D1.3.2 Identification of standards on metadata for ontologies

M. Carmen Suárez-Figueroa

Raúl García-Castro, Asunción Gómez-Pérez, Raúl Palma (UPM)
Lyndon J.B. Nixon, Elena Paslaru (FU Berlin)
Jens Hartmann (UKARL)
Mustafa Jarrar (VUB)

Abstract.
EU-IST Network of Excellence (NoE) IST-2004-507482 KWEB
Deliverable D1.3.2 (WP1.3)
This deliverable presents a set of structured metadata for evaluated and certified ontologies.

Document Identifier:	KWEB/2004/D1.3.2/v1.0
Class Deliverable:	KWEB EU-IST-2004-507482
Version:	V1.0
Date:	31/01/2005
State:	Final
Distribution:	Public

Knowledge Web Consortium

This document is part of a research project funded by the IST Programme of the Commission of the European Communities as project number IST-2004-507482.

University of Innsbruck (UIBK) – Coordinator
Institute of Computer Science,
Technikerstrasse 13
A-6020 Innsbruck
Austria
Contact person: Dieter Fensel
E-mail address: dieter.fensel@uibk.ac.at

École Polythechnique Fédérale de Lausanne (EPFL)
Computer Science Department
Swiss Federal Institute of Technology
IN (Ecublens), CH-1015 Lausanne.
Switzerland
Contact person: Boi Faltings
E-mail address: boi.faltings@epfl.ch

France Telecom (FT)
4 Rue du Clos Courtel
35512 Cesson Sévigné
France. PO Box 91226
Contact person : Alain Leger
E-mail address: alain.leger@rd.francetelecom.com

Freie Universität Berlin (FU Berlin)
Takustrasse, 9
14195 Berlin
Germany
Contact person: Robert Tolksdorf
E-mail address: tolk@inf.fu-berlin.de

Free University of Bozen-Bolzano (FUB)
Piazza Domenicani 3
39100 Bolzano
Italy
Contact person: Enrico Franconi
E-mail address: franconi@inf.unibz.it

Institut National de Recherche en Informatique et en Automatique (INRIA)
ZIRST - 655 avenue de l'Europe - Montbonnot Saint Martin
38334 Saint-Ismier
France
Contact person: Jérôme Euzenat
E-mail address: Jerome.Euzenat@inrialpes.fr

Centre for Research and Technology Hellas / Informatics and Telematics Institute (ITI-CERTH)
1st km Thermi – Panorama road
57001 Thermi-Thessaloniki
Greece. Po Box 361
Contact person: Michael G. Strintzis
E-mail address: strintzi@iti.gr

Learning Lab Lower Saxony (L3S)
Expo Plaza 1
30539 Hannover
Germany
Contact person: Wolfgang Nejdl
E-mail address: nejdl@learninglab.de

National University of Ireland Galway (NUIG)
National University of Ireland
Science and Technology Building
University Road
Galway
Ireland
Contact person: Christoph Bussler
E-mail address: chris.bussler@deri.ie

The Open University (OU)
Knowledge Media Institute
The Open University
Milton Keynes, MK7 6AA
United Kingdom.
Contact person: Enrico Motta
E-mail address: e.motta@open.ac.uk

Universidad Politécnica de Madrid (UPM)
Campus de Montegancedo sn
28660 Boadilla del Monte
Spain
Contact person: Asunción Gómez Pérez
E-mail address: asun@fi.upm.es

University of Karlsruhe (UKARL)
Institut für Angewandte Informatik und Formale Beschreibungsverfahren – AIFB
Universität Karlsruhe
D-76128 Karlsruhe
Germany
Contact person: Rudi Studer
E-mail address: studer@aifb.uni-karlsruhe.de

University of Liverpool (UniLiv)
Chadwick Building, Peach Street
L697ZF Liverpool
United Kingdom
Contact person: Michael Wooldridge
E-mail address: M.J.Wooldridge@csc.liv.ac.uk

University of Manchester (UoM)
Room 2.32. Kilburn Building, Department of Computer Science, University of Manchester,
Oxford Road
Manchester, M13 9PL
United Kingdom
Contact person: Carole Goble
E-mail address: carole@cs.man.ac.uk

University of Sheffield (USFD)
Regent Court, 211 Portobello street
S14DP Sheffield
United Kingdom
Contact person: Hamish Cunningham
E-mail address: hamish@dcs.shef.ac.uk

University of Trento (UniTn)
Via Sommarive 14
38050 Trento
Italy
Contact person: Fausto Giunchiglia
E-mail address: fausto@dit.unitn.it

Vrije Universiteit Amsterdam (VUA)
De Boelelaan 1081a
1081HV. Amsterdam
The Netherlands
Contact person: Frank van Harmelen
E-mail address: Frank.van.Harmelen@cs.vu.nl

Vrije Universiteit Brussel (VUB)
Pleinlaan 2, Building G10
1050 Brussels
Belgium
Contact person: Robert Meersman
E-mail address: robert.meersman@vub.ac.be

Changes

Version	Date	Author	Changes
0.1	24/09/2004	M. C. Suárez-Figueroa	Initial Setup
0.11	5/10/2004	M. C. Suárez-Figueroa	Initial Proposed Metadata
0.12	28/10/2004	M. C. Suárez-Figueroa	Updates from Phone Conference
0.13	30/10/2004	A. Gómez-Pérez	Knowledge Web Standardization Strategy
0.14	1/11/2004	R. Palma M. C. Suárez-Figueroa	SoA on Metadata and Standards
0.15	5/11/2004	R. García-Castro A. Gómez-Pérez	Oyster
0.16	11/11/2004	R. Palma M. C. Suárez-Figueroa	Metadata Examples: Dublin Core and FOAF
0.17	15/11/2004	R. Palma M. C. Suárez-Figueroa	Metadata Examples: CSDGM, EdNA, and MM
0.18	16/11/2004	E. Paslaru L. J.B. Nixon	Metadata Example: LOM Metadata Example of Ontology: SWPATHO Initial Description ot the Proposed Metadata
0.19	16/11/2004	M. C. Suárez-Figueroa	Revision and update of the Proposed Metadata
0.2	16/11/2004	J. Hartmann	Metadata Example: BibTeX Example of Ontology: OntoWeb Ontology
0.21	19/11/2004	A. Gómez-Pérez	Introduction
0.22	22/11/2004	A. Gómez-Pérez	Revision
0.23	29/11/2004	E. Paslaru	Annotated SWPATHO Ontology Classification of the Proposed Metadata
0.24	30/11/2004	J. Hartmann	Annotated OntoWeb Ontology
0.25	3/12/2004	M. Jarrar	Formalization of the Proposed Metadata
0.26	7/12/2004	M. Jarrar	Metadata Example: DogmaModeler Ontology Metadata
0.27	14/12/2004	M. Jarrar M. C. Suárez-Figueroa	Updates from WP Leader
1.0	31/01/2005	Hamish Cunningham M. C. Suárez-Figueroa	Updates from Quality Controller

Executive Summary

This deliverable presents a set of structured metadata for evaluated and certified ontologies. This set of metadata for describing ontologies has taken as input the OntoRoadMap ontology from the OntoWeb[1] project. This deliverable also includes a survey of metadata and standards, and an identification widely used of metadata examples (Dublin Core, LOM metadata, FOAF, BibTeX, etc.).

The document is structured as follows:

- Section 1 briefly describes the main objectives of the tasks 2 and 3 of the Workpackage 1.3.
- Section 2 presents the state of the art on metadata and standards.
- Section 3 lists several metadata examples, which include the description, a subset of elements, and the references.
- Section 4 presents the proposed metadata for describing ontologies. These metadata have been classified depending on their features: syntactical, semantical, heuristic and pragmatic, and contextual.
- Section 5 includes the SWPATHO ontology and the OntoWeb ontology annotated with the metadata defined in the previous section.
- Section 6 presents a short description of the Oyster system.
- Section 7 explains the Knowledge Web standardization strategy.

[1] http://www.ontoweb.org/

Contents

1. Introduction

by UPM.

One of the aims of the Knowledge Web Network of Excellence (NoE) is *Standardization*. Ontologies are meant as a shared means of communication between computers and between humans and computers. To achieve this, ontologies should be represented, exchanged and accessed in agreed-upon open standards.

In order to achieve this aim, the Knowledge Web NoE includes WorkPackage 1.3 whose main objectives are the following:

- To make recommendations about how to develop and to deploy ontologies and semantic web applications in particular business cases and with regard to industrial needs for semantic processing.
- To define registries where ontology developers will be able to publish the structured metadata of evaluated and certified ontologies. Development of end-user methods with corresponding tools support to measure the usability and usefulness of ontologies on the registries for an application.
- Standardization of structured metadata to describe the ontology field in contact with standardization bodies (W3C, ISO, IEEE, CEN, ...).

The WorkPackage 1.3 of the Knowledge Web NoE contains two main tasks to achieve the aforementioned aim.

- **Task 2. Ontology repository**. Registries where ontology developers will be able to publish the structured metadata of evaluated and certified ontologies will be specified. Development of end-user methods with corresponding tools support to measure the usability and usefulness of ontologies in the registries for an application will be defined.

  In addition to the recommendation of ontologies for a particular type of use context, how the ontology is used and processed in semantically driven IT systems is the next question to address. Standard or reusable components or tools can form a development framework, library or toolbox of strategic interest for Industry and enable faster take-up of the technology.

- **Task 3. Standards**. Standards of structured metadata to describe the ontology content are a recognized need. Search will be made to find the suitable standardization bodies (W3C, ISO, IEEE, CEN, etc.). Other areas of possible standards will be examined, such as interoperability of tools and knowledge processing interfaces, as identified during the experimentation on the Semantic Web Framework.

In this deliverable we propose a set of metadata for describing ontologies. These metadata are needed to define and create the ontology resgistries. We propose also these metadata for standardization.

2. State of the Art on Metadata and Standards

by UPM.

2.1. Introduction

The term **Metadata** has become particularly common with the popularity of the World Wide Web. But the underlying concepts have been in use for as long as collections of information have been organized. Until the mid-1990s, Metadata was a term most prevalently used by communities involved with the management and interoperability of geospatial data, and with data management and systems design and maintenance in general. For these communities, Metadata referred to a suite of industry or disciplinary standards as well as additional internal and external documentation and other data necessary for the identification, representation, interoperability, technical management, performance, and use of data contained in an information system [3].

Among the large amount of definitions available for Metadata, we can find some dictionary-oriented. One of this states that the prefix "meta-" is defined (by The Macquarie Dictionary[2]) as meaning "among", "together with", "after" or "behind". That suggests the idea of a "fellow traveler": that Metadata is not fully fledged data, but it is a kind of fellow-traveler with data, supporting it from the sidelines. This means that is inherent in the concept of Metadata that there is an association of some kind between the Metadata and the information resource which it describes [1]. In the same line another definition states that the word Metadata shares the same Greek root as the word metamorphosis. Meta means change (it also denotes a nature of a higher order or more fundamental kind) and Metadata, or "data about data" describes the origins of and track the changes to data [2].

Other definitions are more application-oriented: Metadata, literally "data about data", is an increasingly ubiquitous term that is understood in different ways by the diverse professional communities that design, create, describe, preserve, and use information systems and resources [3].

We also have data-oriented definitions: Metadata describes the content, quality, condition, and other characteristics of data. Metadata answers who, what, when, where, why, and how questions about every facet of the data that are being documented in a consistent and precise format, so that a potential data user can decide whether the data is appropriate for his/her use. That means that Metadata can relieve potential data users of having to have full advance knowledge of a dataset's existence and characteristics [4, 5, 6, 7].

And finally we can find definitions based on the architecture of Metadata (W3C[3]) which states that Metadata consists of assertions about data, and such assertions typically, when

[2] http://www.macquariedictionary.com.au
[3] http://www.w3.org/

represented in computer systems, take the form of a name or type of assertion and a set of parameters. The Metadata can be represented as a set of independent assertions. This model implies that in general, two assertions about the same resource can stand alone and independently. When they are grouped together in one place, the combined assertion is simply the sum (actually the logical AND) of the independent ones. Therefore (because AND is commutative) collections of assertions are essentially unordered sets. This design decision rules out for example, in simple sets of data, assertions which are somehow cumulative or later ones override earlier ones. Each assertion stands independently of others [8].

Whichever definition you use, fundamentally Metadata describes an information resource, helps to increase its accessibility and gives other useful resource information that help to its management (see section 2.3). A collection of such Metadata elements may describe one or many information resources and can help [4]:

- To find datasets.
- To determine whether the dataset is appropriate for certain uses (fitness-for-use).
- To manage datasets (a record of changes, updates, procedures used, etc.).
- To extract information needed to correctly use datasets (perform projection transformations, select on certain attributes, etc.).

Note that we are not dealing in this document with metadata in the sense of semantic annotation that connects ontologies and texts, but in the one of describing ontologies.

