Project no. FP6-507752

MUSCLE

Network of Excellence
Multimedia Understanding through Semantics, Computation and Learning

Analysis of open problems and interoperability limits of existing standards and tools

Start date of project: 1 March 2004 Duration: 48 months

Deliverable Type: CO
Number: 4
Nature: Report
Task: WP 9

Name of responsible: Ovidio Salvetti
Editor & editor's address: Gabriele Pieri

Revision 1.0

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	✓

Keyword List: Multimedia Metadata, Metadata Standards, Interoperability

Contents

Chapter 1

Introduction

The increasing role of multimedia (MM) resources in the real world (pictures, graphics, 3D models, audio, video, speech, etc.) needs to improve the automatic generation and extraction of semantic knowledge from multi-source data and to enhance the possibility of their computational interpretation and processing.

Scientific and industrial communities tend to create their own standards for their particular needs and often they use or implement even their specific tools for MM data processing or retrieval. This could cause an unrestricted growth of the number of possible standards giving rise to a difficult communication to integrate or share MM data.
The tools for MM metadata are somehow independent from the MM data standards, that is some of them are strictly related to specific data, thus limiting their potential application as an effective interoperability tool, while others are more generic due to their relative independency from a specific standard.
On a more abstract level, the problem of knowledge and semantics representation of MM data ranges from a lower processing level (e.g. segmentation) to a higher symbolic level (e.g. semantic description), thus increasing complexity.

From a previous state-of-the-art study performed both outside and inside the NoE we found that many metadata standards have been developed by a number of communities having particular requirements for different specific application contexts. This characteristic appears to be an interoperability limit of the existing standards.
Thus, two main aspects should be taken into consideration to achieve an efficient use of MM metadata in order to read, search and exchange MM data in the network. On one hand, a careful analysis is necessary to investigate the possibility of identifying a common MM metadata standard able to describe and represent these resources, also considering their heterogeneous nature and semantics.
On the other hand, following a more general approach, it should be essential to define a more abstract model, together with proper mapping tools, suitable to represent and transform the different metadata sets, considering the correlation among similar meanings (semantic mapping). This latter possibility would enable MM applications to make use of ontologies.

This document firstly introduces the mostly and commonly used MM metadata standards and tools highlighting the main reasons why they have been introduced, or created from other previously existing standards, in order to point out the reference framework of the project.

Then, a discussion is presented about the open problems regarding MM metadata integration and interoperability from a semantic point of view, also related to the available technologies.

Chapter 2

A reference framework of the project

The description of MM contents for various domains has given rise to a large number of metadata standardization initiatives that have been developed recently.
Many initiatives aimed at allowing the share, the exchange and the interoperability between large networks.

The main distinction among standards can be made according to the representation of MM content: (a) a specific representation for a domain, usually referred to as a standardised description scheme; (b) a representation which aims at the integration of more metadata standards related to different domains. This second one needs to provide richer metadata models and tools that allow the definition of description schemes for arbitrary domains (these standards are usually referred to as standardised metadata frameworks).

The following Table I [1] reports several metadata standards among the most used and cited, which can be intended as representative for a wide range of different domains.

A list of descriptive characteristics for each standard is shown, and in particular information about:

- standardization bodies,
- year of the last released version,
- MM data types described,
- application domains,
- semantic level of description
- capacity of automatic or manual metadata creation.

In Fig. 1, the role of a MM metadata standard in both frames of constructing and using MM information is shown.

An overview regarding the main requirements and motivations behind the creation of these major standards follows, to understand which are the main application domains involved and the directions of the involved initiatives.

