology for the description of semantic user preferences for multimedia content, as the MPEG-7/21 user preference descriptions allow keyword-only descriptions of the semantics of the preferred content. The application ontology is based on a semantic user preference model we have proposed that also allows for the explicit specification of the boolean operators to be used in the different phases of multimedia content search and filtering.

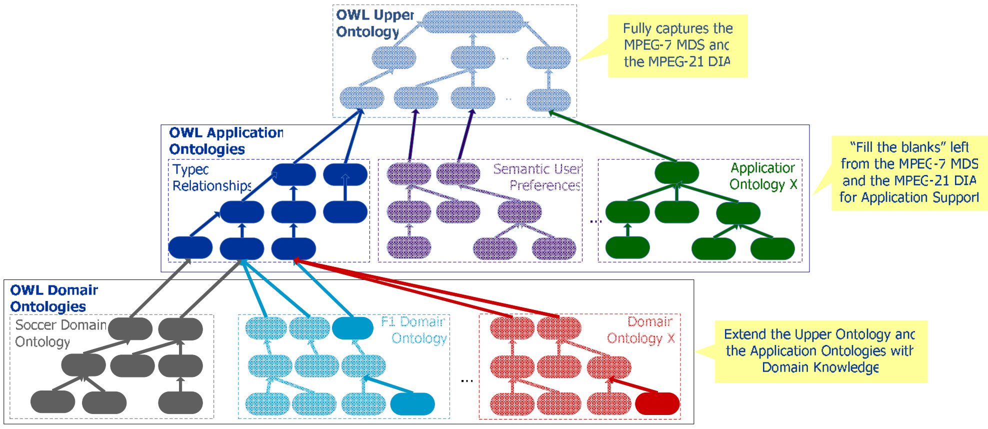

Figure 2: The ontological infrastructure of the DS-MIRF framework

- The *Domain Ontologies*, which extend the Upper Ontology and the Application Ontologies with domain knowledge. For example, consider sports ontologies that extend the abstract semantic description capabilities of the MPEG-7 MDS. We have developed a methodology for defining and integrating domain ontologies in the DS-MIRF framework and we have defined ontologies for soccer and formula 1 in order to test our methodology.

Our approach for interoperability support between MPEG-7/21 and OWL is useful for two distinct target groups: (a) Users that are aware of the DS-MIRF framework and the software offered, and would like to integrate in the framework ontologies they are developing, so that they are interoperable with MPEG-7/21; and (b) Users having ontologies independent of the DS-MIRF framework, which may be integrated in the DS-MIRF framework as domain or application ontologies.

4. The DS-MIRF Framework Ontological Infrastructure

We present in this section the ontological infrastructure provided in the DS-MIRF framework, which is comprised of: (a) The Upper, Application and Domain ontologies that we have defined in order to allow the integration of domain and application knowledge in MPEG-7/21 compliant metadata; and (b) The ontology definition and integration methodologies that we have developed. We present the Upper Ontology in subsection 4.1, and we discuss application ontology related issues in subsection 4.2. In subsection 4.3 we describe the methodology we developed for the integration of OWL Domain Ontologies in DS-MIRF.

4.1. The Upper Ontology of the DS-MIRF Framework

Our approach for interoperability support in the multimedia domain utilizes an Upper Ontology that fully captures the metadata model provided by the MPEG-7 MDS and the MPEG-21 DIA Architecture. In addition, all the constructs of the MPEG-7 Visual [20] that are necessary for the representation of the MPEG-7 MDS are also captured in the Upper Ontology. The Upper Ontology has been implemented in OWL, according to a methodology that essentially defines the steps needed to transform the tree-structured XML Schemas of the MPEG-7 MDS and the MPEG-21 DIA Architecture in the Description Logics based OWL syntax. The Upper Ontology was therefore defined according to the following methodological steps:

1. *MPEG-7/21 Simple Datatype Representation:* The MPEG-7/21 simple datatypes are imported from the XML Schema syntax using the `rdfs:Datatype` construct.
2. *MPEG-7/21 Complex Type Representation:* Every MPEG-7/21 complex type is represented as an OWL class, which has as `rdf:ID` the value of the complex type name.
    2.1. *Representation of Attributes and Simple Type Elements:* The attributes and the simple type elements (of type string, integer etc.) of the complex type are represented as OWL datatype properties, having the newly-defined class as domain and the simple type as range.
    2.2. *Representation of Complex Type Elements:* Complex type elements are represented as OWL object properties. An OWL class for the representation of the type of the complex type element is defined, if it does not already exist.
    2.3. *Subclassing:* For the representation of the subclass/superclass relationships that hold for the complex type, the following actions are performed:

2.3.1. If the complex type is a subtype of another complex type, the subclass relationship is represented using the `rdfs:subClassOf` construct.

2.3.2. If the complex type is a subtype of a simple type, a datatype property is defined with `rdf:ID` "*type_name*Content". Here *type_name* is the type of the supertype (e.g. string, integer etc.) that has the supertype as range and the newly-defined OWL class as domain.

As an example, we present in Figure 4 the definition of the OWL class "AgentType" (subclass of the "DSType" class that represents all the Descriptor Schemes). The "AgentType" class corresponds to the MDS complex type "AgentType" shown in Figure 3. The complex type element "Icon" is also shown in Figure 3 and the corresponding "Icon" object property is shown in Figure 4.

```
<complexType name="AgentType" abstract="true">
    <complexContent>
      <extension base="mpeg7:DSType">
        <sequence>
          <element name="Icon" type="mpeg7:MediaLocatorType" minOccurs="0" 
maxOccurs="unbounded"/>
        </sequence>
      </extension>
    </complexContent>
</complexType>
```

Figure 3: The MPEG-7 MDS type "AgentType"

```
<owl:Class rdf:ID="AgentType">
  <rdfs:subClassOf rdf:resource="#DSType"/>
</owl:Class>
<owl:ObjectProperty rdf:ID="Icon">
  <rdfs:domain rdf:resource="#AgentType"/>
  <rdfs:range rdf:resource="#MediaLocatorType"/>
</owl:ObjectProperty>
```

Figure 4: The "AgentType" OWL class of the Upper Ontology

3. *MPEG-7 Classification Scheme Representation:* MPEG-7 Classification Schemes are represented as instances of the MPEG-7 MDS type "ClassificationSchemeType", comprised of a set of, possibly nested, "Term" elements of type "TermDefinitionType". Thus we represent them as "ClassificationSchemeType" individuals related with a set of "TermDefinitionType" individuals through the "Term" Object Property. As an example, (a part of) the "RoleCS" CS, shown in Figure 5, is represented as shown in Figure 6.

```
<ClassificationScheme uri="urn:mpeg7:mpeg7:cs:RoleCS:2001" do-
main="//CreationInformation/Creation/Creator/Role"
"//UsageInformation/Dissemination/Disseminator/Role">
  <Term termID="AUTHOR">
    <Name xml:lang="en">Author</Name>
  </Term>
</ClassificationScheme>
```