2.2. What is Metadata?

A basic implementation of a Metadata record generally consists of a set of pre-defined elements (sometimes called fields or attributes) which describes different parts of a resource, and each element can have one or more values. For example, a Metadata record describing a book may contain author, title and publisher elements [9, 10].

Each Metadata schema will usually have the following characteristics [9]:

- A limited number of elements.
- The name of each element.
- The meaning of each element.

A broader definition of Metadata is as “the sum total of what one can say about any information object at any level of aggregation”. An information object is anything that can be addressed and manipulated by a human or a system as a discrete entity. The object may be comprised of a single item, or it may be an aggregate of many items. In general all information objects, regardless of the physical or intellectual form they take, have three features all of which can be reflected through Metadata [2]:

- *Content* relates to what the object contains or is about, and is *intrinsic* to an information object.

- *Context* indicates the who, what, why, where, how aspects associated with the object's creation and is *extrinsic* to an information object. The context facilitates the authentication of resources.
- *Structure* relates to the formal set of associations within or among individual information objects and can be *intrinsic* or *extrinsic*. The more highly structured an information object is, the more that structure can be exploited for searching, manipulation, and interrelating with other information objects.

Metadata can be an information resource in its own right. Metadata can describe Metadata, that is, Metadata itself may have attributes such as ownership and an expiry date, and so there is meta-Metadata. For example, a review of a film (which on one level is a piece of Metadata related to the film) is, on another level, a literary work with its own author and its own intellectual property constraints [1, 8, 11].

In an environment where a user can gain unmediated access to information objects over a network, Metadata [2]:

- certifies the authenticity and degree of completeness of the content;
- establishes and documents the context of the content;
- identifies and exploits the structural relationships that exist between and within information objects;
- provides a range of intellectual access points for an increasingly diverse range of users; and
- provides some of the information an information professional might provide in a physical reference or research setting.

There is more to Metadata than description, however. Repositories also create Metadata relating to the administration, accessing, preservation, versionning, and use of collections. Metadata not only identifies and describes an information object; it also documents how that object behaves, its function and use, its relationship to other information objects, and how it should be managed [2].

2.3. Why is Metadata important?

Metadata relieves potential users of resources to have full advance knowledge of their existence or characteristics. Metadata is a systematic method for describing resources and thereby improving access to them (i.e. it supports the creation of a data inventory). If a resource is worth making available, then it is worth describing it with Metadata, so as to maximize the ability to locate it [9].

However, as we can see from the previous definitions, Metadata consists of complex constructs that can be expensive and time-consuming to create and maintain and it may be initially difficult to read. On the other hand, current networked digital information systems have provided professionals many new opportunities while at the same time confront them to new challenges. To exploit these events, professionals need judiciously crafted Metadata that conforms to international standards which provide [2, 9, 12]:

- *Increased accessibility*: Effectiveness of searching can be significantly enhanced through the existence of rich, consistent Metadata. Metadata can also make it possible to search across multiple collections or to create virtual collections from materials that are distributed across several repositories if the descriptive Metadata are the same or can be mapped across each site [2, 10].
- *Data Investment protection*: Mitigates effect of staff turnover and individual memory loss. Sets the stage for data re-use and update and provides documentation of data sources and quality. As personnel change in an organization, institutional knowledge leaves the organization. Undocumented data can lose their value. Subsequent workers may have little understanding of the contents and uses for a digital data base and may find they can't trust results generated from these data. Also, lack of knowledge about other organizations' data can lead to duplication of effort [12, 2].
- *Retention of context*: Many repositories do not simply hold objects. They maintain collections of objects that have complex interrelationships among each other and associations with people, places, movements, and events. Metadata plays a critical role in documenting and maintaining those relationships, as well as in indicating the authenticity, structural and procedural integrity, and degree of completeness of information objects [2].
- *Expanding use*: Many digital information systems make it easier to disseminate digital versions of objects to users around the globe who, for reasons of geography, economics, or other barriers, might otherwise never have had an opportunity to view them. However, this new communities of users may have significantly different needs to those of the traditional users. Metadata can document these changes in uses of systems and content, and that information can in turn feed back into systems development decisions. Well-structured Metadata can also facilitate an almost infinite number of ways to search for information, present results, and even manipulate information objects without compromising the integrity of those information objects [2].
- *Multi-versioning*: The existence of objects in digital form has heightened interest in the ability to create multiple and variant versions of those objects. In either case, there must be Metadata to link the multiple versions and capture what is the same and what is different about each version [2].
- *Legal issues*: Metadata allows repositories to track the many layers of rights and reproduction information that exist for information objects and their multiple versions. Metadata also documents other legal or donor requirements that have been imposed on objects - for example, privacy concerns or proprietary interests. Can prevent data from being inappropriately used or provides protection if data is inappropriately used [2, 12].
- *Preservation*: For the digital information objects to have a chance of surviving migrations through successive generations of computer hardware and software, or removal to entirely new delivery systems, they will need to have Metadata that

enables them to exist independently of the system that is currently being used to store and retrieve them. Technical, descriptive, and preservation Metadata that documents how a digital information object was created and maintained, how it behaves, and how it relates to other information objects will all be essential. For the information objects to remain accessible and intelligible over time, it will also be essential to preserve and migrate this Metadata [2].

- Metadata is also a key component of *data lineage*. It provides basic information about the source and derivation of a data set [5].
- *System improvement and economics*: Benchmark technical data, much of which can be collected automatically by a computer through Metadata, is necessary to evaluate and refine systems in order to make them more effective and efficient from a technical and economic standpoint. The data can also be used in planning for new systems [2].
- *Evidence of prudent data stewardship*: an organization that takes the time to create and maintain quality Metadata will also mostly likely take the time to develop good quality, clean data [12].
- *Help to users to understand data*: Having Metadata available insures that potential data users can make an informed decision about the appropriate use of a dataset. Metadata provides consistency in terminology focuses on key elements of data and facilitates data transfer and interpretation by new users [5, 12].
- *Reduced workload associated with questions about data*: users don't have to keep asking producers questions [12].
- *Cuts overall costs*: allows automation of tools which ease overall burden and cost of data population and maintenance [12].
- *Help to publicize and support the data you or your organization have produced* [6].

After all the previously benefit we can say that benefits exceed associated costs and that [12]:

- Costs associated in creating and maintaining Metadata are identifiable, manageable and short term.
- Most cost-effective to generate Metadata as integrated step of data creation.
- Costs vary with complexity, level of detail, and age of data set.
- Benefits associated with having and using Metadata are identifiable, immediate, and increase over time.

All of the above benefits can be better exploited if we use Metadata that conforms to **standards**. A Metadata standard is a common set of terms and definitions that describe data. Implementing Metadata standards increases the value of data by facilitating data sharing through time and space [2].

2.4. What types of Metadata are there?

Due to the broad conception of Metadata as we saw previously, it is helpful to break it down into distinct categories that reflect key aspects of Metadata functionality to understand this conception better. Some common Metadata categories and examples of common functions that each might perform in a digital information system are [2]:

- **Administrative**: Metadata used in managing and administering information resources. This type of Metadata can be used for:
  - Acquisition information.
  - Rights and reproduction tracking.
  - Documentation of legal access requirements.
  - Location information.
  - Selection criteria for digitization.
  - Version control and differentiation between similar information objects.
  - Audit trails created by recordkeeping systems.
- **Descriptive:** Metadata used to describe or identify information resources. This type of Metadata can be used for:
  - Cataloging records.
  - Finding aids.
  - Specialized indexes.
  - Hyperlinked relationships between resources.
  - Annotations by users.
  - Metadata for recordkeeping systems generated by records creators.
- **Preservation:** Metadata related to the preservation management of information resources. This type of Metadata can be used for:
  - Documentation of physical condition of resources.
  - Documentation of actions taken to preserve physical and digital versions of resources, e.g., data refreshing and migration.
- **Technical:** Metadata related to how a system functions or Metadata behave. This type of Metadata can be used for:
  - Hardware and software documentation.
  - Digitization information, e.g., formats, compression ratios, scaling routines.
  - Tracking of system response times.
  - Authentication and security data, e.g., encryption keys, passwords.

- **Use:** Metadata related to the level and type of use of information resources. This type of Metadata can be used for:
  - Exhibit records.
  - Use and user tracking.
  - Content re-use and multi-versioning information.

2.5. Who and How does Metadata create?

Metadata creation is typically considered to be an obligation of the data producer or data provider. In fact, anyone who creates or modifies data should know how to create Metadata. This means that they should know what the basic content of almost all Metadata standards is. Creating correct Metadata is like library cataloguing, except the creator needs to know more of the information behind the data in order to properly document them. It is important to ensure that there is good communication between the Metadata producer and the data producer; the former will have to ask questions of the latter [6, 4, 5, 13].

The information needed to create Metadata is often readily available when the data are collected. Data producers and users cannot afford to be without documented data. That is, the initial expense of documenting data clearly outweighs the potential costs of duplicated or redundant data generation [2].

First is important to understand both the data you are trying to describe and the standard itself. Then you need to decide about how you will encode the information. You then use some tool to enter information so that the Metadata conform to the standard [5].

The more easily the Metadata can be created and collected at point of creation of a resource or at point of publication, the more efficient the process and the more likely it is to take place. There are many such tools available and the number continues to grow. Such tools can be standalone or part of a package of software, usually with a backend database or repository to store and retrieve the Metadata records [9].

However, as the number of such tools continues to grow, Metadata creation and management have become a very complex mix of manual and automatic processes and layers created by many different functions and individuals at different points in the life of an information object. This means that Metadata is created, modified, and sometimes even disposed of at many points during the life of a resource. As they move through each phase, the objects acquire layers of Metadata that can be associated with the objects in several ways [2].

The phases through which information objects typically move during their life in a digital environment are [2]:

- *Creation and multi-versioning*: Objects enter a digital information system by being created digitally or by being converted into digital format. Multiple versions of the same object may be created for preservation, research, dissemination, or even product development purposes. Some administrative and descriptive Metadata may be included by the creator.

- *Organization*: Objects are automatically or manually organized into the structure of the digital information system and additional Metadata for those objects may be created through registration, cataloging, and indexing processes.
- *Searching and retrieval*: Stored and distributed objects are subject to search and retrieval by users. The computer system creates Metadata that track retrieval algorithms, user transactions, and system effectiveness in storage and retrieval.
- *Utilization*: Retrieved objects are utilized, reproduced, and modified. Metadata related to user annotations, rights tracking, and version control may be created.
- *Preservation and disposition*: Information objects undergo processes such as refreshing, migration, and integrity checking to ensure their continued availability. Information objects that are inactive or no longer necessary may be discarded. Metadata may document both preservation and disposition activities.

At each phase of the object's life, the creation of Metadata may involve several steps. One approach of Metadata development organizes it in the following stages: creation, verification and (optionally) distribution [5]:

Metadata Creation: Novice and expert Metadata developers must develop strategies and implement techniques to Metadata creation [5].

Metadata Verification: What this really means is once a record has been completed, the information it contains must be reviewed, or parsed. It is also important to review the content of the Metadata, verifying that the information describes the subject data completely and correctly [5, 6].

Metadata Distribution: When a Metadata record has passed the parsing routine, it is ready for distribution. There are many ways to distribute Metadata, the internet being the most common [5].

A more general approach adopted successfully by many organizations in their development and management of Metadata includes [14, 2]:

- **Adopt an operational plan**: two to three pages that form a planning perspective address a typical approach to training staff and management, data set inventory and prioritization, adopting procedures and best practices, etc.
- **Establish Procedures**:
    - Identifying which Metadata schema or schemas should be applied in order to best meet the needs of the information creator, repository and users.
    - Deciding which aspects of Metadata are essential for what you wish to achieve (what level of detail to collect Metadata at), and how granular they need each type of Metadata to be. There will likely always be important tradeoffs between the costs of developing and managing Metadata to meet current needs, and creating sufficient Metadata that can be capitalized

upon for future, often unanticipated uses; ensuring that the Metadata schemas being applied are the most current versions.

- **Train people**: identify needed skills for all associated personnel and set aside time for training.
- **Assess technology infrastructure**: identify collection tools compatible with existing IT structure and that supports procedures for creation, maintenance and distribution of Metadata. Also important to consider if tool is adaptable/upgradeable.
- **Identify suitable Metadata repository**: where Metadata resides should be based on consideration of internal access, external access and maintenance.

2.6. Where is Metadata stored?

Metadata can be contained within the same envelope as the information object - for example, in the form of header information for an image file, (e.g. Dublin Core), or through some form of bundling, (for example, with the Universal Preservation Format (UPF)). Metadata can also be attached to the information object through bi-directional pointers or hyperlinks while the relationships between Metadata and information objects, and between different aspects of Metadata, can be documented by registering them with a Metadata registry (for more information see [15]). However, in any instance where it is critical that Metadata and content coexist, then it is recommended that the Metadata become an integral part of the information object and not be stored elsewhere [2].