	MARC	Dublin Core	CDWA	VRA Core	CSDGM	Z39.87	LOM	DIG35	METS	JPX	SMPTE Metadata Dictionary
Standardization Body	Library of Congress	Dublin Core Metadata Initiative (DCMI)	Art Information Task Force (AITF)	Visual Resource Association	Federal Geographic Data Committee (FGDC)	National Information Standard Organization (NISO)	IEE (LTSC)	Digital Imaging Group (DIG of I3A)	Digital Library federation (DLF)	Joint Photographic Experts Group (JPEG)	Society of Motion Picture and Television Engineers (SMPTE)
Year	Current version MARC 21 since 1999	Current version 1.1 since 1999	Current version 2.0 since 2000	Current version 3.0 since 2002	Update version since from 1998	2002	2002	Current version 1.1 April 2001	Last review 2001	2000	Last review 2004
MM Type	Any	Any	Any	Images	Any	Images	Any	Images	Any	Images	Any
Domain	Bibliographic media description	Bibliographic media description	Description of Art works	Description of images of Art works	Description of Geographic media	Description of still images	Description of educational media	Description of digital images	Description of digital objects	Description of digital images	Description of audio/video documents
Level	Largely semantic	Largely semantic	Largely semantic	Largely semantic	Semantic and technical	Technical	Largely semantic	Semantic and technical	Semantic and technical	Semantic and technical	Semantic and technical
Producibility	Mainly manual	Mainly manual	Mainly manual	Mainly manual	Manual and Automatic	Mainly automatic	Mainly manual	Mainly manual	Mainly manual	Mainly manual	Manual and Automatic

Table I. Selection of several MM metadata standards.

The Machine-Readable Cataloging (MARC) standard for the representation and communication of bibliographic and related information in machine-readable form was implemented following the requirements for a generalized interchange format to accommodate data describing all forms of materials susceptible to bibliographic description, as well as related information such authority, classification, community information, and holdings data.

The Dublin Core Metadata Initiative began in 1995 as a program to bring together librarians, digital library researchers, content providers and text-mark-up experts to improve discovery standard for information resources.
The original Dublin Core, merging small sets of descriptors, quickly drew the global interest of a wide variety of information providers in the arts, sciences, education, business, and government sectors.

CDWA stands for Categories for the Description of Works of Art, a metadata schema designed by of the Art Information Task Force, to "describe the content of art databases by articulating a conceptual framework for describing and accessing information about objects and images". It was released in February 1996 and its last version dates back to September 2000.

While CDWA is exhaustive in its list of elements needed to describe museum objects, it is not entirely satisfactory for the description of images, and in particular, does not cover all of the elements needed for the description of architectural and other site-specific works. In order to expand the concept to non-art objects and visual documents, the VRA Core Categories was created. Compared with CDWA, VRA Core Categories was designed to cover most visual materials. Even if it has not such comprehensive categories as CDWA, as Dublin Core, it provides a core set of elements, which can be expanded by adding new elements as needed.

The Content Standard for Digital Geospatial Metadata (CSDGM) was developed by the Federal Geographic Data Committee with the perspective of defining the information required by a prospective user to determine the availability of a set of geospatial data, to determine the fitness of the set of geospatial data for an intended use, to determine the means of accessing the set of geospatial data, and to successfully transfer the set of geospatial data.

The metadata standard defined by the National Information Standard Organization (NISO) for still images is the Z39.87 and was developed with the purpose of defining a standard set of metadata elements for digital images. Standardizing the information allows users to develop, exchange, and interpret digital image files. It has been designed to facilitate interoperability among systems, services, and software, as well as to support the long-term management of and the continuous access to digital image collections.
The purpose of this standard is to facilitate the development of applications oriented to validate, manage, migrate, and process images of enduring value. Such applications are viewed to be essential components of large-scale digital repositories and digital asset management systems.

The IEEE Learning Technology Standards Committee defined metadata as information about an object, both physical and digital. Considering that the number of objects continues to grow exponentially and equally the need for learning is dramatically expanding, the lack of information or metadata about objects has produced a critical and fundamental constraint on the ability to discover, manage and use the objects themselves. To address this problem, the IEEE LTSC LOM working group has created a standard for "Learning Object Metadata".

Formed in April 1999, the vision of the DIG35 Initiative Group was to "provide a standardized mechanism which allows end-users to see digital image use as being equally as easy, as convenient and as flexible as the traditional photographic methods while enabling additional benefits that are possible only with a digital format".
By establishing standards, the Initiative Group sought to overcome a variety of challenges that have arisen as the sheer volume of digital images being used was increased. Among these there are efficient archiving, indexing, cataloguing, reviewing, and retrieving of individual images, whenever and wherever needed.