Figure 5: A part of the "RoleCS" Classification Scheme of the MPEG-7 MDS

```
<ClassificationSchemeType rdf:ID="RoleCS">
  <uri>urn:mpeg7:mpeg7:cs:RoleCS:2001</uri>
  <domain>"//CreationInformation/Creation/Creator/Role"</domain>
  <domain>"//UsageInformation/Dissemination/Disseminator/Role"</domain>
  <Term>
    <TermDefinitionType rdf:ID="AUTHOR">
      <termID rdf:datatype="&xsd;NMTOKEN">AUTHOR</termID>
      <Name>
        <TermNameType rdf:ID="AUTHORNT">
          <lang>en</lang>
          <content>Author</content>
        </TermNameType>
      </Name>
    </TermDefinitionType>
  </Term>
</ClassificationSchemeType>
```

Figure 6: The "RoleCS" OWL individual that represents the part of the "RoleCS" Classification Scheme of the MPEG-7 MDS shown in Figure 5

The complete Upper Ontology (including the MPEG-21 DIA and the MPEG-7 MDS DS and CS representations) has been defined using the above rules but is not shown here due to space limitations. It is an OWL-DL ontology, available at http://astral.ced.tuc.gr/delos/content/testbeds/MPEG7MDS_MPEG21DIA.zip, and has been validated by the WonderWeb OWL species ontology validator[2].

4.2. Application Ontologies

The *Application Ontologies* in the DS-MIRF framework provide general-purpose constructs that are either implied in the MPEG-7/21 texts (but missing in their syntax) or are necessary for advanced multimedia service provision but are not available in MPEG-7/21, thus making multimedia content service support tedious and sometimes limited. We have implemented two application ontologies. The first application ontology is for the representation of typed relationships (see paragraph 4.2.1). It is an example of application ontology that provides

[2] The OWL species validator, available at http://phoebus.cs.man.ac.uk:9999/OWL/Validator, validates OWL ontologies and checks if they conform to one of the OWL species.

general-purpose constructs that are implied in the MPEG-7/21 text but missing from the MPEG-7/21 syntax. The second application ontology is for the representation of semantic user preferences (see paragraph 4.2.2). It is an example of application ontology that provides general-purpose constructs, which are not provided in the MPEG-7/21.

4.2.1. Application Ontology for Typed Relationship Representation

We present here the typed relationship application ontology we have developed which contains a set of extensions for the MPEG-7 MDS that allow the full representation of typed relationships that are described in the MPEG-7 MDS text but their features are not fully captured in the MPEG-7 MDS syntax. The users are not forced to use this ontology, but if they do so, the definition of relationships in MPEG-7 metadata descriptions becomes much easier.

The relationships are represented in the MPEG-7 MDS as instances of the "RelationType" class. Every relationship has its type, its source and target items, its strength and a characterization as directed or non-directed. Although this is not explicitly stated in the MPEG-7 MDS syntax, it is stated in the MPEG-7 MDS textual description that the relationship types should take the values defined in the *RelationBase CS* (for the representation of basic relationship types like equals, inside, refines etc.), the *TemporalRelation CS* (for the representation of temporal relationship types like precedes, follows, overlaps, contains etc.), the *SpatialRelation CS* (for the representation of spatial relationship types like over, below, north etc.), the *GraphRelation CS* (for the representation of relationships existing among graph nodes like identity, equivalent etc.) and the *SemanticRelation CS* (for the representation of semantic relationship types like shows, references, agent, patient, causer etc.). In addition, it is specified in the MPEG-7 MDS textual description for every relationship type, if it is directed and, if so, which is its inverse relationship. Consequently, the annotator who wishes to correctly define a relationship should have all this textual information at hand.

We have defined an OWL class hierarchy rooted in the "TypedRelationType" (which is a subclass of the "RelationType" class of the Upper Ontology), and we captured in the OWL classes all the information existing in the MPEG-7 MDS text. This forms an application ontology that can greatly facilitate application development by the users in the large majority of the cases. As shown in Figure 7, the direct subclasses of “RelationType” are homonyms of the classification schemes where the relationship types are defined. Each of them has a number of subclasses, which correspond to the relationship types defined in the homonym classification scheme, together with the information literally described about them in the MPEG-7 MDS text. The annotator that uses the typed relationship application ontology does not have to be aware of the textual description of the MPEG-7 MDS, since all the information is captured in the ontology. The typed relationship ontology is an OWL-DL ontology, it has been validated by the OWL species validator, and is available at: http://elikonas.ced.tuc.gr/ontologies/AppOntos/TRAO/TypedRelationships.

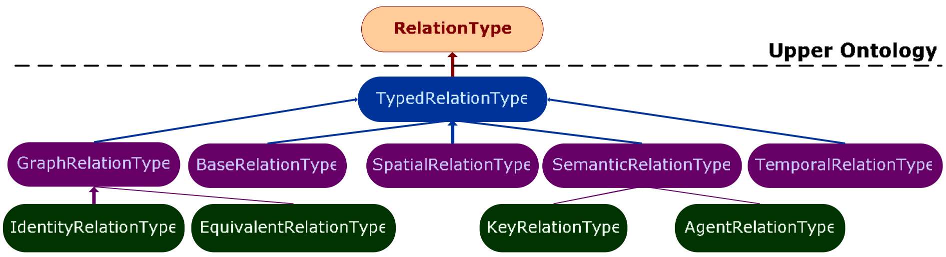

Figure 7: The OWL class hierarchy defined for the representation of typed relationships

As a typed relationship example, consider the OWL definition of the “AgentRelationType” class, which represents the agent relationship, and is shown in Figure 8.

```
<owl:Class rdf:ID="AgentRelationType">
  <rdfs:label>Relation</rdfs:label>
  <rdfs:subClassOf rdf:resource="#SemanticRelationType"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#type"/>
      <owl:hasValue>&SemanticRelationCS;agent</owl:hasValue>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#directed"/>
      <owl:hasValue>true</owl:hasValue>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#inverse"/>
      <owl:hasValue>#AgentOfRelationType</owl:hasValue>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>
```

Figure 8: OWL Representation of the "AgentRelationType" class

4.2.2. Application Ontology for Semantic User Preference Representation

In this subsection we present an application ontology for the description of semantic user preferences regarding multimedia content search, filtering and consumption. Our work was motivated from the observation that, although the MPEG-7 MDS constructs have been shown to allow encoding multimedia content semantics and domain knowledge, this knowledge cannot be integrated in a systematic way in the MPEG-7/21 User Preferences. Therefore, the users cannot express precisely their preferences about the multimedia content semantics and the structured semantic descriptions of the multimedia content cannot be fully exploited by the MPEG-7/21 User Preferences, where the end-users may specify the desired semantics of the content that is either delivered to them, as stated in their *FilteringAndSearchPreferences (FASP)*, or comprises the summaries automatically created for them, as stated in their *BrowsingPreferences*. For example, in the MPEG-7/21 user preference model, a user may state that the goals of a soccer game should be delivered to him (or be used to automatically create a summary for him), but he cannot state that only the goals scored by France should be delivered to him (or create a summary for him). The advanced preference descriptions are useful for demanding users particularly interested in a domain like soccer as well as people working in the domain, for example for a soccer team coach that would like to see three days before the next game all the goals scored by the next opponent of his team in the last two games.