Metadata may be created, stored, and used in a variety of formats. The most basic form of Metadata is an ASCII text document. An ASCII document is easy to transfer to other users independent of the hardware/software platform they use. Another common format is Hypertext Markup Language (HTML). HTML provides an attractive way to view Metadata using a browser [4].

For example, Metadata for resources that can be accessed through the World Wide Web, may be deployed in a number of ways [9]:

- Embedding the Metadata in the Web page by the creator or their agent using META tags in the HTML coding of the page.
- As a separate HTML document linked to the resource it describes.
- In a database linked to the resource. The records may either have been directly created within the database or extracted from another source, such as Web pages.

As systems designers increasingly respond to the need to incorporate and manage Metadata in information systems and to address how to move them forward through time, many additional mechanisms for associating Metadata with information objects are likely to become available. Metadata registries and schema recordkeeping systems are also more likely to develop as it becomes increasingly necessary to document schema evolution and to alert implementers to version changes [2].

2.7. Metadata on the World Wide Web

Recently there has been a focus on Metadata in relation to those information resources which can be accessed through the World Wide Web. In this context Metadata is "data describing web resources". However, whether in the traditional context or in the Internet context, one key purpose of Metadata is to facilitate and improve the retrieval of information. But one of the major problems of the World Wide Web today is that it is really hard to automate any task which has to be performed on the web [1, 11].

So far, the web is mainly built as a forum for human interaction; because most web documents are written for human consumption, the only available form of searching on the web (for example) is to simply match words or sentences contained in documents. This means that typing a few keywords in a web search service and receiving a few thousand hits is not necessarily very useful. This is what information scientists' term "high recall" and "low precision". The high recall refers to this (frustrating) experience of using an Internet search engine and receiving thousands of hits. It is popularly known as information overload. The low precision refers to not being able to locate the most useful documents [11, 9].

The Working Group on Government Information Navigation[4] outlined the problems with Internet search engines. These problems and how Metadata can be used to solve them are [9, 1]:

- The search engines can return a lot of irrelevant information because they have no means (or very few means) of distinguishing between important and incidental words in document texts. If we could target our searches onto words which are used as significant terms, we could achieve an enormous improvement in precision. Metadata can be used to achieve this by identifying just the major concepts of the information resource.
- Precision can also be improved if we could target searches onto words or phrases that identify their correct role. Metadata can be used to achieve this by identifying the different characteristics of the information resource: the author, subject, title, publisher and so on.
- There is also a need to improve search recall, that is, to retrieve information resources that would otherwise be missed. For example, relevant information can be missed because sites contain types of resource in addition to HTML text (e.g. images, databases, PDF documents). Metadata can support retrieval of these resources by identifying them, thus ensuring they are not missed by harvesting engines.
- Recall can also be improved due to other factors. For example, it is known that most harvesting engines do not index every page on a site, but often only the top two or three hierarchical levels. Thus, these engines miss significant documents which, on larger and more complex sites, may be located in lower levels of the

[4] http://www.nla.gov.au/lis/esd4.html

hierarchy. A better harvesting process would gather Metadata from a repository created locally from a complete coverage of the local site. The data in this repository could then be gathered regularly by the harvesting engine.

- Search engines, especially the more comprehensive ones, may index sites on an infrequent basis and may therefore not contain the most current data.

2.8. Metadata Standards

There are a variety of Metadata standards. On one hand, many highly detailed Metadata standards are now emerging to describe and provide access to a particular type of information resources. This type of standards (such as the Encoded Archival Description (EAD)[5] , the Australian Recordkeeping Metadata Schema (RKMS)[6], and the FGDC standard developed by the US Federal Geographic Data Committee[7]) attempt to articulate the communities' mission-specific differences as well as to facilitate mapping between common data elements. By contrast, the Dublin Core Metadata Element Set (DC)[8] identifies a small, simple set of Metadata elements that can be used by any community to describe and search across a wide variety of information resources on the World Wide Web. Such Metadata standards are necessary in order to ensure that different kinds of descriptive Metadata are able to interoperate with each other [1, 2].

Another standard is the Resource Description Framework (RDF)[9] recommended by the W3C. One could say that the history of Metadata at W3C begins with Platform for Internet Content Selection (PICS)[10]. The development of RDF as a general Metadata framework and in a way as a general knowledge representation mechanism for the web was heavily inspired by PICS [11].

RDF aims at facilitating the creation and exchange of Metadata as any other Web data, since Metadata can be widely used in order to fully exploit information resources (e.g., sites, documents, data, images, etc.) available on the WWW (content and/or structure). It is a foundation for processing Metadata; it provides interoperability between applications that exchange machine-understandable information on the Web and emphasizes facilities to enable automated processing of Web resources [11, 16].

The Warwick Framework [1]

Before examining one particular Metadata standard it is useful to observe that an architecture has been developed to handle a variety of Metadata sets. This architecture is known as the Warwick Framework[11].

[5] http://www.loc.gov/ead/

[6] http://www.sims.monash.edu.au/research/rcrg/research/spirt/onresearch.html

[7] http://www.fgdc.gov/Metadata/Metadata.html

[8] http://dublincore.org/

[9] http://www.w3.org/RDF/

[10] http://www.w3.org/PICS/

[11] http://www.dlib.org/dlib/july96/lagoze/07lagoze.html

The Warwick Framework uses the “container-package” model. It provides a conceptual framework only: the method of handling the containers and packages must be provided by any particular application of the model. A container is simply any mechanism for aggregating packages. A package may be of three types:

- a “primitive” package, which contains one or more pieces of Metadata; each primitive package has a type (for example, a MARC package, a Dublin Core package, an FGDC package).
- an “indirect” package, which refers to another information resource, for example through a link to its URL.
- a “container” package: a package may itself be a container, and there is no limit to the degree of nesting involved here.

The Warwick Framework seems extraordinarily simple, but in fact it provides a powerful model for handling Metadata. Its advantages are that it is modular (Metadata is assembled in packages), extensible (there is no limit to the types of package which may be assembled in a container), distributed (through the use of indirect packages) and recursive (since a package may also be a container).

2.9. Conclusions

The necessity of Metadata is more evident each day, especially when the large amounts of information available currently (i.e. the World Wide Web) makes it difficult to search and retrieve information. Metadata is the information necessary to help to create order in the information, providing description, classification and organization.

However, Metadata relates to more than the description of an object. While the term is most familiar with description or cataloging, Metadata can also indicate the context, management, processing, preservation and use of the resources being described. Besides, one information object’s Metadata can simultaneously be another information object’s data [2].

Although Metadata has become more popular with digital era, Metadata does not have to be digital. Cultural heritage and information professionals have been creating Metadata for as long as they have been managing collections. Increasingly, such Metadata are being incorporated into digital information systems [2].

The use of these digital information systems, however, can help in the creation of Metadata because Metadata can come from a variety of sources. It can be supplied by a human (a creator, information professional, or user), created automatically by a computer, or inferred through a relationship to another resource such as a hyperlink. But, the implementation of Metadata does not include just its creation. Metadata continue to accrue during the life of an information object or system. Metadata is created, modified, and sometimes even disposed of at many points during the life of a resource [2].

Metadata is like interest - it accrues over time. Carefully designed Metadata, results in the best information management in the short and long-term. If thorough, consistent Metadata has been created, it is possible to conceive of it being used in an almost infinite number of new ways to meet the needs of non-traditional users, for multi-versioning, and for data mining [2].

Future of Metadata on the Web

The solutions presented in the section 2.7 for the search problem and for the general issue of letting automated "agents" roam the web performing useful tasks could elevate the status of the web from *machine-readable* to something we might call *machine-understandable*. In the future, when the Metadata languages and engines are more developed, it should also form a strong basis for a web of machine understandable information about anything: about the people, things, concepts and ideas [11, 8].

The phrase machine-understandable is key. We are talking about information which software agents can use in order to make life easier for us, ensure we obey our principles, the law, check that we can trust what we are doing, and make everything work more smoothly and rapidly [8].

So, once the web has been sufficiently populated with rich Metadata, what can we expect? First, searching on the web will become easier as search engines have more information available, and thus searching can be more focused. Doors will also be opened for automated software agents to roam the web, looking for information for us or transacting business on our behalf. The web of today, the vast unstructured mass of information, may in the future be transformed into something more manageable and thus something far more useful [11].

3. Metadata Examples

In this section we present several metadata example widely used, like:

- Dublin Core, which is used for describing networked resources [23].
- LOM metadata, which is a set of elements for describing learning resources.
- FOAF, which can be used for managing on-line communities [24, 25].
- BibTeX, which is a metadata format for modelling bibliography entries.
- Content Standard for Digital Geospatial Metadata, which is used for providing a common terminology for the documentation of digital geospatial data.
- Education Network Australia Metadata Standard, whose main aim is to support interoperability across all sectors of education and training in Australia in the area of online resource discovery and management.
- MusicBrainz Metadata Initiative, which is a model for describing digital audio and video tracks.
- DogmaModeler ontology metadata, which is being used in the DogmaModeler ontology engineering tool [20, 21].

3.1. Dublin Core

by UPM.

Description:

The Dublin Core (DC) metadata standard[12] is a simple yet effective element set for describing a wide range of networked resources. It was developed during 1995 and 1996 as a response to the need to improve retrieval of information resources, especially on the World Wide Web. The DC standard includes two levels: Simple and Qualified. Simple DC comprises fifteen elements; Qualified DC includes an additional element, Audience, as well as a group of element refinements (or qualifiers) that refine the semantics of the elements in ways that may be useful in resource discovery. The semantics of DC have been established by an international, cross-disciplinary group of professionals from librarianship, computer science, text encoding, the museum community, and other related fields of scholarship and practice.

Another way to look at DC is as a “small language for making a particular class of statements about resources”. In this language, there are two classes of terms: elements (nouns) and qualifiers (adjectives), which can be arranged into a simple pattern of statements. The resources themselves are the implied subjects in this language. (For additional discussion of DC Grammar, see[13]) In the diverse world of the Internet, DC can be seen as a “metadata pidgin for digital tourists”: easily grasped, but not necessarily up

[12] http://dublincore.org/documents/dces/

[13] http://dublincore.org/usage/documents/principles/

to the task of expressing complex relationships or concepts. It was intended to be descriptive, rather than evaluative. It does not provide for rating systems.

Each element in the DC basic element set is optional and may be repeated. Most elements also have a limited set of qualifiers or refinements, attributes that may be used to further refine (not extend) the meaning of the element. The DC Metadata Initiative (DCMI)[14] has established standard ways to refine elements and encourage the use of encoding and vocabulary schemes. The full set of elements and elements refinements[15] conforming to DCMI "best practice" is available, with a formal registry in process.

Although the DC was originally developed with an eye to describing document-like objects (because traditional text resources are fairly well understood), DC metadata can be applied to other resources as well. Its suitability for use with particular non-document resources will depend to some extent on how closely their metadata resembles typical document metadata and also what purpose the metadata is intended to serve.

The DC standard was deliberately limited to a small set of elements which would have applicability over a wide range of types of information resource.

Set of Elements:

Table 1 includes the simple DC elements.

Element Name	Label	Type[16]	Definition	Comment
Title	Title	C	A name given to the resource.	Typically, Title will be a name by which the resource is formally known.
Creator	Creator	IP	An entity primarily responsible for making the content of the resource.	Examples of Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity.
Subject	Subject and Keywords	C	A topic of the content of the resource.	Typically, Subject will be expressed as keywords, key phrases or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.
Description	Description	C	An account of the content of the resource.	Examples of Description include, but are not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.
Publisher	Publisher	IP	An entity responsible for	Examples of Publisher include a person, an organization, or a service.

[14] http://dublincore.org/

[15] http://dublincore.org/documents/dcmi-terms/

[16] Type indicates if the element relates more to the content (C), Intellectual Property (IP), Instantiation (I) of the item

			making the resource available	Typically, the name of a Publisher should be used to indicate the entity.
Contributor	Contributor	IP	An entity responsible for making contributions to the content of the resource.	Examples of Contributor include a person, an organization, or a service. Typically, the name of a Contributor should be used to indicate the entity.
Date	Date	I	A date of an event in the lifecycle of the resource.	Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 and includes (among others) dates of the form YYYY-MM-DD.
Type	Resource Type	C	The nature or genre of the content of the resource.	Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type Vocabulary). To describe the physical or digital manifestation of the resource, use the FORMAT element.
Format	Format	I	The physical or digital manifestation of the resource.	Typically, Format may include the media-type or dimensions of the resource. Format may be used to identify the software, hardware, or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types defining computer media formats).
Identifier	Resource Identifier	I	An unambiguous reference to the resource within a given context.	Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Formal identification systems include but are not limited to the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).
Source	Source	C	A Reference to a resource from which the present resource is derived.	The present resource may be derived from the Source resource in whole or in part. Recommended best practice is to identify the referenced resource by means of a string or number

				conforming to a formal identification system.
Language	Language	I	A language of the intellectual content of the resource.	Recommended best practice is to use RFC 3066 which, in conjunction with ISO639, defines two- and three-letter primary language tags with optional subtags. Examples include "en" or "eng" for English, "akk" for Akkadian", and "en-GB" for English used in the United Kingdom.
Relation	Relation	C	A reference to a related resource.	Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.
Coverage	Coverage	C	The extent or scope of the content of the resource.	Typically, Coverage will include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [TGN]) and to use, where appropriate, named places or time periods in preference to numeric identifiers such as sets of coordinates or date ranges.
Rights	Rights Management	IP	Information about rights held in and over the resource.	Typically, Rights will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the Rights element is absent, no assumptions may be made about any rights held in or over the resource.