Many standards have been designed to encode metadata for objects held within digital library collections: what the Metadata Encoding and Transmission Standard (METS) attempted to fill was the lack of an overall *framework* within which these schemes could be integrated. This was a new emergent standard designed to encode metadata for electronic texts, still images, digitised videos, sound files and other digital materials contained in electronic library collections. Mainly, it attempted to address the lack of standardization in digital library metadata practices, which is one of the reasons that currently contrast the growth of coherent digital collections.

The Society of Motion Picture and Television Engineer developed a Metadata Dictionary (SMPTE 335M-2001) to define a registry of metadata element descriptions useful for associating both specific information and other metadata. The metadata dictionary structure covers the use of metadata for all types of essence (video, audio, and data in their various forms).

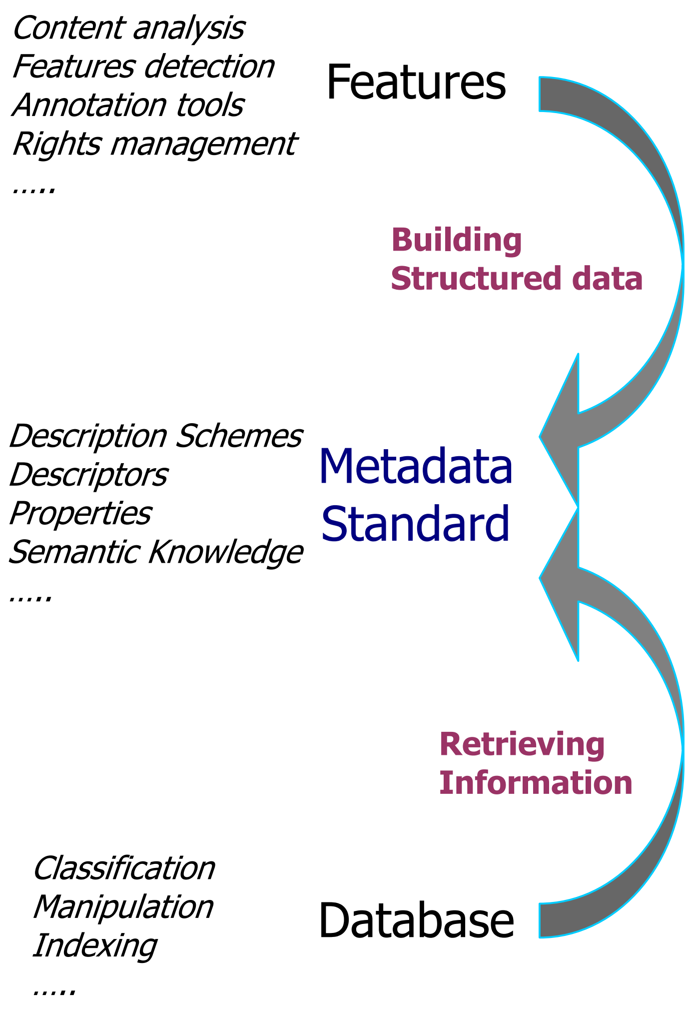

Fig. 1 – The role of a multimedia metadata standard.

Chapter 3

Open problems and interoperability criteria analysis

Aiming at the integration of different applications and community-specific metadata vocabularies, individual metadata efforts should concentrate on classifying and expressing semantics tailored toward focused functional and community needs.
A challenge for this integration is the interoperability of multiple metadata standards that are associated with and across resources (see Fig. 2). Different metadata standards are not semantically distinct, but overlap and relate to each other in numerous ways.
A first straight approach to interoperability would be a one-to-one crosswalk, but such an approach does not scale to the many metadata vocabularies that will continue to develop. A more scalable solution is through the exploitation of the fact that many entities and relationships are so frequently present and that they do not belong into the domain of a particular vocabulary but are across all of them [2].