Another limitation of the MPEG-7/21 User Preference model is that in the hierarchical filtering and search preferences (FASP) the users cannot explicitly state which boolean operator should be applied on which of the criteria stated in the different levels of the hierarchy; it is rather left to the retrieval system to decide how to combine them. Thus, a user may state that he is interested in penalties with weight 0.3, in goals with weight 0.7 and in France with weight 0.9, but cannot state that he is interested in (goals AND France) with weight 0.9 OR in (penalties AND France) with weight 0.5.

We have developed a model that allows the description of semantic user preferences, which can exploit all the power of the MPEG-7 semantic content descriptions. The model allows the users to explicitly specify the boolean operators to be used in the FASP hierarchies during content filtering, in order to overcome the serious limitations of the MPEG-7/21 User Preferences. The proposed user preference model is compatible with the MPEG-7 Semantic DS, and can exploit its constructs and the domain knowledge encoded with them. The user preference model has been expressed both as an OWL application ontology integrated in the DS-MIRF framework and as an XML Schema extension of MPEG-7/21. In the remaining of this section, we describe the semantic user preference model, its MPEG-7/21 compliant implementation as well as how we derive the OWL application ontology from it.

Semantic User Preference Model. The DS-MIRF semantic user preference model extends the MPEG-7/21 user preferences with semantic entities and boolean operators, thus allowing the precise specification of the desired audiovisual content semantics. Like the MPEG-7/21 user preference model, the DS-MIRF semantic user preference model allows the users to specify their *Browsing Preferences (BP)* and their *Filtering and Search Preferences (FASP)*. The browsing preferences allow the specification of the *Summary Preferences (SuP)* of the users, which are used to specify the desired summary properties. They also allow the specification of the *Preference Conditions (PC)*, which describe the usage conditions for each summary preference description. The DS-MIRF semantic user preference model extends the MPEG-7/21 summary preference descriptions with a set of *semantic summary descriptions (SSs)* that specify the preferred audiovisual content. The audiovisual content semantics may be described, in addition to the MPEG-7/21 weighed textual theme list, with a set of weighed semantic entity collections. The summary preference descriptions follow the regular expression syntax shown in Expression 1. They specify the *preferred, minimum* and *maximum summary duration* (*SD, MaxSD* and *MinSD* respectively), the *preferred, minimum* and *maximum number of keyframes* (*KF, MaxKF* and *MinKF* respectively), the *preferred, minimum* and *maximum number of characters for textual summaries* (*C, MaxC* and *MinC* respectively), the *preferred summary type (SType)*, and the *preferred summary theme (STheme)*. The *preference value (pv)* attribute denotes the importance of an element and is an integer value set (implicitly or explicitly) by the user, in the range [-100, 100]. The default value of pv is 10, while negative pv values denote negation. The formal syntax of a semantic summary (SS) description, which is a weighed collection of semantic entity descriptions (T) that describe the desired audiovisual content, is shown in Expression 2.

SP = (SType pv)*(STheme pv)*(**SS pv**)*[SD][MaxSD] [MinSD][NoKF][MaxNoKF][MinNoKF][NoC] [MaxNoC][MinNoC]

Expression 1: Summary Preferences (SP) formal syntax

SS = (T* pv)*

Expression 2: Formal syntax of semantic summary descriptions

A semantic entity description may contain: (a) The *id* of the semantic entity *(Tid)*, which may plays the role of a variable name; (b) The semantic entity *type (TType)*; (c) Desired attribute value specifications, represented by *attribute name (AName)* – desired *attribute value (AValue)* pairs; (d) Descriptions of the desired semantic entity relationships consisting of the necessary relationship *type (RType)* and *target (RTarget)*, and the optional relationship *source (RSource)* and *strength (RStrength)*. The relationships may be of the relationship types that are specified in the standard classification schemes of the MPEG-7 MDS; and (e) Descriptions of the desired values of the semantic entity elements, including for the *element name (EName)*, the list of the desired element attribute values represented by *attribute name (EAName)* – desired *attribute value (EAValue)* pairs and the list of the desired values of its sub-elements *(E)*. When several criteria are specified for the same semantic entity inside a semantic summary preference description, they are implicitly assumed to be logically ANDed. The formal syntax of a semantic entity is shown in Expression 3.

T = (Tid TType) | (Tid TType) AND ((EName (EAName EAValue)* (E)*)|(RType RTarget [RSource] [RStrength])|(AName AValue)) (AND((EName (EAName EAValue)* (E)*)| (RType RTarget [RSource] [RStrength])|(AName AValue)))*

Expression 3: Formal syntax of a semantic entity (T)

As an example of semantic browsing preferences, consider the preferences of a user who wants a summary containing the goals scored by France. They are expressed in the formal syntax shown in Expression 4. We assume in this example that France's goals are bound to the "FGoal" variable and that an abstract semantic entity exists, which has as id "Goal", which represents the class of all the goals. We also assume the existence of a semantic entity having as id "France", which represents the national soccer team of France. The "exemplifies" relation states that "FGoal" is an example of the abstract "Goal" event, the "agent" relation states that the "FGoal" has been scored by "France".

UP1 = ((FGoal, EventType) AND ((exemplifies, Goal) AND (agent, France)) 100)

Expression 4: Formal syntax of the preferences of a user who wants a summary containing the goals scored by France.

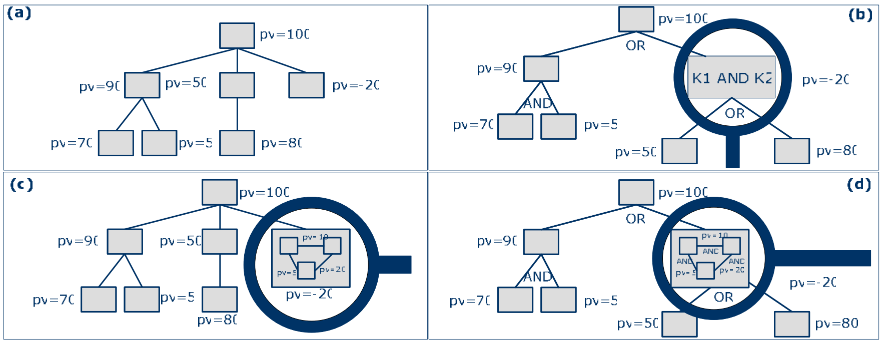

Figure 9: FASP structures supported by the semantic user preference model: (a) Weighed Keyword Hierarchies; (b) Weighed Keyword Hierarchies with Boolean Operators; (c) Weighed Hierarchies of Keywords and Semantic Entity Collections; and (d) Weighed Hierarchies of Keywords and Semantic Entity Collections with Boolean Operators.