Table 1. Dublin Core Metadata Element Set

References:

- Dublin Core Projects: http://dublincore.org/projects/
- Using Dublin Core: http://dublincore.org/documents/usageguide/

3.2. LOM Metadata (metadata for e-learning objects)

by FU Berlin.

Description:

Learning Object Metadata (LOM) is the result of an initiative to develop open market-based standards for online learning, including specifications of learning content meta-data. LOM is made up of a conceptual data schema as well as a binding of the schema to XML. Of interest to Semantic Web practitioners there is also a draft RDF binding in development[17].

The IEEE LOM standard defines a set of meta-data elements for describing learning resources. This includes element names, definitions, data types and field lengths. The conceptual data schema is organized hierarchically. A sample view of this hierarchy is shown in figure 1.

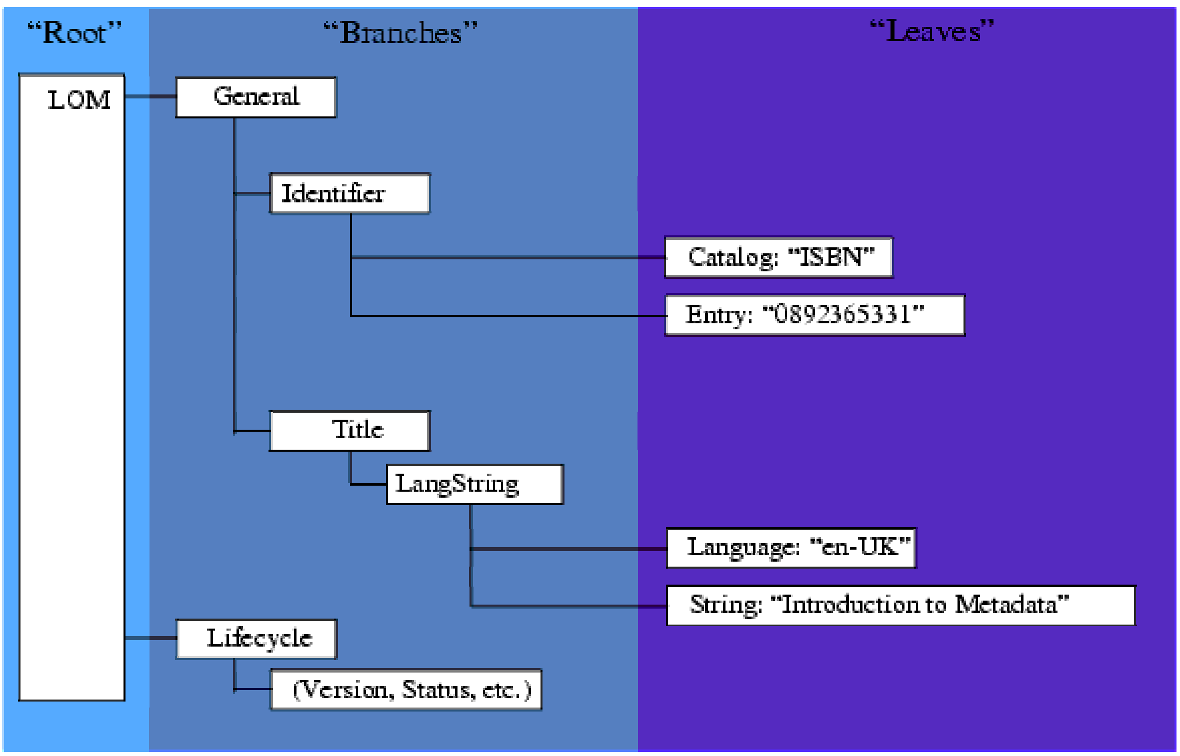

Figure 1. LOM standard conceptual data schema

Set of Elements:

LOM elements are divided into 9 top level categories: General, Life Cycle, Meta-metadata, Technical, Educational, Rights, Relation, Annotation and Classification. These "branches" (in the hierarchy of the conceptual schema) contain other elements, some of

[17] http://kmr.nada.kth.se/el/ims/metadata.html

which are leaves and some are sub-branches which lead to leaves. Figure 2 shows a graphical illustration of the LOM elements.

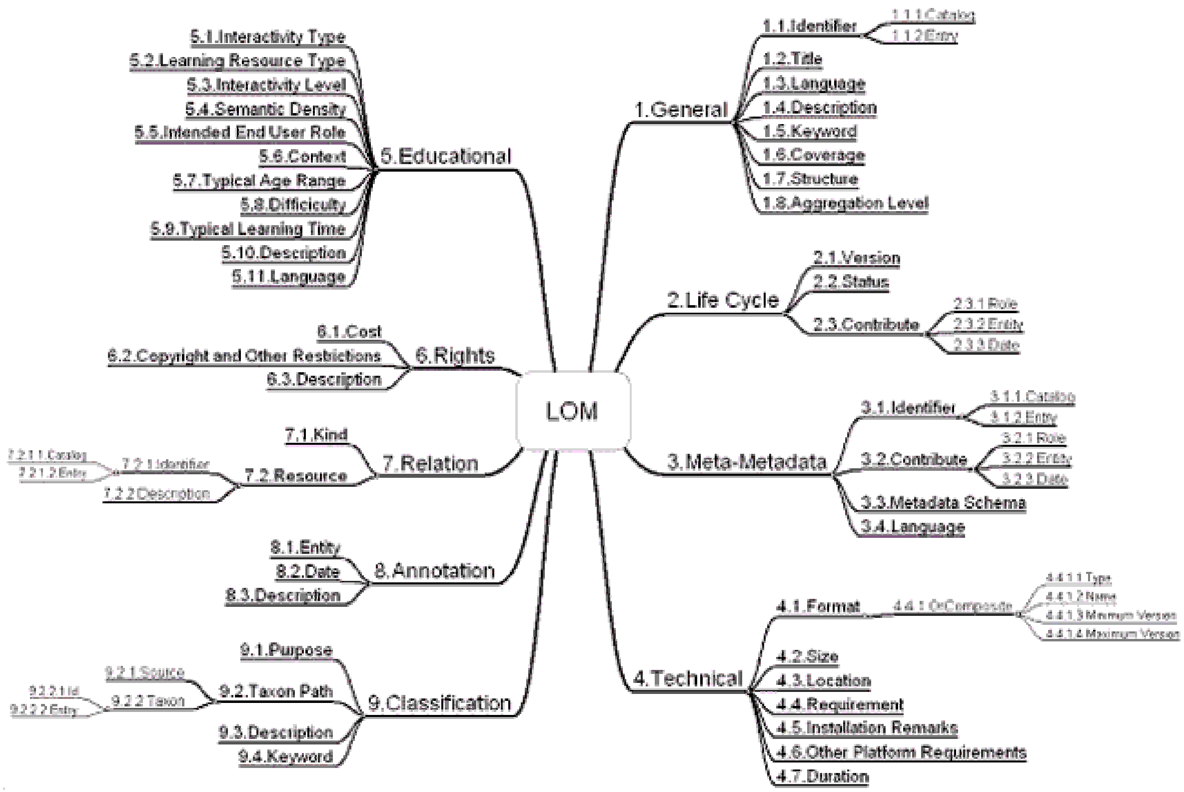

Figure 2. Graphical view of the LOM elements

The elements of LOM, with their proposed RDF binding (which uses Dublin Core and ten LOM specific namespaces) and data type, are shown in table 2. LOM Restricted Vocabulary datatypes represent controlled vocabularies which are used as permissible values for the element. Formal ontologies have not been specified for these vocabularies.

LOM Element	RDF Encoding	Data Type
General		
1.1 Identifier	<dc:identifier>	#PCDATA
1.2 Title	<dc:title>	#PCDATA
1.3 Language	<dc:language>	RFC1766
1.4 Description	<dc:description>	#PCDATA
1.5 Keyword	<dc:subject>	#PCDATA
1.6 Coverage	<dc:coverage>	#PCDATA
1.7 Structure	<lom-gen:structure>	LOM restricted vocabulary
1.8 Aggregation Level	<lom-gen:aggregationLevel>	LOM restricted vocabulary
Lifecycle		
2.1 Version	<lom-life:version>	#PCDATA
2.2 Status	<lom-life:status>	#PCDATA
2.3 Contribute	<dc:publisher> <dc:editor>	VCard or #PCDATA VCard or #PCDATA

	<dc:creator>	VCard or #PCDATA
Metametadata		
3.1 Identifier	See 1.1	
3.2 Contribute	See 2.3	
3.3 Metadata Scheme	<meta:metadataScheme>	Resource
3.4 Language	See 1.3	
Technical		
4.1 Format	<dc:format>	dcterms:IMT
4.2 Size	<lom-tech:ByteSize>	#PCDATA
4.3 Location	<lom-tech:location>	#PCDATA URL
4.4 Requirements	<lom-tech:operatingSystem> <lom-tech:browser>	#PCDATA #PCDATA
4.5 Installation Remarks	<lom-installationRemarks>	#PCDATA
4.6 Other Platform Requirements	<lom-other: PlatformRequirements>	#PCDATA
4.7 Duration	<dcterms:extent>	ISO8601
Educational		
5.1 Interactivity Type	<lom-edu:interactivityType>	LOM restricted vocabulary
5.2 Learning Resource Type	<lom-edu:type>	LOM restricted vocabulary
5.3 Interactivity Level	<lom-edu:interactivityLevel>	LOM restricted vocabulary
5.4 Semantic Density	<lom-edu:semanticDensity>	LOM restricted vocabulary
5.5 Intended End-User Role	<lom-edu: intendedEndUserRole>	LOM restricted vocabulary
5.6 Context	<lom-edu:context>	LOM restricted vocabulary
5.7 Typical Age Range	<dcterms:audience>	
5.8 Difficulty	<lom-edu:difficulty>	LOM restricted vocabulary
5.9 Typical Learning Time	<lom-edu:typicalLearningTime>	ISO8601
5.9 Description	<lom-edu:description>	#PCDATA
5.10 Language	<lom-edu:language>	RFC1766
Rights		
6.1 Cost	<lom-rights:cost>	LOM restricted vocabulary
6.2 Copyright and Other Restrictions	<lom-rights: copyrightAndOtherRestrictions>	LOM restricted vocabulary
6.3 Description	<dc:rights>	#PCDATA
Relation		
7 Relation	Use DCTerms	#PCDATA
Annotation Category		
8 Annotation	<lom-ann:annotation>	#PCDATA
Classification		
9 Classification	<lom-cls:classification> <lom-cls:prerequisite> <lom-cls:educationalObjective> <lom-cls:accessibilityRestrictions <lom-cls:educationalLevel> <lom-cls:skillLevel> <lom-cls:securityLevel> <lom-cls:competency.>	LOM restricted vocabulary LOM restricted vocabulary LOM restricted vocabulary LOM restricted vocabulary LOM restricted vocabulary LOM restricted vocabulary LOM restricted vocabulary LOM restricted vocabulary

Table 2. LOM elements

References:

- 1484.12.1 IEEE Standard for Learning Object Metadata (LOM) conceptual data schema can be found at: http://standards.ieee.org/
- Information relating to the IEEE 1484.12.3 Draft Standard for eXtensible Markup Language (XML) Schema Definition Language Binding for Learning Object Metadata can be found at: http://ltsc.ieee.org/wg12/par1484-12-3.html
- Further information about the IEEE Learning Object Metadata Working Group is available at: http://ltsc.ieee.org/wg12/index.html
- The IMS Learning Resource Meta-data specifications versions 1.0 - 1.3 can be found at: http://www.imsglobal.org/metadata/

Acknowledgements:

- Diagrams from IMS Meta-data Best Practice Guide for IEEE 1484.12.1-2002 Standard for Learning Object Metadata Revision: 20 May 2004. http://www.imsglobal.org/metadata/mdv1p3pd/imsmd_bestv1p3pd.html#1535098. Copyright © 2004 IMS Global Learning Consortium, Inc.
- Element list with RDF binding from http://www.downes.ca/xml/RSS_LOM.htm: "RDF Site Summary 1.0 Modules: Learning Object Metadata". Copyright © 2003 by the Authors: Stephen Downes, NRC E-Learning Group, Moncton.
  Permission to use, copy, modify and distribute the RDF Site Summary 1.0 Specification and its accompanying documentation for any purpose and without fee is hereby granted in perpetuity, provided that the above copyright notice and this paragraph appear in all copies. The copyright holders make no representation about the suitability of the specification for any purpose. It is provided "as is" without expressed or implied warranty.