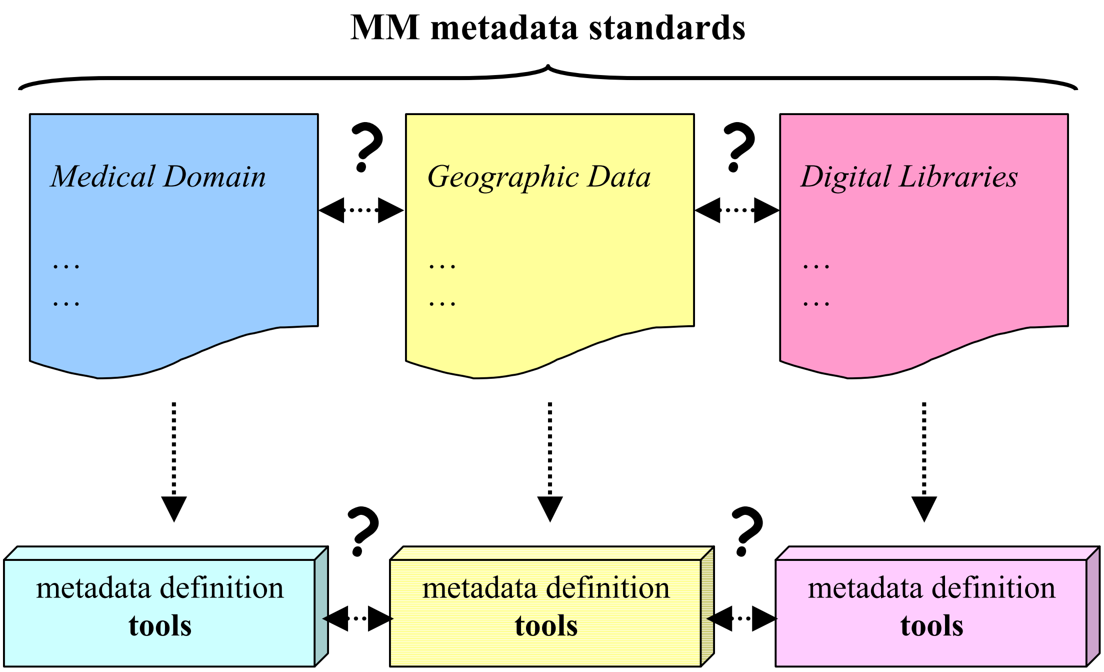

Fig. 2 - The integration of MM data is problematic, due to different MM metadata standards defined for specific domains, often with specific associated tools.

Semantics expression

The development of MM mining tools for the knowledge and semantic analysis brought to a need for ontologies which can express key entities and relationships which are used in the MM data description, two main examples of this ontologies are presented: the Resource Description Framework (RDF) Schema and the Web Ontology Language (OWL). These two examples can be identified as ontologies that provide knowledge representation able to develop complex services and tools to perform automatic knowledge-based reasoning, integration, sharing, and acquisition, especially working with audiovisual MM data having a specific and detailed semantic.

RDF is defined as a general purpose language to represent information in the web. This definition introduces RDF Schema which is a standard to describe how to use RDF to express RDF vocabularies on the Web [3].

RDF Schema can be seen as a semantic extension of the RDF language, providing mechanisms for the description of resources, groups of them, and most important relationships among them. The resources described can be used to establish characteristics of other resources, such as domains and properties.

These properties can be identified as the attributes of the described resources, thus corresponding to traditional attribute-values, but they can also represent the semantic relationship between resources.
On the opposite, RDF does not provide mechanisms for the description of these properties, and more important for the description of the relationships between properties and other resources (general axioms) [4].

In the effort to provide a better integration and interoperability of data among descriptive communities, the World Wide Web Consortium (W3C) Web Ontology Working Group is building upon the RDF Core work a language for defining structured web based ontologies (OWL Web Ontology Language) [5].

The OWL Web Ontology Language is intended to provide a language that can be used to describe the classes, and relations between them, that are inherent in Web documents and applications. The semantics of this ontology specifies how to derive logical consequences, i.e. facts which are not present in the ontology but that are derived through the semantics. These derivations can be based on multiple distributed documents by means of defined OWL mechanisms.