The MPEG-7/21 user preference model allows the representation of the filtering and search preferences (FASP) as *Keyword and Non-Semantic Preference Weighed Hierarchies* (see Figure 9a). The DS-MIRF semantic user preference model extends the semantic description capabilities of the MPEG-7/21 FASPs: in addition to the MPEG-7/21 FASPs, it allows boolean operator specification, thus supporting the weighed keyword hierarchies with boolean operators shown in Figure 9b, as well as semantic entity hierarchies with and without boolean operators (see Figure 9c and Figure 9d respectively). It is clear that the MPEG-7/21 FASP structure of Figure 9a is a special case of the FASP structure of Figure 9c.

The knowledgeable users may express which boolean operators should be used and in which place within their FASP preferences, while the naïve users are not forced to do so. The users may also specify the *preference conditions (PC)* that should hold so that the user preferences in the current FASP apply. A FASP contains the user's *classification preferences (CP)*, the *semantic creation preferences (BSCrP)* and the *source preferences (SP)* as well as nested FASPs, as shown formally in Expression 5.

FASP = (CP|PC|SP|**BSCrP**|FASP)pv ((**AND|OR**) (CP|PC|SP|**BSCrP**|FASP)pv)*

Expression 5: Formal syntax of a FASP

The semantic creation preferences (BSCrP) specify the boolean operators and the desired properties of the audiovisual content, including the preferred *title (TT)*, *keywords (K)*, *creator (Cr)*, *location (L)*, *creation tool (CT)* and the desired *semantic preferences (BSP)*. The formal syntax of the semantic creation preferences are expressed in regular expression syntax as shown in Expression 6.

BSCrP = ((TT|K|Cr|L|D|CT|**BSP**) pv) **((AND|OR)** ((TT|K|Cr|L|D|CT|**BSP**) pv))*

Expression 6: Formal syntax of the semantic creation preferences (BSCrP)

The semantic preferences are weighed semantic entity collections with boolean operators. The formal syntax of the semantic preferences is shown in Expression 7.

BSP = (T((OR|AND) T)*) pv)

Expression 7: Formal syntax of content preferences (BSP).

A special case of the generalized FASP syntax is the syntax of FASPs with implicit boolean operators. The semantic creation preferences for FASPs with implicit boolean operators do not contain boolean operators and are a special case of the semantic creation preferences of Expression 6. They have the MPEG-7 creation preferences, which in the special case may not have semantic content preferences.

An example of filtering and search preferences are the preferences of a user who is interested in France's goals (pv=90) and in France's penalties (pv=50). They are expressed in the formal syntax as shown in Expression 8. We have made the same assumptions as the example of Expression 7 and we additionally assume here that France's penalties are bound to the "APenalty" variable. We also assume that an abstract semantic entity exists, which has as id "Penalty", which represents the class of all the penalties.

UP2 = (((FGoal, EventType) AND ((exemplifies, Goal) AND (agent, France)) 90) OR ((APenalty, EventType) AND ((exemplifies, Penalty) AND (agent, France)) 50))

Expression 8: Formal syntax of the preferences of a user who is interested in goals (pv=80), in France's goals (pv=90) and in France's penalties (pv=50).

MPEG-7/21 Implementation of the Semantic User Preference Model. In order to implement the semantic user preference model described above as an extension of MPEG-7/21 we extended the MPEG-7/21 type hierarchy as shown in Figure 10 and described next.

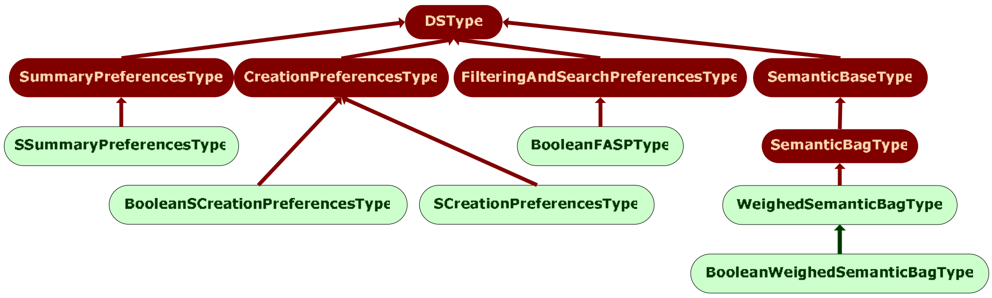

Figure 10: Extensions of the MPEG-7/21 Type Hierarchy

- We defined the *SCreationPreferencesType* type, for the representation of semantic creation preferences with implicit boolean operators. SCreationPreferencesType is a subtype of the MPEG-7 MDS *CreationPreferencesType* that represents the creation preferences of the users. The semantic preferences are set in its *SemanticPreferences* element, which is a weighed semantic entity collection. For the representation of weighed semantic entity collections we defined *WeighedSemanticBagType*, a subtype of the MPEG-7 type *SemanticBagType* (which is used for the representation of semantic entity collections).
- We defined the *SSummaryPreferencesType* type for the representation of semantic summary preferences. SSummaryPreferencesType is a subtype of the MPEG-7 MDS *SummaryPreferencesType* that represents summary preferences. SSummaryPreferencesType has a set of semantic summary preference description elements (namely *SummarySemantics*), which are of WeighedSemanticBagType.
- We defined the *BooleanFASPType* type for the representation of FASPs with explicit boolean operators. BooleanFASPType is a subtype of the MPEG-7 type *FilteringAndSearchPreferencesType*. BooleanFASPType has two attributes, both of *operatorType* (that takes one of the values "AND" and "OR"): (a) the *operatorFASP*, which represents the operator that specifies how the current FASP is combined with the FASPs existing in the same hi-

erarchy level; and (b) the *operator*, which represents the operator applied by the current FASP to combine its non-FASP components.

- We defined the *BooleanSCreationPreferencesType* type for the representation of semantic creation preferences with explicit boolean operators. BooleanSCreationPreferencesType is a subtype of CreationPreferencesType and extends it with *SemanticPreferences* elements, that essentially are weighed collections of semantic entities combined with a boolean operator. For the representation of collections of semantic entities on which a boolean operator is applied, we defined the *BooleanWeighedSemanticBagType*, a subtype of WeighedSemanticBagType.

An example of a semantic browsing preference description is shown in Figure 11, which corresponds to the formal syntax of Expression 4.

```
<UserPreferences id="UP1">
  <BrowsingPreferences>
    <SummaryPreferences xsi:type="SCreationPreferencesType">
      <SummarySemantics preferenceValue="100">
          <Label/>
          <SemanticBase xsi:type="EventType" id="FGoal">
             <Label/>
             <Relation type="agent" target="#France"/>
             <Relation type="exemplifies" target="#Goal"/>
          </SemanticBase>
        </SummarySemantics>
    </SummaryPreferences>
  </BrowsingPreferences>
</UserPreferences>
```

Figure 11: Semantic browsing preference description that corresponds to the regular expression of Expression 4.