3.3. FOAF

by UPM.

Description:

FOAF[18], or "Friend of a Friend", provides a way to create machine-readable Web homepages for people (their interests, relationships and activities), groups, companies and other kinds of thing.

To achieve this, the FOAF project use the "FOAF vocabulary" to provide a collection of basic terms that can be used in these Web pages. At the heart of the FOAF project is a set of definitions designed to serve as a dictionary of terms that can be used to express claims about the world. The initial focus of FOAF has been on the description of people, since people are the things that link together most of the other kinds of things we describe in the Web: they make documents, attend meetings, are depicted in photos, and so on.

[18] http://www.foaf-project.org/

Founded by Dan Brickley and Libby Miller, FOAF is an open community-lead initiative which is tackling head-on the wider Semantic Web[19] goal of creating a machine processable web of data.

Technically, FOAF is an RDF[20]/XML Semantic Web vocabulary. FOAF documents are most commonly used as a way of representing information about people in a way that is easily processed, merged and aggregated.

FOAF provides conventions for saying the sorts of things that you might say in your homepage (‘My name is...’, ‘I work for ...’, ‘I’m interested in ...’, ‘I live near ...’, ‘I’m pictured in these photos...’, ‘I write in this weblog...’, etc.), but in a way that is easy for computers to process. Since computers are pretty dumb, and can’t read human languages, we provide simplistic FOAF descriptions, to help them answer questions such as ‘Show me pictures of Weblog authors interested in ... who live near here’, ‘Show me recent articles written by people at this meeting’, ‘Is this person vegetarian?’, etc.

FOAF is a Semantic Web project, which is an effort to make the Web easier for machines to help us navigate.

Set of Elements:

FOAF files are just text documents (well, Unicode documents). They are written in XML syntax, and adopt the conventions of the Resource Description Framework (RDF), so that Semantic Web applications can use those terms in a variety of RDF-compatible document formats and applications. In addition, the FOAF vocabulary defines some useful constructs that can appear in FOAF files, alongside other RDF vocabularies defined elsewhere. For example, FOAF defines categories (*classes*) such as foaf:Person, foaf:Document, foaf:Image, alongside some handy *properties* of those things, such as foaf:name, foaf:mbox (ie. an internet mailbox), foaf:homepage etc., as well as some useful kinds of *relationship* that hold between members of these categories.

Table 3 shows the FOAF classes.

Class	Status	Definition	Comment
foaf:Agent	unstable	An agent (e.g. person, group, software or physical artifact).	The foaf:Agent class is the class of agents; things that do stuff. A well known sub-class is foaf:Person, representing people. Other kinds of agents include foaf:Organization and foaf:Group.
foaf:Document	testing	A document.	The foaf:Document class represents those things which are, broadly conceived, ‘documents’. The foaf:Image class is a sub-class of foaf:Document, since all images are documents.
foaf:Group	unstable	A class of	The foaf:Group class represents a

[19] http://www.w3.org/2001/sw/

[20] http://www.w3.org/2001/sw/

		Agents.	collection of individual agents (and may itself play the role of a foaf:Agent, ie. something that can perform actions). This concept is intentionally quite broad, covering informal and ad-hoc groups, long-lived communities, organizational groups within a workplace, etc.
foaf:Image	testing	An image.	The class foaf:Image is a sub-class of foaf:Document corresponding to those documents which are images. Digital images (such as JPEG, PNG, GIF bitmaps, SVG diagrams etc.) are examples of foaf:Image.
foaf:OnlineAccount	unstable	An online account.	A foaf:OnlineAccount represents the provision of some form of online service, by some party to some foaf:Agent. Sub-classes include foaf:OnlineChatAccount, foaf:OnlineEcommerceAccount and foaf:OnlineGamingAccount.
foaf:OnlineChatAccount	unstable	An online chat account.	A foaf:OnlineChatAccount is a foaf:OnlineAccount devoted to chat / instant messaging.
foaf:OnlineEcommerceAccount	unstable	An online e-commerce account.	A foaf:OnlineEcommerceAccount is a foaf:OnlineAccount devoted to buying and/or selling of goods, services etc. Examples include Amazon, eBay, PayPal, thinkgeek, etc.
foaf:OnlineGamingAccount	unstable	An online gaming account.	A foaf:OnlineGamingAccount is a foaf:OnlineAccount devoted to online gaming. Examples might include EverQuest, Xbox live, Neverwinter Nights, etc., as well as older text-based systems (MOOs, MUDs and suchlike).
foaf:Organization	unstable	An organization.	The foaf:Organization class represents a kind of foaf:Agent corresponding to social instititutions such as companies, societies etc.
foaf:Person	testing	A person.	The foaf:Person class represents people. Something is a foaf:Person if it is a person. We don’t nitpic about whether they’re alive, dead, real, or imaginary. The foaf:Person class is a sub-class of the foaf:Agent class, since all people are considered ‘agents’ in FOAF.
foaf:PersonalProfileDocument	testing	A personal profile RDF	The foaf:PersonalProfileDocument class represents those things that

		document.	are a foaf:Document, and that use RDF to describe properties of the person who is the foaf:maker of the document.
foaf:Project	unstable	A project (a collective endeavour of some kind).	The foaf:Project class represents the class of things that are 'projects'. These may be formal or informal, collective or individual. It is often useful to indicate the foaf:homepage of a foaf:Project.

Table 3. The FOAF Vocabulary Description

References:

- The FOAF FAQ: http://rdfweb.org/topic/FAQ

3.4. BibTeX

by UKARL.

Description:

BibTeX is a metadata format for modelling bibliography entries used within the LaTeX [26] document preparation system. LaTeX which can be used on most every operating system is a well known system for typesetting documents and is used often in the scientific and academia communities or by commercial publishers. LaTeX itself is based on TeX an initial typesetting system developed by Donald Knuth.

BibTeX is on the one hand a metadata standard and on the other hand a system for managing bibliographic entries obtaining these data from bibliographic databases. The metadata format and program was developed in 1985 by Oren Patashnik and Leslie Lamport (see References).

BibTeX provides metadata attributes (entry types) for nearly every kind of bibliographic entry which has its own set of attributes describing a reference.

Example:

```
@article {Bibliography_Entry.2004-05-19.1152,
      author = {Jens Hartmann and York Sure},
      title = {An Infrastructure for Scalable, Reliable Semantic
      Portals},
      journal = {IEEE Intelligent Systems},
      volume = {19},
      number = {3},
      pages = {58-65},
      month = {MAY},
      year = {2004},
}
```

Set of Elements:

Table 4 shows the standard entry types for BibTeX.

Element Name	Definition
article	An article from a journal or magazine.
book	A book.
boklet	A work that is printed and bound, but without a named publisher or sponsoring institution.
conference	See inproceedings.
inbook	A part of a book, which may be a chapter (or section) and/or a range of pages.
incollection	A part of a book.
inproceedings	An article in a conference proceedings.
manual	Technical documentation.
mastersthesis	A Master's thesis.
misc	Miscellaneous publication.
phdthesis	A PhD thesis.
proceedings	The proceedings of a conference.
techreport	A technical report.
unpublished	A document having an author and title, but not formally published.

Table 4. Standard Entry Types

Each entry type in BibTeX can consist of the fields presented in table 5.

Element Name	Definition
address	The address of the publisher.
author	The name(s) of the author(s).
booktitle	Title of the part of the cited book.
chapter	A chapter number.
crossref	Normally used as reference key, e.g. for a database where additional data is stored.
edition	The edition of a publication.
editor	The name(s) of the editor(s).
howpublished	Description how the publication was published.
institution	The sponsoring institution.
journal	Name of journal.
key	Allows to specify a key for referring this publication. Note, this key is not the key used for “cite”.
month	The month in which the work was published.
note	Notes or comments can be inserted here.
number	The number of a work in series (journal, magazine, etc.).
organization	The organization which organizes a conference or publishes a work.
pages	Page numbers or page rank.
publisher	The name of the publisher.
school	The name of the school.
series	The name of the series (optional).
title	The title of the work.

type	The type of a work / technical report.
volume	The volume number.
year	The year of the work in which the work was published.

Table 5. Standard BibTeX Fields

References:

- ftp://sunsite.unc.edu/pub/packages/TeX/biblio/bibtex/
- BibTeXing, by Oren Patashnik, February 1988, (BibTeX distribution), ftp://sunsite.unc.edu/pub/packages/TeX/biblio/bibtex/distribs/doc/btxdoc.tex

3.5. Content Standard for Digital Geospatial Metadata (CSDGM)

by UPM.

Description:

The Federal Geographic Data Committee (FGDC)[21] initiated work on the first version of the standard in June, 1992, through a forum on geospatial metadata. The first version was approved on June, 1994 and the current version (June 1998)[22] is fully backward compatible with and supersedes the first version.

The objective of the standard is to provide a common set of terminology and definitions for the documentation of digital geospatial data. The standard establishes the names of data elements and compound elements (groups of data elements) to be used for these purposes, the definitions of these compound elements and data elements, and information about the values that are to be provided for the data elements.

The major uses of metadata are:

- to maintain an organization's internal investment in geospatial data
- to provide information about an organization's data holdings to data catalogues, clearinghouses, and brokerages, and
- to provide information needed to process and interpret data to be received through a transfer from an external source.

The information included in the standard was selected based on four roles that metadata play:

- Availability. Data needed to determine the sets of data that exist for a geographic location.
- Fitness for use. Data needed to determine if a set of data meets a specific need.
- Access. Data needed to acquire an identified set of data.

[21] http://www.fgdc.gov/index.html
[22] http://www.fgdc.gov/metadata/contstan.html

- Transfer. Data needed to process and use a set of data.

These roles form a continuum in which a user cascades through a pyramid of choices to determine what data are available, to evaluate the fitness of the data for use, to access the data, and to transfer and process the data. The exact order in which data elements are evaluated, and the relative importance of data elements, will not be the same for all users.

This standard is intended to support the collection and processing of geospatial metadata. It is intended to be useable by all levels of government and the private sector. The standard is not intended to reflect an implementation design.

Set of Elements:

The standard is organized in a hierarchy of data elements and compound elements that define the information content for metadata to document a set of digital geospatial data. The starting point is "metadata" (section 0). The compound element "metadata" is composed of other compound elements representing different concepts about the data set. Each of these compound elements has a numbered section in the standard. In each numbered section, these compound elements are defined by other compound elements and data elements.

Each section begins with the name and definition of the compound element that defines the section. The name and definition are followed by production rules (using Yourdan syntax) that define this compound element in terms of data elements, either directly or by the use of intermediate compound elements.

The basic compound elements of the standard are shown in table 6.