Through the OWL ontologies, tools can be available to provide general support that is non-specific to a particular domain.
The effort to build a useful and manageable reasoning system is not simple, while the construction of ontology is a more tractable problem. The benefits arising from tools based on the formal properties of the OWL language, which deliver a variety of competences, will make the ontology construction likely to be a good path.

On the other hand, OWL has expressive limitations, particularly concerning what can be said about properties. In some application domains, the restriction imposed by these limitations can be difficult to walk around. In particular, it may be necessary to relate inputs and outputs of composite processes to the inputs and outputs of their component processes, and in the language provided by OWL there is no composition constructor.
For example, an obvious relationship between the composition of *parent* and *brother* properties and the implied *uncle* property cannot be expressed [6].

Chapter 4

A dynamic description scheme

Within the Moving Pictures Expert Group (MPEG) the development of the Multimedia Content Description Interface (MPEG-7), aims at the definition of standardized tools for the automatic production and understanding of MM data descriptions in the tasks of retrieving, categorising and filtering information.
While prominent progress has been accomplished in automatic segmentation, recognition and detection of low-level features in MM data, less development has been made about the automatic generation of the semantic description of MM data.

The ISO/IEC standard MPEG-7 for MM content description can be classified as a MM metadata standard, but with respect to others above mentioned, it is not defined for a restricted application domain. This standard is intended to be applicable to an open range of application domains [7].

The latest initiative of the MPEG group, MPEG-21 [8], can be referred to as a framework whose goals are to:

- provide a framework to enable the use of MM resources through networks;
- integrate components and standards in order to harmonise the technologies that are able to create, manage and operate with MM content;
- collaborate with other standardisation bodies for the development of specifications and standards responding to functional requirements in order to go toward a MM framework

However, the actual main effort of the MPEG-21 initiative is focused on the identification of digital items and the definition of a Right Expression Language and a Rights Data Dictionary.

Regarding the definition of their description schemes, both MPEG-7 and MPEG-21 use the XML Schema language, which provides support for the definition of constraints for the structure, the cardinality and the data-types.
On the other hand, it lacks support for the definition of semantic knowledge needed for an efficient mapping, integration and knowledge acquisition.

Considering the trend toward the exponential growth of MM data resources distributed in many different places, the use of tools provided by MPEG-7 gives the capability to identify, search, index and disseminate information about content [9].

Chapter 5

Toward the definition of an upper-ontology

Reporting words from the W3C organisation, "The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation" [10].

The idea behind the Semantic Web is to allow data sharing and reusing among applications in industrial and scientific communities. It has been defined in a joint collaboration between various researchers and industrial partners.

Furthermore, we can find two main goals for this collaboration:

- data integration

- intelligent support for end-user

Usually, solutions for the first problematic are implemented ad hoc, through specific mapping between different schemas.
The second goal is to give a more automatic support to the users in finding, filtering and combining information from different sources.

Through the definition of ontologies, the functionality to improve Web searches can be powered by putting in relation the content information about a resource with specific structures representing associated knowledge and defined inference rules.

A formal definition of ontology is given in [11]: an explicit formal specification of how to represent the objects, concepts and other entities that are assumed to exist in some area of interest and the relationships that hold among them.

In other words, it consists of concepts, axioms and relationships describing a domain. Furthermore, it can be seen as a detailed and structured dictionary enabling an automatic processing of content information.

Ontologies can be distinguished between single domain ontologies (i.e. describing a specific field) and upper level ontologies (i.e. describing information concepts and relationships in a generic domain as expressively as the natural language). An upper level ontology defines structures and concepts upon which single domain ontologies could be implemented. An upper ontology is defined through abstract concepts, which are generic enough to be exploited by a wide range of domains. They are especially suitable for MM data interoperability and integration.

In practice, ontologies have to be defined in such a way to grant that different resources can be gathered to the same ontology when they have a shared meaning. Furthermore, an upper ontology should be defined in such a way to be extended 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.

Research communities working on standards are developing upper ontologies in order to achieve interoperability among metadata, and integration of MM data.