An example of a semantic filtering and search preference description is shown in Figure 12, which corresponds to the formal syntax of Expression 8.

```
<UserPreferences id="UP2">
  <FilteringAndSearchPreferences xsi:type="BooleanFASPType" operator="OR"
operatorFASP="OR">
    <CreationPreferences operator="OR"
xsi:type="BooleanSCreationPreferencesType">
      <SemanticPreferences preferenceValue="90"
xsi:type="BooleanWeighedSemanticBagType" operator="OR">
        <Label/>
        <SemanticBase xsi:type="EventType" id="FGoal">
          <Label/>
          <Relation type="agent" target="#France"/>
          <Relation type="exemplifies" target="#Goal"/>
        </SemanticBase>
```

```
      </SemanticPreferences>
      <SemanticPreferences preferenceValue="50" operator="OR">
        <Label/>
        <SemanticBase xsi:type="EventType" id="APenalty">
          <Label/>
          <Relation type="agent" target="#France"/>
          <Relation type="exemplifies" target="#Penalty"/>
        </SemanticBase>
      </SemanticPreferences>
    </CreationPreferences>
  </FilteringAndSearchPreferences>
</UserPreferences>
```

Figure 12: Semantic filtering and search preference description that corresponds to the regular expression of Expression 8.

Our semantic user preference model is available in XML Schema as an MPEG-7/21 extension at: http://elikonas.ced.tuc.gr/ontologies/MP7FASPext/semUP.xsd.

The Application Ontology that captures the Semantic User Preference Model. In order to express the semantic user preference model in OWL, we defined an OWL application ontology using the methodology that we described in subsection 4.1 for the definition of the Upper Ontology. The integration of the semantic user preference model in the DS-MIRF framework allows us to take advantage of domain-specific extensions in the user preference descriptions. These extensions become available to be used for user preference specifications through the integration in DS-MIRF of domain knowledge expressed in OWL domain ontologies (see next subsection, 4.3, for the methodology of integration of OWL domain ontologies in DS-MIRF). The semantic user preference application ontology is an OWL-DL ontology available at http://elikonas.ced.tuc.gr/ontologies/AppOntos/SFASPAO/SUserPreferences.

4.3. Methodology for the Integration of OWL Domain Ontologies

We describe in this subsection the methodology that we developed for the definition and integration of OWL *Domain Ontologies* that extend the semantics encapsulated in the DS-MIRF Upper Ontology and the Application Ontologies with domain knowledge. The classes representing the domain-specific entities are defined in a way that extends the Upper Ontology and the Application Ontologies, since these classes are domain-specific specializations of the general-purpose constructs of the Upper Ontology and the Application Ontologies. The domain ontologies are defined according to the following methodological steps:

1. Domain-specific entity types (e.g. soccer players in the soccer domain) are represented by OWL classes that are subclasses of the appropriate Upper or Application Ontology classes. Domain-specific entities represent semantic entities appearing in the audiovisual content. They extend the *SemanticBaseType* class (that represents the semantic entities) and its subclasses (*EventType, ObjectType, AgentObjectType, SemanticPlaceType, Semantic-TimeType, SemanticStateType* and *ConceptType*). For example, in a soccer application the "PlayerObject" subclass of the "AgentObjectType" Upper Ontology class, is defined for the representation of soccer players.
   1.1. Features not present in the superclass are represented as additional object or datatype properties (e.g. the shirt number of a soccer player is represented as a datatype property in the domain of the "PlayerObject" class).
   1.2. Additional constraints may be applied on the properties inherited from the parent class, in order to guide the indexers to produce valid metadata (e.g. a soccer player should have an affiliation with a soccer team).
2. Additional restrictions for the general-purpose relationships expressed in the Upper Ontology and the typed relationship application ontology are usually needed (e.g. a "Goal" event may be related to player instances as goal agents). In these cases, properties are defined that permit relating relationships to the allowed domain-specific entities only.
   2.1. A subproperty of the "Relation" property ("Relation" links semantic entities with relationships) is defined. The domain of the property is the union of the classes to which belong individuals that are capable of being sources of a typed relationship and its range is the typed relationship class.

        As an example, assume that we would like to express the restriction that goals should be scored only by players. The "ScoresRelation" object property (subproperty of the "Relation" property), should be defined, having the "PlayerObject" class as domain and as range the "AgentRelationType" class (see Figure 13).

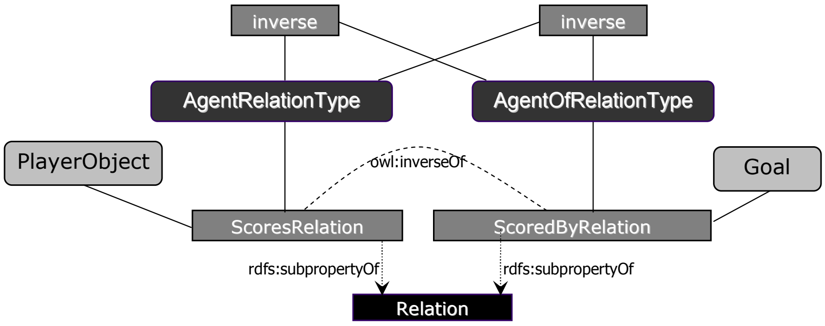

Figure 13: The "ScoresRelation" and "ScoredByRelation" object properties

2.2. The inverse property of the one defined in step 2.1 is defined in the domain of the classes the individuals of which are capable of being targets of the relationship. If the relationship used in step 2.1 is not directed, it becomes the range of the newly-defined property. If the relationship used in step 2.1 is directed, its inverse relationship becomes the range of the property. Since the "AgentRelationType" relationship is directed, in the example of step 2.1, the "ScoredByRelation" object property (inverse of the "ScoresRelation" property) should be defined, having the "Goal" class as domain and as range the "AgentOfRelationType" class.

In order to test the methodology that we described here, we have developed an OWL-DL soccer ontology (available at http://elikonas.ced.tuc.gr/ontologies/soccer.zip) and an OWL-DL Formula 1 ontology (available at http://elikonas.ced.tuc.gr/ontologies/F1.zip).

5. Multimedia Content Service Support

We present in brief in this subsection the support provided in the DS-MIRF framework for multimedia content services. In particular, we present in subsection 5.1 in brief the OWL to MPEG-7/21 transformation rules and in subsection 5.2 the retrieval and filtering support.

5.1. Transformation Rules

We present in this subsection the transformation rules that have been implemented in the GraphOnto component [25] and allow the transformation of: (a) Domain ontologies defined according to the methodology described in subsection 4.3 into abstract MPEG-7/21 semantic descriptions; (b) OWL individuals defined based on the domain ontologies into MPEG-7/21 semantic descriptions. The produced descriptions are valid MPEG-7/21 (parts of) documents.