Name	Definition	Short Name	Production Rule
Metadata	data about the content, quality, condition, and other characteristics of data.	metadata	Metadata = Identification_Information + 0{Data_Quality_Information}1 + 0{Spatial_Data_Organization_Information}1 + 0{Spatial_Reference_Information}1 + 0{Entity_and_Attribute_Information}1 + 0{Distribution_Information}n + Metadata_Reference_Information
Identification Information	basic information about the data set.	idinfo	Identification_Information = Citation + Description + Time_Period_of_Content + Status + Spatial_Domain + Keywords + Access_Constraints + Use_Constraints + (Point_of_Contact) + (1{Browse_Graphic}n) + (Data_Set_Credit) + (Security_Information) + (Native_Data_Set_Environment) +

			(1{Cross_Reference}n)
Data Quality Information	a general assessment of the quality of the data set.	dataqual	Data_Quality_Information = 0{Attribute_Accuracy}1 + Logical_Consistency_Report + Completeness_Report + 0{Positional_Accuracy}1 + Lineage + (Cloud_Cover)
Spatial Data Organization Information	the mechanism used to represent spatial information in the data set.	spdoinfo	Spatial_Data_Organization_Information = 0{Indirect_Spatial_Reference}1 + 0{Direct_Spatial_Reference_Method + ([Point_and_Vector_Object_Information | Raster_Object_Information])}1
Spatial Reference Information	the description of the reference frame for, and the means to encode, coordinates in the data set.	spref	Spatial_Reference_Information = 0{Horizontal_Coordinate_System_Definition}1 + 0{Vertical_Coordinate_System_Definition}1
Entity and Attribute Information	details about the information content of the data set, including the entity types, their attributes, and the domains from which attribute values may be assigned.	eainfo	Entity_and_Attribute_Information = [1{Detailed_Description}n | 1{Overview_Description}n | 1{Detailed_Description}n + 1{Overview_Description}n]
Distribution Information	information about the distributor of and options for obtaining the data set.	distinfo	Distribution_Information = Distributor + 0{Resource_Description}1 + Distribution_Liability + 0{Standard_Order_Process}n + 0{Custom_Order_Process}1 + (Technical_Prerequisites) + (Available_Time_Period)
Metadata Reference Information	information on the currentness of the metadata information, and the responsible party.	metainfo	Metadata_Reference_Information = Metadata_Date + (Metadata_Review_Date) + (Metadata_Future_Review_Date) + Metadata_Contact + Metadata_Standard_Name + Metadata_Standard_Version + 0{Metadata_Time_Convention}1 + (Metadata_Access_Constraints) + (Metadata_Use_Constraints) + (Metadata_Security_Information) + 0{Metadata_Extensions}n
Citation Information	the recommended reference to be used for the data set.	citeinfo	Citation_Information = 1{Originator}n + Publication_Date +

			(Publication_Time) + Title + 0{Edition}1 + 0{Geospatial_Data_Presentation_Form}1 + 0{Series_Information}1 + 0{Publication_Information}1 + 0{Other_Citation_Details}1 + (1{Online_Linkage}n) + 0{Larger_Work_Citation}1
Time Period Information	information about the date and time of an event.	timeinfo	Time_Period_Information = [Single_Date/Time | Multiple_Dates/Times | Range_of_Dates/Times]
Contact Information	Identity of, and means to communicate with, person(s) and organization(s) associated with the data set.	cntinfo	Contact_Information = [Contact_Person_Primary | Contact_Organization_Primary] + (Contact_Position) + 1{Contact_Address}n + 1{Contact_Voice_Telephone}n + (1{Contact_TDD/TTY_Telephone}n) + (1{Contact_Facsimile_Telephone}n) + (1{Contact_Electronic_Mail_Address}n) + (Hours_of_Service) + (Contact_Instructions)

Table 6. Basic elements of the CSDGM

References:

- Content Standard for Digital Geospatial Metadata (CSDGM). http://www.fgdc.gov/metadata/contstan.html
- American Society for Testing and Materials (ASTM). ASTM Section D18.01.05 Draft Specification Content Specification for Digital Geospatial Metadata. http://www.ifla.org/documents/libraries/cataloging/metadata/astmd18.pdf
- Frequently Asked Questions concerning the FGDC's Content Standard for Digital Geospatial Metadata. http://biology.usgs.gov/fgdc.metadata/version2/faqa.htm

3.6. Education Network Australia Metadata Standard

by UPM.

Description:

The Education Network Australia (EdNA) is a national framework for collaboration on the use of the Internet in education and training.

The EdNA Metadata Standard is based on the internationally recognised Dublin Core Metadata Element Set (DCMES)[23] and is consistent with the Australian Government Locator Service (AGLS)[24]. The work of maintaining the EdNA Metadata Standard is

[23] http://www.dublincore.org/documents/dces/

[24] http://www.agls.gov.au/

conducted by the EdNA Metadata Standard Working Group which reports to the AICTEC[25] Standards Sub-Committee and the education.au limited[26] Board.

The purpose of the EdNA Metadata Standard is to support interoperability across all sectors of education and training in Australia in the area of online resource discovery and management. Adoption of the standard will assist people across education and training engaged in the production and use of well-described digital content. It will also support the technical requirements for well-structured coding of this content to exchange and serve up data on request. The principal application of the standard at present is to facilitate the aggregation of metadata about educational resources, from all states and territories, and all sectors of education and training, for EdNA Online.

The EdNA Metadata Standard comprises a set of guiding principles together with a set of metadata elements which are situated within the Dublin Core Metadata Initiative (DCMI) framework. These elements include the 15 Dublin Core (DC) elements (some with EdNA qualifiers) and EdNA elements. The EdNA elements extend the scope of description that can be included in a metadata record with information that has particular educational relevance. Some of the EdNA elements are specifically for the administration of EdNA Online.

Version 1.0 of the standard was first published in August 1998 after stakeholder consultation encompassing a period of 18 months. The current version (Version 1.1) was ratified by the EdNA Standards sub-committee (now known as the AICTEC Standards Sub-Committee) in December 2000.

Subset of Elements:

Table 7 includes the EdNA elements (the 15 DC elements are presented in section 3.1).

Element Name	Definition	Comment
Audience	A category of user for whom the resource is intended	This element provides the basis on which searchers may find resources relevant to particular audience(s). The element may be refined to include the education/training sector or level at which the resource is intended to be used.
Approver	Email of person or organization approving the item for inclusion in EdNA Online.	This may be the same as the publisher when organizations are describing their own resources, but different when nominating external resources. Recommended best practice is to use a persistent form for the value if available e.g., an email alias of the form: contact@organisation_name.edu.au is preferred. This element is used solely for EdNA Online purposes.
CategoryCode	A numerical code derived from the database tables	This element should contain numbers representing categories represented in the EdNA Online Browse

[25] http://www.aictec.edu.au/aictec/go

[26] http://www.educationau.edu.au/index.html

	which support the EdNA Online Browse Categories.	Tree[27]. This element is used solely for EdNA Online purposes.
Entered	Date item was entered as an entry in the EdNA Online item database (used for management purposes).	Created automatically by EdNA Online database software and should not be encoded into the metadata of source documents. This element is used solely for EdNA Online purposes.
Indexing	A parameter which identifies the extent to which the EdNA Online indexing ("spidering") software should follow hyperlinks from the described page.	The EdNA Online spider indexer, in performing full text indexing can optionally follow links in web pages which appear on evaluated websites and create indexes of these additional pages. These additional pages are not displayed through Browse but can be discovered in EdNA Online search results. This element is used solely for EdNA Online purposes.
Review	A third-party commentary or formal review of the resource. This element is defined such that two forms of review are accommodated.	Where the short form is used then recommended best practice is to use EDNA.Reviewer (see below).
Reviewer	Name of person and/or organization or authority affiliated with the review.	Recommended best practice is to include details of organizational affiliation.
Version	The nature or genre of the content of the resource. Version of the EdNA Metadata Standard applied.	For administrative tracking purposes it is recommended best practice to include this information.

Table 7. EdNA Elements

References

- EdNA Metadata Standard. http://www.edna.edu.au/edna/go/pid/385
- EdNA Online's Newsletters. http://www.edna.edu.au/edna/go/pid/1469
- "Metadata". Curriculum Materials Information Services. The Department of Education and Training. Government of Western Australian. http://www.eddept.wa.edu.au/cmis/cat/meta.htm

[27] http://www.edna.edu.au/edna/browse/0

3.7. MusicBrainz Metadata Initiative

by UPM.

Description:

The MusicBrainz Metadata (MM) Initiative[28] is designed to create a portable and flexible means of storing and exchanging metadata related to digital audio and video tracks. The MusicBrainz Metadata Initiative is a content description model for audio and video tracks on the Internet.

The MM Initiative uses RDF/XML[29] to facilitate the exchange of audio/video related metadata. This initiative describes an RDF namespace that should be used in conjunction with the Dublin Core[30] metadata recommendation.

This specification will be used by the MusicBrainz.org metadata server to communicate metadata queries. All the data returned from the server will be in the MM/Dublin Core RDF format.

This initiative defines the namespaces shown in table 8.

Namespace Abbreviation	Name	Namespace URI
mm	MusicBrainz Metadata	http://musicbrainz.org/mm/mm-2.1#
mq	MusicBrainz Query	http://musicbrainz.org/mm/mq-1.1#
mem	MusicBrainz Extended Metadata	http://musicbrainz.org/mm/mem-1.0#

Table 8. MM namespaces

Set of Elements:

The MusicBrainz Metadata namespace defines RDF classes and properties for expressing basic music related metadata. This namespace intentionally focuses on the primitives for metadata. There are 3 classes and 12 properties. The classes are shown in table 9.

Element Name	Definition
Artist	This class is used to describe an Artist. Inside of Artist, use dc:title, mm:sortName, and mm:albumList to describe the name and sortname of the artist as well as a list of albums available by that artist.
Album	Use this to describe an Album. The dc:title, dc:creator and mm:trackList properties should be used to describe the album.
Track	Use Track to describe one audio track. Use the dc:title, mm:trackNum, dc:creator, mm:trmId properties to describe a track.

Table 9. MusicBrainz Metadata classes

[28] http://www.musicbrainz.org/MM/index.html

[29] http://www.w3.org/RDF/

[30] http://dublincore.org/documents/dces/

The properties are shown in table 10.

Element Name	Definition
sortName	Use this to indicate the name of an artist for use in an alphabetically sorted list. If the dc:title field for an Artist contains "The Beatles", this field should contain "Beatles, The".
trackNum	This indicates the sequence number for a track in a given album. trackNum will only be used if a track number must be specified outside of the context of an album. When tracks are described inside of an album context, then the track list order defines the track number implicitly. Track numbers start with 1.
trmid	This property is used to indicate a Relatable TRM Id for an audio track.
cdindexid	Used to describe a CD Index Disk Id for a CD-ROM.
duration	Used to indicate the duration of a track in milliseconds.
releaseType	Indicates the type of a release. The possible values are: TypeAlbum, TypeSingle, TypeEP, TypeCompilation, TypeSoundtrack, TypeSpokenword, TypeInterview, TypeAudiobook, TypeLive, TypeRemix, TypeOther.
releaseStatus	Indicates the status of a release. The possible values are: StatusOfficial, StatusPromotion, StatusBootleg.
artistList	This property is used to convey a list of artists.
albumList	This property is used to describe a list of albums.
trackList	This property is used to convey a list of tracks.
cdindexidList	This property is used to describe a list of CD Index ids in an album.
trmidList	This property is used to describe a list of TRM ids in a track.

Table 10. MusicBrainz Metadata classes

References:

- MusicBrainz Metadata Initiative 2.1. http://www.musicbrainz.org/MM/index.html
- MusicBrainz Metadata Examples. http://www.musicbrainz.org/MM/mm_examples.html

3.8. DogmaModeler Ontology Metadata

by VUB.

Description:

This ontology metadata is being used in the DogmaModeler ontology engineering tool [20, 21]. This metadata has been created and is being successfully used, since 2002, for building libraries of ontological resources. It provides a specialization and an extension of the Dublin-Core metadata elements.

Set of Elements:

Table 11 presents the definition of the DogmaModeler metadata elements. The formal specification of these elements is presented in figure 3.

Element Name	Definition
Acronym	An abbreviation formed from the initial letter or letters of words in the ontology title. E.g. ‘CCOntology’, or ‘DOLCE’.
Title	The full and official heading name of the ontology, it may give a brief summary of the matters it deals with. E.g. ‘Customer Complaint

	Ontology', or 'Descriptive Ontology for Linguistic and Cognitive Engineering'.
Version	Information about the edition of this ontology. Typically, it includes version number, label, and date. Whenever the ontology is enhanced, updated or improved, it is often assigned a new version. Although versions represent the different states of an ontology during its life cycle, different versions are seen as different ontologies.
Number	A unique code assigned to the ontology for identification. This number is usually assigned by an ontology registration entity.
URI	Uniform Resource Identifier, the W3C's codification of the address syntax of an ontology. In its most basic form, a URI consists of a scheme name (such as file, http, ftp) followed by a colon, followed by a path whose nature is determined by the scheme that precedes it (see RFC 1630). URI is the umbrella term for URNs, URLs, and all other Uniform Resource Identifiers.
Genericity	The level of generalization where the ontology. The genericity level of an ontology is typically one of the {'Application', 'task', 'Domain', 'Core', 'Foundational', 'Linguistic', 'Metamodel'}. Examples: The CCOntology is a 'core' ontology; DOLCE is a 'foundational' ontology; "WordNet" is a 'Linguistic' Ontology. etc.
Language	The human language in which the ontology terms (i.e. labels of concepts, roles, etc) is expressed. In case this terminology is expressed in more than one language, the value of this attribute is 'Multilingual'. Recommended best practice is to use RFC 3066 which, in conjunction with ISO639, defines two- and three-letter primary language tags with optional subtags. Examples include "en" or "eng" for English, "akk" for Akkadian", and "en-GB" for English used in the United Kingdom.
DevelopmentStatus	The completion status or condition of this ontology, typically one of {Draft, Final, Revised, Unavailable}.
DomainSubject	A heading descriptor indicating the subject matter and the domain of the ontology. For example, e-business, sport, book-shopping, car-rental, etc. Typically, it is expressed as keywords, key phrases, or classification codes. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.
Context	Information about of the scope of the ontology, in which the interpretation (i.e. the intended meaning) of the ontology terminology is bounded. For example: the context of the WordNet ontology could be the English language, the context of the "CCOntology" is the EU complaint regulations, etc.
Description	Further information about the ontology. It may include but is not limited to: an abstract, reference to a graphical representation, a free-text account of the content, the methodology used to build this ontology, documentation, etc.
Creator	An entity primarily responsible for creating the ontology. Examples of creator include a person, an organization, or a service. Typically, the name of a creator should be used to indicate the entity.
Contributor	An entity responsible for making contributions to the ontology content. Examples of a Contributor include a person, an organization, or a service. Typically, the name of a contributor should be used to indicate the entity.
CreationDate	The date that is associated with the creation of the ontology. In other words, the first date in the ontology lifecycle. Recommended best practice for encoding the date value is defined in a profile of ISO 8601