From what mentioned, an upper ontology is necessary to permit the integration of multi-source MM information. In fact, it is essential for accurate mappings between metadata vocabularies and the construction of services, such as cross-domain searching, tracking, browsing, data mining and knowledge acquisition. By combining metadata from various initiatives (Dublin Core, MPEG-7, MPEG-21, CIDOC/CRM, …), an upper ontology could also provide a basis for semantic interoperability and the development of services based on deductive inferencing [12][13].

The advantages of defining domain-independent upper ontologies can be identified in the overlaps, redundancies and incompatibilities between the semantics of terms used for example in both MPEG-7 and MPEG-21. Moreover, providing a common model having a single set of semantic definitions, would facilitate the efficiency and interoperability of MM systems based on the lower-level integrated standards.

Interesting projects focused to the development of upper ontologies are available in literature (see for instance [14]).

In the following, one of these (ABC [15]) is described in more detail.

ABC: an example of core ontology

Within the Harmony Project a top-level ontology has been defined with the goal of integrating MM data in the domain of digital libraries and making possible the interoperability between metadata schemas (model called ABC Ontology [15]).

The ABC upper level ontology follows the above discussed problem regarding many metadata approaches which are based on the assumption that predefined objects are invariable, together with their attributes [12]. This last hypothesis is not sufficient to describe the creation, evolution, and usage over time of different objects. Also advanced queries are not possible with such approach.
The ABC Ontology is basically a primitive ontology, that is a core for building other ontologies [16], providing domain independent base classes that can be used with domain specific properties or sub-classes.
In this model, tools can be implemented to obtain a mapping between metadata standards. Thus, ABC Ontology is planned to be a basis to put together semantics of single domain standards or even ontologies. In such a way, the main objective of semantic interoperability is fully achieved.
In particular, ABC Ontology is based upon the semantics of MPEG-7 standard and the right data dictionary in MPEG-21.
The model is designed for modelling generic physical, digital and similar MM objects present in libraries, various archives or museums and on the Web.
Abstract concepts can also be described through this ontology, for example changes in times or life-cycle events occurring to an object.

Chapter 6

Conclusions

After a brief introduction of the reference framework for our project, recalling the mostly and commonly used MM metadata standards in the scientific and industrial communities, this document focused on the discussion regarding the open problems of integration and interoperability between different MM metadata standards and tools.

Considering the actual situation, where each scientific and industrial community tends to create its own standards and tools required for its particular needs, an increase of the available standards is to be expected.
This growth will probably generate difficulties for MM data integration and sharing.

Thus, two main possibilities to reach an efficient interoperability have been identified for the search and exchange of MM data: on one hand, a common MM metadata standard format to represent heterogeneous MM resources and their semantics, and on the other hand more abstract models or mapping tools, able to represent and transform among metadata sets having correlated elements (see Fig. 3).