During the metadata transformation from OWL to MPEG-7/21, the individuals representing MPEG-7/21 constructs or constructs defined in the application ontologies are transformed into XML elements. The object properties are transformed into elements and the datatype properties are transformed into the constructs they represent (attributes, elements or simple values). In order to produce valid MPEG-7/21 descriptions, information regarding the MPEG-7/21 XML element order and the MPEG-7/21 representation of the datatype properties is needed. This information is kept in a transformation rule ontology (available at http://elikonas.ced.tuc.gr/ontologies/Rules/OWL2MPEG7Rules) and is utilized during both ontology and metadata transformations.

A different approach is taken for the transformation of the OWL domain ontologies and the OWL individuals defined using them, which are both represented as MPEG-7 semantic elements, of type "SemanticBaseType". The "AbstractionLevel" element of the "SemanticBaseType" together with MPEG-7 semantic relationships are used to capture the ontology semantics. "AbstractionLevel" has the "Dimension" attribute, of non-negative integer type, which denotes if a semantic element is abstract and represents a class (when "AbstractionLevel.Dimension"≠0) or is non-abstract and describes an individual (when "AbstractionLevel" is absent or "AbstractionLevel.Dimension"=0). An abstract semantic entity that represents a domain-specific class is related with each of its subclasses through: (a) a "Relation" element of type "generalizes", having as source the class and as target the subclass; and (b) a "Relation" element of type "specializes", having as source the subclass and as target the class. In addition, an abstract semantic entity that represents a domain-specific class is related with each of the semantic entities representing the class individuals through pairs of "exemplifies"/"exemplifiedBy" relationships between the individual and the class.

The properties defined in the domain ontologies are transformed into "Property" elements (if they are of simple type) or into pairs of "property"/"propertyOf" relationships that associate semantic entities (if they are of complex type).

The reason for using this approach (instead, for example, from XML Schema subtyping for the representation of the domain ontology classes) is that this way full compatibility with MPEG-7 is maintained so that all the tools and applications that use MPEG-7 still work transparently with the produced MPEG-7 metadata.

In order to allow interoperability with both standard MPEG-7/21 applications and applications supporting the DS-MIRF semantic user preference model, we have defined, in addition to the above-described rules, a set of transformation rules for the conversion of the OWL user preferences into standard MPEG-7/21. During these transformations the semantic entities in the user preferences are systematically transformed into sets of keywords and the boolean operators are eliminated.

5.2. Retrieval and Filtering Support – Evaluation

The DS-MIRF retrieval and filtering support is based on semantic queries that may be specified by the end-users using appropriate query editors on top of the DS-MIRF framework. The semantic queries may have implicitly or explicitly specified preference values and boolean operators. Thus, the DS-MIRF semantic queries are distinguished into: (a) *Semantic queries (Q) with implicit boolean operators and preference values*, which are described by the regular expression of Expression 9, where T is a semantic entity formed according to Expression 3; (b) *Semantic queries (Qpv) with implicit boolean operators and explicit preference values*, which are described by the regular expression of Expression 10, where pv is a preference value in the range [-100, 100]; (c) *Semantic queries (QB) with explicit boolean operators and implicit preference values*, which are described by the regular expression of Expression 11; and (d) *Semantic queries (QBpv) with explicit boolean operators and preference values*, which are described by the regular expression of Expression 12.

$$Q = T^{+}$$

Expression 9: Formal syntax of a semantic query (Q) with implicit preference values and boolean operators

$$Qpv = (T\ pv)^{+}$$

Expression 10: Formal syntax of a semantic query (Qpv) with explicit preference values and implicit boolean operators

QB = T ((AND|OR) T)*

Expression 11: Formal syntax of a semantic query (QB) with implicit preference values and explicit boolean operators

QBpv = (T pv) ((AND|OR) (T pv))*

Expression 12: Formal syntax of a semantic query (QBpv) with explicit preference values and explicit boolean operators

In the Table 1 below we present some examples of DS-MIRF semantic queries using the formal syntax specified in Expression 9, Expression 10, Expression 11 and Expression 12. The queries use either the general constructs provided by MPEG-7/21 (queries 1, 2) or the MPEG-7/21 constructs and domain knowledge (queries 3, 4 and 5).

Query	Natural Language Description
`1. (Zidane, AgentObjectType) 100`	Give me the segments where Zidane appears (not only as a player!)
`2. ((D, SemanticTimeType) AND (after, D1)) AND ((D1, SemanticTimeType) AND (Time, 11/6/2004)) 100)`	Give me the segments referring to time after 11/6/2006
`3. ((Zidane, AgentObjectType) AND (exemplifies, Player)) 100`	Give me the segments where Zidane appears as a player
`4. ((BGoal, EventType) AND ((exemplifies, Goal) AND (agent, France)) 100)`	Give me the segments where France scores
`5. ((ZGoal, EventType) AND ((exemplifies, Goal) AND (agent, Zidane) AND (patient, Buffon)) 100)`	Give me the segments where the player Zidane scores against Buffon

Table 1: Semantic Query Examples

The knowledge-based capabilities of the DS-MIRF framework allow, in addition to the semantic querying, the support of advanced multimedia filtering functionality since the semantic user preferences essentially are semantic queries. As an example, the 4th query of Table 1 is equivalent with the user preference description expressed in Expression 4.

The examples of Table 1 clearly demonstrate that the queries that make use of the domain knowledge are more expressive than the ones that do not use domain knowledge. This will affect the performance in terms of precision/recall. We expect that the semantic queries supported by the DS-MIRF framework will also be more efficient in terms of precision/recall compared with keyword-based ones, as they allow to express precisely the user's criteria and are less sensitive in the completeness of the annotation (for example, as soon as Ronaldinho is affiliated with Barcelona, the goals he scores in the Spanish Championship are considered as Barcelona's goals in DS-MIRF, but this is not true in keyword annotations if "Barcelona" is not explicitly stated in a description).

In order to evaluate the effectiveness of the DS-MIRF framework, we have performed some experiments for multimedia content retrieval. In particular, we have compared the effectiveness of the semantic queries supported by the DS-MIRF framework with both keyword-based queries and queries expressed using the general-purpose constructs provided by MPEG-7/21 together with the textual descriptions of the semantic entities.

In order to perform reliable experiments, we have prepared a set of multimedia object annotations for multimedia content from FIFA World Cup (Mundial) 2006. In particular, we have annotated images of all the teams that participated in Mundial 2006, images of some of the players that participated in Mundial 2006, images, videos and video segments of the games of Mundial 2006 from the round of 16 on (several multimedia objects may show the same event – e.g. a goal may be shown in a separate video segment, in the soccer game video and in and image). These annotations contain keyword-based descriptions (in the *TextAnnotation* element of the MPEG-7 multimedia content descriptions) as well as semantic descriptions enriched with domain knowledge (in the *Semantic* element of the MPEG-7 multimedia content descriptions). The multimedia object annotations have been stored in the DS-MIRF MPEG-7/21 metadata repository.