	and includes (among others) dates of the form YYYY-MM-DD.
Rights	Information about rights held in and over the ontology. Typically, rights will contain a copyrights statement and other restriction for the ontology, and the cost description in case the use of this ontology requires payment. If the Rights element is absent, no assumptions may be made about any rights held in or over the resource.
SpecificationLanguage	The formal language in which the ontology is being specified; for example, OWL, DAML-OIL, ORM-ML, UML, KIF, etc.
Validation	An evidence about the testing activities of the ontological content. Such tests might be conceptual or ontological quality, syntax validation, etc. Typically, one should indicate the validation methodology and comments about the results.
Tool	The name of the tool by which the ontology has been developed, e.g. Protégé, DogmaModeler, etc.
Application	Citation to the application(s) using/has used this ontology. Typically, one should provide the name, URL, and some description about the application.
NumberOfConcepts	Statistics about the number of concepts in the ontology.
NumberOfRelations	Statistics about the number of relations in the ontology.
NumberOfAxioms	Statistics about the number of axioms in the ontology, an axiom typically is a formal definition/expression.
NumberOfInstances	Statistics about the number instances in the ontology.
IncludesOntology/ IncludedInOntology	A reference to another ontology, which is supposed to be included as part of this ontology. examples of this relation between ontologies is “Imports” in OWL, “inclusion” in Ontolingua, “Compose” in Dogma, etc. The formal semantics of such relationships is necessarily the same.
StepVersionOf/ PreviousVersionOf	A reference to the step/previous version of this ontology.

Table 11. DogmaModeler metadata elements

Figure 3 presents the formal specification of the DogmaModeler ontology metadata represented using the Object Role Modeling (ORM) graphical notation [22].

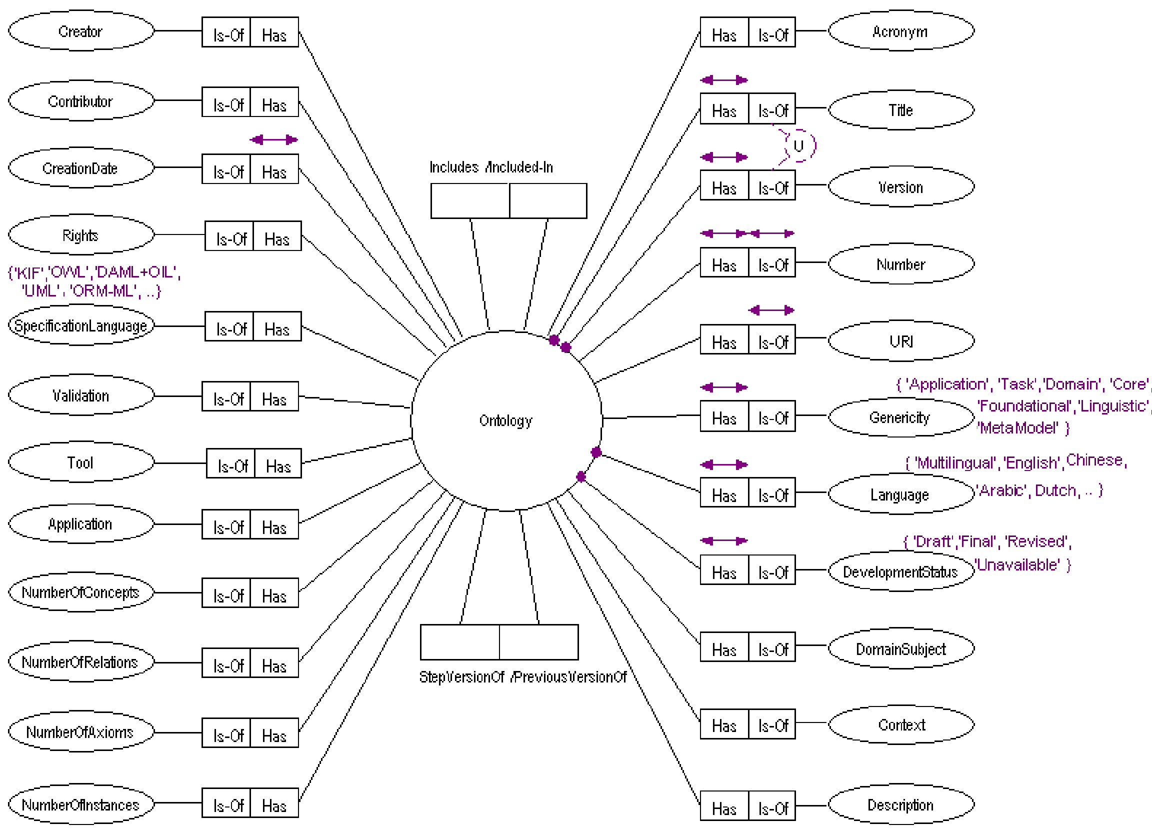

Figure 3. Formal specification of the DogmaModeler ontology metadata

In ORM, ellipses denote concepts (e.g. "ontology", "title") and rectangles denote relationships (e.g. "Has/IsOf", "Includes/IncludedIn"). Each relation consists of two roles. Rules in ORM can also be represented graphically: the mandatory rule "●" between a concept and a role denotes that it is mandatory for the concept to play this role. The uniqueness rule "⟷" on top of a role denotes that the concept can play this role only once. The uniqueness rule (U) between "title" and "Version" indicates that an ontology can be identified by the values of its title and version. In other words, different values of title and version together refer to diferent ontologies. The value constraint in ORM indicates the possible values (i.e. instances) for an object type, similar to the 'OneOf' restriction in OWL. A value constraint is denoted as a set of values {v1, …, vn} depicted near an object type, e.g. {Draft, Final, Revised, Unavailable}.

4. Definition of the Proposed Metadata

by UPM, FU Berlin, and VUB

In this section we propose a set of metadata elements to describe ontologies. For each metadata element, we provide: the element name, the label (the shortened form of the element name), a natural language definition, the cardinality, and a comment (with additional information regarding type, format, etc.).

Our proposed metadata elements for describing ontologies has taken as input the OntoRoadMap ontology from the OntoWeb[31] project. Furthermore, our proposed set of metadata elements (and their definitions) is an *adoption and an extension* to the metadata standards that we have presented in section 3, mainly Dublin Core and DogmaModeler metadata elements and specifications. In this way, we gain more adoptability of our elements and compatibility with legacy resources and systems.

We divided the set of metadata into three different levels, depending on the desired detail of each metadata:

- **Core** level for mandatory elements. The elements included in this level are needed for describing the ontology.
- **Desirable** level for optional, but still important, information about the ontology.
- **Extra** level for additional information about the ontology.

The proposed metadata elements, divided in the three aforementioned level, are shown in table 12. A formal specification of these elements, using the ORM notation [22], is presented in figure 4.

Element Name	Label	NL Definition	Cardinality	Comment
---------- CORE ----------				
Name	name	The name of the ontology.	{1, 1}	Typically, name is how the ontology is commonly known.
URI	URI	A Uniform Resource Identifier (URI) is a compact string of characters for identifying the ontology.	{1, 1}	URIs provide a simple and extensible means for identifying a resource. The specification of URI syntax and semantics[32] is derived from concepts introduced by the World Wide Web global information initiative.
URL	URL	The Uniform Resource Locator (URL) where the	{1, N}	The URLs give basically physical addresses of ontologies which are

[31] http://www.ontoweb.org/
[32] http://www.ietf.org/rfc/rfc2396.txt

		ontology can be accessed.		retrievable using protocols already deployed on the net [33].
Type	type	The type of information of the ontology.	{1, 1}	Examples of Type include but are not limited to: {Application, Domain, Generic, Representation. Foundational, Upper Level, Task, Core, Foundational, Linguistic, Metamodel, etc.}
Domain	domain	The nature of the contents of the ontology.	{1, 1}	The domain specifies the field in which the ontology is specialized.
Status of Development	status	Information that specifies the ontology status.	{1, 1}	The status of development tells how mature and tested the ontology is. Examples include: {Draft, Applied, Final, Revised, Unavailable, etc.}
Is implemented in	isImplemented	The ontology language in which the ontology is implemented.	{1, N}	Ontology Languages include: {RDF(S), DAML+OIL, OWL, KIF, etc.}
Language	language	The natural language of the ontology elements.	{1, N}	This element includes basically all existing natural languages like English, Spanish, etc. It is also possible to have the value 'Multilingual' if the ontology is in multiples natural languages.
Creator	creator	An entity primarily responsible for making the structure and content of the ontology.	{1, 1}	This creator could be a person and/or an organization.
Element Name	**Label**	**NL Definition**	**Cardinality**	**Comment**
---------- DESIRABLE INFORMATION ----------				
Acronym	acronym	The acronym or abbreviation	{0, N}	The acronym (a word formed from the initial

[33] http://www.w3.org/Addressing/URL/Overview.html

		given to the ontology.		letters of a multi-word name) or the abbreviation (a shortened form of a word) by which the ontology can be known.
Has been developed using Ontology Methodology	usedMethodology	The methodology used for developing the ontology.	{0, N}	The methodology used for developing the ontology could be: {None, Uschold and King, Grüninger and Fox, METHONTOLOGY, etc.}
Has been developed using Ontology Tool	usedTool	The tool used for developing the ontology.	{0, N}	The tool used for developing the ontology could be: {None, Protégé-2000, OILEd, OntoEdit, WebODE, etc.}
Contributor	contributor	An entity responsible for making contributions to the content of the ontology.	{0, N}	This contributor could be a person and/or an organization.
Is used in	usedIn	The project and/or the application in which the ontology is used	{0, N}	This element could be used to choose between different ontologies, in order to reuse them.
Has been evaluated	evaluated	The information about the ontology has been evaluated or not.	{0, 1}	This element means whether the ontology has been evaluated ({Yes, Not, Unknown}).
Has Documentation	documentation	A long description of the content of the ontology	{0, N}	Examples of Documentation could include: an abstract, an overview of how the ontology has been developed, a graphical representation of the ontology, etc. This element could include files in different formats.
Rights	rights	Information about rights held in and over the ontology	{0, N}	Rights could contain a rights management statement for the ontology (Intellectual Property Rights, Copyright, and various

				Property Rights).
Context	context	Information about of the scope of the ontology, in which the interpretation (i.e. the intended meaning) of the ontology terminology is bounded.	{0, N}	For example: the context of the WordNet ontology could be the English language, the context of the "CCOntology" is the EU complaint regulations, etc.
KR Paradigm	paradigm	The technique of Knowledge Representation that was used to create the ontology.	{0, 1}	KR Paradigm includes: {Frames, DL, etc.}
Version Number	versionN	The version of the ontology.	{0, 1}	Versioning could be useful for tracking, comparing and merging ontologies. The number could be incremented by 1, or a smaller or larger value, depending on the personal preference of the author. It would be recommended to use a standard.
Version Date	versionD	The date this version of the ontology was created.	{0, 1}	The format of the date should be a standard.
Includes Ontology	includes	The ontology could include ontologies.	{0, N}	This element refers to the different ontologies included in the ontology.
Element Name	**Label**	**NL Definition**	**Cardinality**	**Comment**
---------- EXTRA INFO----------				
Number of Concepts	numConcepts	The number of concepts included in the ontology.	{0, 1}	This element is expected to contain a numerical value, which represents the number of concepts in the ontology.
Number of Relations	numRelations	The number of relations included in the ontology.	{0, 1}	This element is expected to contain a numerical value, which represents the number of relations in the ontology.
Number of Instances	numInstances	The number of instances	{0, 1}	This element is expected to contain a numerical

		included in the ontology.		value, which represents the number of instances in the ontology.
Number of Axioms	numAxioms	The number of axioms included in the ontology.	{0, 1}	This element is expected to contain a numerical value, , which represents the number of axioms in the ontology.
Number of Metaclasses	numMetaclasses	The number of metaclassesinclu ded in the ontology.	{0, 1}	This element is expected to contain a numerical value, which represents the number of metaclasses in the ontology.