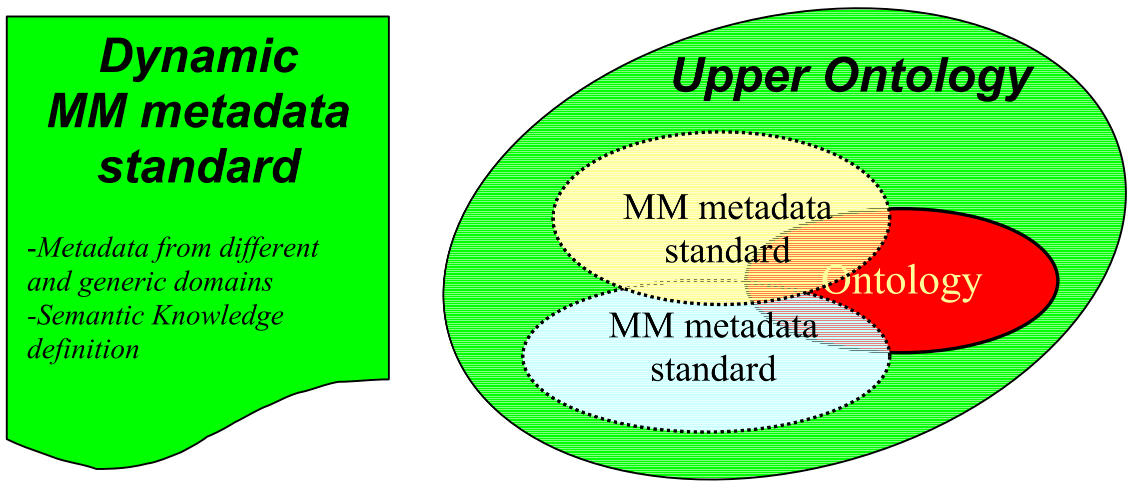

Fig. 3 - Two possible solutions for MM data integration and interoperability: dynamic single metadata standard (left) or definition of an upper ontology (right).

As an example of the first typology, MPEG-7 has been discussed. This metadata standard owns the characteristic of being dynamic and thus it is adaptable to different application domains, also regarding the semantic knowledge description.

Regarding the second typology, which implements an abstract model, its main concept of representing metadata sets from different domains with correlated elements has been discussed. As an actual example, the ABC core ontology has been introduced.

The first solution has the advantage of being a commonly established and widely used standard, but has a big disadvantage in the hard task to find the application domains which can be defined as the base for an almost-universal crosswise standard.

The second solution has the advantage of being more independent from the lower metadata standards, in such a way the definition of upper ontologies covering multiple domains is more efficient, but on the other hand there is an intrinsic complexity in the definition of a complete new ontology.

Both solutions for MM data integration and interoperability have advantages and disadvantages, but they represent the only possible solutions to solve the limits introduced by the unlimited expansion of the possible metadata standards.

In a future work, at the light of the NoE requirements, our study will be focused on two tracks, (a) by exploring the task of defining an upper ontology based on specific standards developed by the various scientific and industrial communities, and (b) by directly investigating the possibility to exploit a single dynamic metadata standard.

References

[1] Di Bono M. G., Pieri G., Salvetti O.: "WP9: a Review of Data and Metadata Standards and Techniques for Representation of Multimedia Content", Deliverable 9.1, MUSCLE Network of Excellence, FP6-5077-52, August 2004.

[2] Lagoze C., Hunter J., Brickley D.: "An Event-Aware Model for Metadata Interoperability", ECDL 2000, Lisbon, September 2000.

[3] Resource Description Framework (RDF), http://www.w3.org/RDF/

[4] Heflin J., Hendler J.: "Semantic Interoperability on the Web", Proceedings of Extreme Markup Languages 2000. Graphic Communications Association, pp. 111-120, 2000.

[5] Web Ontology Language (OWL), http://www.w3.org/TR/owl-guide/

[6] Horrocks I., Patel-Schneider P. F.: "A Proposal for an OWL Rules Language", Proceedings WWW2004. May 17-24, New York, USA, 2004.

[7] MPEG-7, http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm

[8] MPEG-21, http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm

[9] Westermann U., Klas W.: "An Analysis of XML Database Solutions for the Management of MPEG-7 Media Descriptions", ACM Computing Surveys, Vol. 35, N. 4, pp. 331-373, December 2003.

[10] Semantic Web, http://www.w3.org/2001/sw/

[11] Gruber T. R.: "A translation approach to portable ontologies", Knowledge Acquisition, 5(2):199-220, 1993.

[12] Hunter J.: "Enhancing the Semantic Interoperability of Multimedia through a Core Ontology", IEEE Transactions on Circuits and System for Video Technology, Special Issue on Conceptual and Dynamical Aspects of Multimedia Content Description, January 2003.

[13] Hunter J.: "Working Towards MetaUtopia - A Survey of Current Metadata Research", Library Trends, Organizing the Internet, Edited by Andrew Torok, 52(2), Fall 2003.

[14] Hunter J.: "Combining the CIDOC/CRM and MPEG-7 to Describe Multimedia in Museums", Museums on the Web, Boston, April 2002.

[15] The Harmony Project, http://metadata.net/harmony/

[16] Lagoze C., Hunter J.: "The ABC Ontology and Model", Journal of Digital Information, Volume 2, Issue 2, November 2001.