We have specified a set of evaluation queries for soccer using the following procedure: First, we visited the website of a popular betting company[3] during Mundial 2006 and studied the bets available for the Mundial 2006 games. Then, we transformed the statistics-oriented bet expressions used in the website of the betting company to content-oriented queries. For ex-

[3] The betting company is *BetandWin*, www.betandwin.com

ample, the bet expression *"How many goals have been scored by Team X?"* has been transformed to the query *"Give me the multimedia objects showing the goals scored by Team X"*. This way, we collected more than 100 query types (not concrete queries) for soccer games (available at: http://elikonas.ced.tuc.gr/Queries/SoccerQueryTypes.doc). All these queries can be expressed according to our semantic query model, using the domain knowledge captured in our soccer ontology. Then, we ranked the importance of these queries together with 6 users, experts in multimedia applications and funs of soccer viewing, who are potentially interested to utilize a semantic service such as the one of the DS-MIRF. We also took into account the bets available for soccer games of lower importance that are also available in the web site of the betting company. Finally, based on all the above considerations, we selected for the first phase of the evaluation the natural language queries shown in Table 2.

Query ID	Natural Language Query
Q1	Give me the multimedia objects showing the goals scored by Italy during Mundial 2006, from the Round of 16 on.
Q2	Give me the multimedia objects showing the goals scored by France during Mundial 2006, from the Round of 16 on.
Q3	Give me the multimedia objects showing the goals scored by Italy during Mundial 2006, in the finals and semi-finals of Mundial 2006.
Q4	Give me the multimedia objects showing the goals scored by France against Brazil during the quarter-finals of Mundial 2006.
Q5	Give me the multimedia objects showing the goals scored by Italy against France during the final of Mundial 2006.
Q6	Give me the multimedia objects showing the penalty kicks of France during the game time of the games of Mundial 2006, from the Round of 16 on.

Table 2: Queries used during the Evaluation Phase

Our retrieval evaluation metrics were *precision (P)* and *recall (R)*. Precision is defined as *P=RDR/RD*, where RDR is the number of the relevant descriptions retrieved and RD is the total number of retrieved descriptions. Recall is defined as *R=RDR/TRD*, where RDR is the number of the relevant descriptions retrieved and TRD is the total number of relevant descriptions.

The queries were posed using keywords only, general-purpose MPEG-7 semantic constructs and their textual annotations (existing in the *Label* and *Definition* elements of the semantic entities) and MPEG-7 semantic constructs together with domain knowledge. The evaluation results for keyword-based queries are shown in Table 3, for queries utilizing general-purpose MPEG-7 semantic constructs and their textual annotations are shown in Table 4 and for queries MPEG-7 semantic constructs together with domain knowledge in Table 5.

Query ID	*Number of Relevant Descriptions*	*Number of Relevant Descriptions Retrieved*	*Total Number of Descriptions Retrieved*	*Precision*	*Recall*
Q1	6	6	8	0,75	1
Q2	11	11	15	0,73	1
Q3	4	4	6	0,67	1
Q4	3	3	3	1	1
Q5	3	3	5	0,6	1
Q6	4	4	7	0,57	1

Table 3: Query Evaluation Results for Keyword-based Queries

Query ID	*Number of Relevant Descriptions*	*Number of Relevant Descriptions Retrieved*	*Total Number of Descriptions Retrieved*	*Precision*	*Recall*
Q1	6	6	7	0,86	1
Q2	11	11	11	1	1
Q3	4	4	5	0,8	1
Q4	3	3	3	1	1
Q5	3	3	4	0,75	1
Q6	4	4	5	0,8	1

Table 4: Query Evaluation Results for Queries using general-purpose MPEG-7 Semantic Constructs and Textual Annotations

Query ID	*Number of Relevant Descriptions*	*Number of Relevant Descriptions Retrieved*	*Total Number of Descriptions Retrieved*	*Precision*	*Recall*
Q1	6	6	6	1	1
Q2	11	11	11	1	1
Q3	4	4	4	1	1
Q4	3	3	3	1	1
Q5	3	3	3	1	1
Q6	4	4	4	1	1

Table 5: Query Evaluation Results for Queries using general-purpose MPEG-7 Semantic Constructs and Domain Knowledge (DS-MIRF Semantic Queries)

We observe in the evaluation results that the two semantic-based approaches clearly outperform the keyword-based approach in terms of precision, and this is due to the unstructured form of the keyword-based descriptions, where the goals scored by Italy against France are described in the same manner with the goals scored by France against Italy. It is clear that for some query types such as *"Give me the multimedia objects showing the goals scored in the game Italy-France in the final of Mundial 2006"* the precision of the keyword-based approach will be 1, so there will be no difference between a semantic-based approach and a keyword-based one. This also happens if only one team has scored in a game and the query asks for the goals of the scoring team, like for example in Q4 that refers to the game France-Brazil, where only France scored.

The slight difference in the precision between the approach that uses general-purpose MPEG-7 semantic constructs and their textual annotations and the DS-MIRF approach that uses general-purpose MPEG-7 semantic constructs together with systematically expressed domain knowledge is due to the fact that a textual description may cause some false drops due to the way some words are used in the text (here, the false drops in Q1, Q3 and Q5 are due to the "near-goal" characterization of a single shot), while the systematic utilization of domain knowledge in DS-MIRF allows for the unambiguous description of the audiovisual content. The false drops are expected to increase when the textual annotations of the semantic entities describe – even in brief – that an event happened. Consider, as an example, the query *"Give me the multimedia objects showing the misconducts of Mundial 2006"*. If a segment shows the event where the referee shows a red card, and the textual annotation of the semantic entity that represents the red card explains that the red card is the result of a misconduct which however is not shown in this segment, false drops will occur.

The recall is 1 in all the evaluated queries, and for all the approaches. This is due to the simplicity of the queries, but we expect that, in more complex queries, both the keyword-based approach and the approach that uses general-purpose MPEG-7 semantic constructs and their textual annotations will have reduced recall. Consider, as an example, the query *"Give me the multimedia objects showing the misconducts of Mundial 2006"*. The descriptions which contain specific words for the misconducts (e.g. kick, butt etc.) will be retrieved only by the DS-MIRF approach that uses MPEG-7 semantic constructs together with domain knowledge. In addition, queries of the form *"Give me the multimedia objects showing the goals scored by*

Barcelona's players in Mundial 2006" cannot be expressed using the keyword-based approach. Both the precision and the recall are 1 for the DS-MIRF approach, since the detailed soccer ontologies used precisely describe the semantics of all the events used in the queries.

We plan to perform more extensive evaluation in the soccer domain using more semantic entities of the domain (players, red cards etc.) and more complex queries (the whole set of the query types we have collected will be evaluated). In addition, we will extend our evaluation in the formula 1 domain, for which we are selecting potential evaluation query types in the same way we did for the soccer domain. We plan also to collect such sets of query types and perform evaluation for other important domains outside the sports domain, including subdomains of broadcasted news (such as economical news).

This evaluation framework only allows evaluation of the approach in terms of retrieval effectiveness. Extensions will be needed in order to evaluate the framework in terms of other applications, like for example multimedia content filtering, knowledge acquisition, constraint checking, query formation, etc.

6. Conclusions – Future Work

In this paper we have presented the DS-MIRF framework, a software engineering framework that aims to facilitate the development of knowledge-based multimedia applications utilizing and extending MPEG-7/21. The DS-MIRF Framework provides support for interoperability of OWL with the MPEG-7/21 so that domain and application ontologies described in OWL can be transparently integrated with MPEG-7/21 metadata. This allows applications that recognize and use the MPEG-7/21 constructs (e.g. indexing, retrieval, filtering etc.) to make use of domain and application ontologies, resulting in more effective user retrieval and interaction with the audiovisual material. It also allows the standard-based semantic web methodologies and tools to be directly applied in MPEG-7/21 multimedia descriptions, which become first class citizens in the Semantic Web environment. We have presented a systematic methodology and software that supports this framework. We have also presented a generalized mechanism for supporting user preferences, which has as special case the MPEG-7/21 user preferences and it fully exploits the semantic content descriptions in MPEG-7/21. We have also presented a retrieval evaluation methodology and comparative retrieval results which are encouraging.

We believe that the methodology that we have presented here is generic, and its basic philosophy can be reused in other environments to make important standards first class citizens in the Semantic Web. For example, using similar steps, we have developed a framework for defining 3D scenes that follow standards like X3D and VRML compatible with the Semantic Web [26].

Our future research in the area includes: (a) More extensive experimentation using the evaluation framework that we described in subsection 5.2; (b) Extension of the Upper Ontology to fully represent the concepts of *MPEG-7 Visual* [20]; and (c) Utilization of a well-accepted top-level ontology such as the Suggested Upper Merged Ontology (SUMO) [14] and the Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) [27].

7. Acknowledgments

The work presented in this paper was partially funded in the scope of the DELOS II Network of Excellence in Digital Libraries (IST – Project Record Number 507618).

8. References

1. M. Doerr, J. Hunter, C. Lagoze, "Towards a Core Ontology for Information Integration", Journal of Digital Information, Volume 4 Issue 1, April 2003.
2. D. Fallside, “XML Schema Part 0: Primer”, W3C Recommendation, 2001. (http://www.w3.org/TR/xmlschema-0/).
3. D. L. McGuinness and F. van Harmelen (eds.). OWL Web Ontology Language: Overview. W3C Recommendation, 10 Feb. 2004. (http://www.w3.org/TR/owl-features).
4. J. Hunter, "Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology", International Semantic Web Working Symposium (SWWS), Stanford, July 30 - August 1, 2001.
5. ISO/IEC: 15938-5:2003: Information Technology – Multimedia content description interface – Part 5: Multimedia description schemes (2003) First Edition.
6. MPEG Group, “MPEG-7 (Multimedia Content Description Interface)”. (http://www.chiariglione.org/mpeg/index.htm).
7. M. Naphade, J. Smith: “A Hybrid Framework for Detecting the Semantics of Concepts and Context”, in Proceedings of CIVR 2003, Urbana, IL, July 24-25, 2003, pp. 196-205.
8. R. Troncy, “Integrating Structure and Semantics into Audio-visual Documents”, 2nd International Semantic Web Conference (ISWC), 20-23 October 2003, Sanibel Island, Florida, USA.

9. Tsinaraki C., Polydoros P., Christodoulakis S.: "Integration of OWL ontologies in MPEG-7 and TV-Anytime compliant Semantic Indexing". 16th International Conference on Advanced Information Systems Engineering (CAISE), pp. 398-413, June 2004, Riga, Latvia.
10. Tsinaraki C., Polydoros P., Christodoulakis S.: "Interoperability support for Ontology-based Video Retrieval Applications". Conference on Image and Video Retrieval (CIVR) 2004, pp. 582-591, July 2004, Dublin, Ireland.
11. J. Hunter, J. Drennan, S. Little: "Realizing the Hydrogen Economy through Semantic Web Technologies", IEEE Intelligent Systems, January/February 2004
12. D. L. McGuinness, R. Fikes, J. Hendler, L. A. Stein. "DAML+OIL: An Ontology Language for the Semantic Web", IEEE Intelligent Systems, Sep.-Oct. 2002, 17(5):72-80
13. C. Lagoze, J. Hunter, "The ABC Ontology and Model", Journal of Digital Information, 2(2), 6 Nov. 2001, http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Lagoze/lagoze-final.pdf
14. SUO WG, the "Standard Upper Ontology Working Group" (IEEE P1600.1) web site, http://suo.ieee.org/
15. H. Agius, M. Angelides: "Modelling and Filtering of MPEG-7-Compliant Meta-Data for Digital Video". In the Proc. of the ACM SAC, 2004, pp. 1248-1252
16. D. Rogers, J. Hunter, D. Kosovic: "The TV-Trawler Project". In Special Issue of the International Journal of Imaging Systems and Technology on Multimedia Content Description and Video Compression, March 2003.
17. B. Tseng, C.-Y. Lin, J. Smith: Using MPEG-7 and MPEG-21 for personalizing video. IEEE Multimedia, 11(1):42-52, Jan.-Feb. 2004.
18. Q. Wang, W.-T. Balke, W. Kiessling, A. Huhn: P-News: Deeply Personalized News Dissemination for MPEG-7 based Digital Libraries. In the Proc. of ECDL 2004.
19. ISO/IEC 21000-7:2004, "Information Technology – Multimedia Framework (MPEG-21) – Part 7: Digital Item Adaptation," 2004.
20. ISO/IEC: 15938-3:2001: Information Technology – Multimedia content description interface – Part 3 visual (2001) Version 1.
21. F. Pereira, The MPEG-21 standard: *Why an open multimedia framework?*, in Proceedings of the 8th International Workshop on Interactive Distributed Multimedia Systems (IDMS 2001), LNCS 2158, Lancaster, September 2001, Springer-Verlag, Heidelberg, pp. 219–220.
22. Guarino N.: "Formal Ontology and Information Systems", 1st International Conference "Formal Ontology in Information Systems" (FOIS '98), pp. 3-15, June 6-8 1998, Trento, Italy, Ed. Guarino N., IOS Press
23. D. Brickley and R. V. Guha (eds.), "RDF Vocabulary Description Language 1.0: RDF Schema", W3C Recommendation, 10 Feb. 2004. (http://www.w3.org/TR/rdf-schema)
24. F. Manola and Eric Milles (eds.), "RDF Primer. W3C Recommendation", 10 Feb. 2004. (http://www.w3.org/TR/rdf-primer)
25. P. Polydoros, C. Tsinaraki, S. Christodoulakis: "GraphOnto: OWL-Based Ontology Management and Multimedia Annotation in the DS-MIRF Framework". In the proceedings of the Workshop on Multimedia Semantics 2006 (WMS 2006), pp. 14-24, 19-21 June 2006, Chania, Crete.
26. Kalogerakis V., Moumoutzis N., Christodoulakis S.: "Coupling Ontologies with Graphics Content for Knowledge Driven Visualization". In the proceedings of the IEEE VR 2006.
27. The DOLCE Ontology, http://www.loa-cnr.it/DOLCE.html