Table 12. Set of Proposed Metadata

These metadata can be divided into several categories:

- **Syntactical features** offer quantitative and qualitative information about the model and its underlying (graph) topology. An important set of quantitative metrics are typical graph-based metrics. Examples for syntactical features include the number of concepts and properties for each class, the depth of an inheritance tree, the number of incoming properties, number of concept instances, average path length, number of connected components. Since we consider ontologies in a networked context like the Semantic Web, one can also indicate how a model is embedded into the external structure by analyzing links to other networked information sources and references to external concepts.
- **Semantical features** are related to the formal semantics of the representation language and the meaning of the ontology content:
    - level of formality (i.e. highly informal, semi-informal, semi-formal, rigorously formal), representation language
    - ontology type (ontology: upper-level, domain ontology, thesaurus etc., UML diagram: state-transition diagram, activity diagram etc.)
    - ontology domain.
- **Heuristic and pragmatic features** refer to information about the engineering process the ontology resulted from: information about the guidelines, methodologies and tools used during the development process and the input information sources, status of the development, additional documentation or input used to generate the ontology (e.g. in knowledge acquisition tasks which might be part of the engineering process).
- **Contextual features** express information about the history of the ontology, for example when and by whom it was developed, whether multiple versions are available, or about the projects the ontology was used in.

The metadata elements can be categorized following the aforementioned categories as table 13 shows.

	Syntactical Features	Semantical Features	Heuristic and Pragmatic Features	Contextual Features
Core Elements	---	type domain isImplemented	name URI URL status language	creator
Desirable Elements	versionN versionD includes	paradigm	acronym usedMethodology usedTool documentation rights	contributor usedIn evaluated
Extra Information	numConcepts numRelations numInstances numAxioms numMetaClasses	---	---	---

Table 13. Proposed Metadata

Figure 4 shows a formal specification of the proposed metadata in the ORM graphical notation [22], and drawn using DogmaModeler ontology engineering tool [20,21]. For more details on knowing how to read the ORM notation, please see the end of section 3.8.

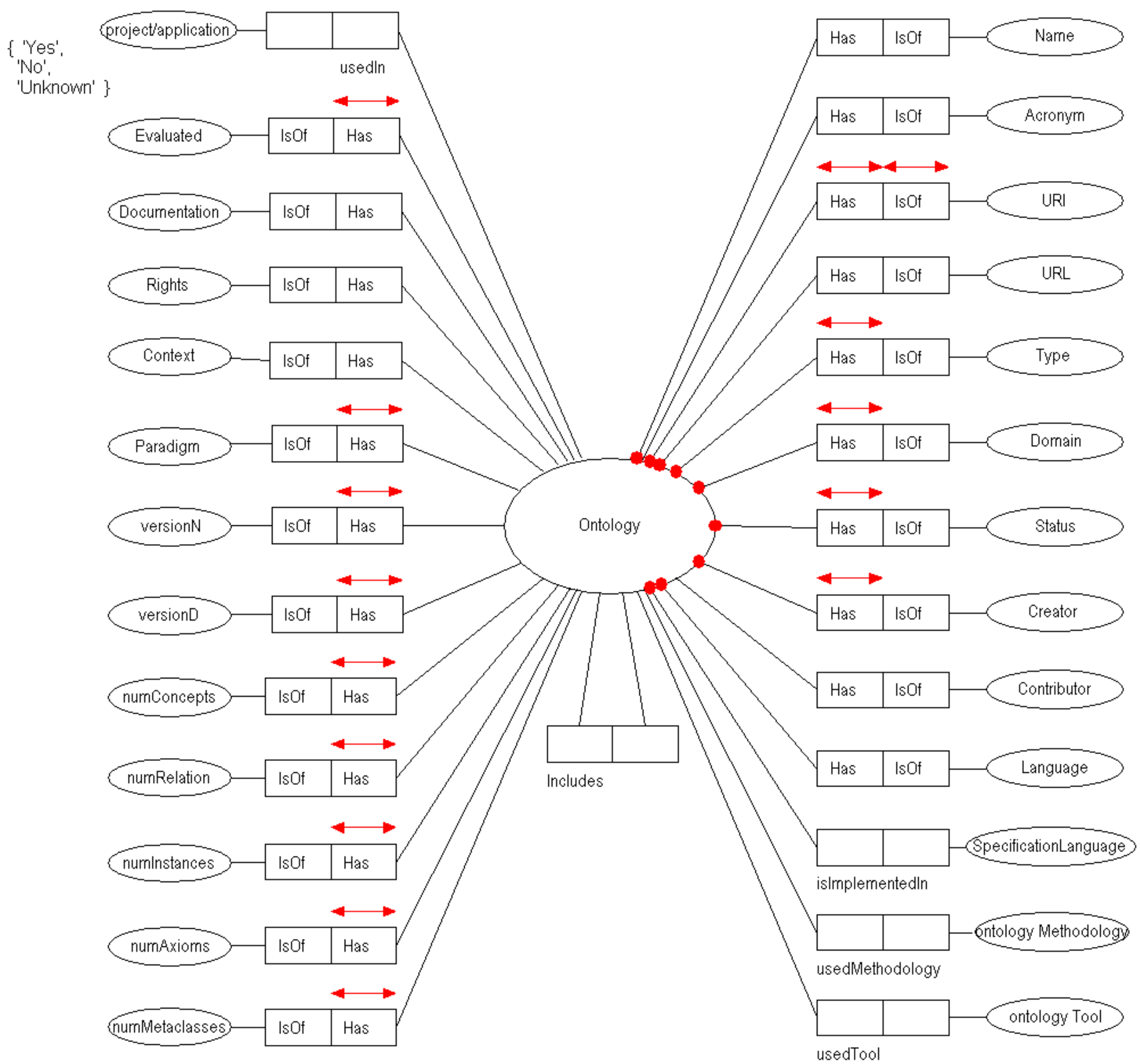

Figure 4. A formal specification of the Proposed Metadata, using the ORM notation [22]

5. Examples of Ontologies Annotated with the Proposed Metadata

In this section we make explicit the metadata proposed in section 4, using them to annotate two ontologies: the SWPATHO ontology and the OntoWeb ontology.

5.1. SWPATHO Ontology

by FU Berlin.

This ontology was generated in the project "A Semantic Web for Pathology" at the FU Berlin as underlying model for a retrieval system for image and text descriptions in the medicine domain. The ontology describes anatomy and diagnosis aspects related to lung pathology. The metadata associated with this ontology is shown in table 14.

Element	Value
---------- CORE ----------	
name	Semantic Web for Pathology
URI	http://nbi.inf.fu-berlin.de/research/swpatho/owldata/swpatho1.owl
URL	http://nbi.inf.fu-berlin.de/research/swpatho/owldata/swpatho1.owl
type	Domain ontology
domain	Pathology
status	Draft
isImplemented	OWL
language	English, German
creator	FU Berlin
---------- DESIDERABLE INFO----------	
acronym	swpatho
usedMethodology	None
usedTool	Protégé-2000, OilEd, SWOOP
contributor	Elena Paslaru, Sonja Niepage, Thomas Leuthold
usedIn	SWPATHO Project
documentation	The ontology describes concepts of lung anatomy and lung diseases and is used as basis for retrieval and semantic annotation tasks on medical data. It is based on common medical libraries like SNOMED, DigitalAnatomist and the UMLS Semantic Network.
paradigm	OWL DL
versionN	0.2
versionD	15.07.2004
includes	http://nbi.inf.fu-berlin.de/research/swpatho/owldata/umlssn.owl
---------- EXTRA INFO----------	
numConcepts	1000
numRelations	50

Table 14. SWAPATHO Ontology Annotated with the Proposed Metadata

5.2. The OntoWeb Ontology

by UKARL.

The OntoWeb portal is based on the OntoWeb Ontology, which represents the concepts and relations from the OntoWeb Portal (AIFB & VUB), the OntoWeb Edu Portal and the OntoWeb Ontology Roadmap Portal.

The OntoWeb ontology can be downloaded from the portal (http://ontoweb.aifb.uni-karlsruhe.de/Ontology/).

We integrated following content types (concepts) of the OntoWeb Ontology Roadmap Portal, the OntoWeb Edu Portal and the OntoWeb AIFB & VUB Portal.

The portal ontology contains the following concepts: "Application", "Business-Scenario", "Document", "Educational Ressource", "Event", "File", "Favorite", "Folder", "Image", "Job", "Language", "Link", "Methodology", "News Item", "Ontology", "Organisation", "Person", "Project", "Tool", "Topic", and "FeaturedContent".

Following the provided standards for ontology metadata, we present in table 15 the metadata for the OntoWeb ontology.

Element	Value
---------- CORE ----------	
name	Ontoweb Portal Ontology
URI	http://ontoweb.aifb.unikarlsruhe.de/Ontology/
URL	http://ontoweb.aifb.unikarlsruhe.de/Ontology/
type	Domain Ontology
domain	Knowledge Mgmt & Representation, Semantic Web, Ontologies
status	Final
isImplemented	RDF(S)
language	English
creator	Jens Hartmann (AIFB)
---------- DESIDERABLE INFO----------	
acronym	OW
usedTool	OntoEdit
usedIn	OntoWeb.org
documentation	OntoWeb.org
paradigm	Frames
versionN	1.0
---------- EXTRA INFO----------	
Number of Concepts	21

Table 15. OntoWeb Ontology Annotated with the Proposed Metadata

Example:


```
<rdf:Description                     rdf:about="http://ontoweb.aifb.uni-
karlsruhe.de/Organisations/Organisation.2003-04-04.0753">
<rdf:type rdf:resource="http://www.OntoWeb.org/extended#Organisation" />
   <title>AIFB, University of Karlsruhe, Germany</title>
   <description>Research Group Knowledge Management</description>
   <dc:indentifier>Organisation.2003-04-04.0753</dc:indentifier>
   <dc:contributor></dc:contributor>
   <dc:coverage></dc:coverage>
   <dc:format></dc:format>
   <dc:language>english</dc:language>
   <dc:relation></dc:relation>
   <dc:publisher>OntoWeb.org</dc:publisher>
   <dc:creator>hartmann</dc:creator>
      <dc:subject>AIFB</dc:subject>
      <dc:subject>Participant</dc:subject>
      <dc:date>2003-05-07 15:46:04</dc:date>
   <dc:source>http://ontoweb.aifb.uni-
karlsruhe.de/Organisations/Organisation.2003-04-04.0753</dc:source>
   <dc:rights>2001-2003(C)OntoWeb.org</dc:rights>
</rdf:Description>
```

6. Technological Support for using Metadata for Ontology Description: Oyster

In order to speed up the adoption of the set of structured metadata proposed in this document and to assist people from industry in evaluating how suitable is a given set of ontologies for a given application, in this section we present a brief description of Oyster, a peer-to-peer system for exchanging ontologies on the internet.

Oyster will be an instance of the SWAP System architecture as presented in [18]. The scenario envisioned is that ontology developers and ontology users share structured metadata for describing different types of ontologies with a Peer-to-Peer system, as researchers are sharing bibliographic references with Bibster [19]. Oyster could be enriched with more complex functionalities that currently are not supported by SWAP system architecture for measuring the usability and usefulness of ontologies in the registries for an application.

One of the Oyster key components will be the ontology that describes the main properties of ontologies. For this purpose, we will use the properties previously identified in section 4 in this document.

We distinguish three types of functionalities:

1. **Local repository management for ontology descriptions**. It includes:
   - Create, remove and update ontology descriptions in the local repository either by editing them, extracting them from ontology's libraries, or by taking them from others peers.
   - Identify duplicate ontology descriptions in the local repository.
   - Visualize ontology descriptions in the local repository, and whenever it is requested the upper level of the taxonomy.
   - Visualize duplicate ontology descriptions in the local repository.
   - Merge duplicate ontology descriptions in the local repository.
   - Import ontology descriptions into the local repository manually or by (semi)automatic methods.
   - Export ontology descriptions from the local repository into several formats, including RDF, XML or HTML.
   - Cluster different versions of a single ontology.
2. **Peer-to-peer network search of ontology descriptions**. It includes:
   - Search for ontology descriptions (including different versions of a given ontology) in the peer-to-peer network either using keyword searches or using the semantic facilities provided by the tool.
   - Define the scope of a query in the P2P network, either identifying a single peer, a set of peers or the entire network.

- The outputs of the search and query results could be a set of ontology descriptions, that might be imported in the local repository or merging them with an existing description.
- If ontology descriptions are duplicated, means for visualizing and merging the ontology descriptions with others from the local repository will be given.
- Export search results into different formats or languages.

3. **Processing of ontology description queries**. It includes:
   - Provide information to the peer-to-peer network about the network topology and query routing.
   - Obtain information from the peer-to-peer network about the network topology and query routing.
   - Advertise the available ontology descriptions to the peer-to-peer network.
   - Route queries through the network by sending new queries or forwarding queries from other peers.
   - Process queries and return answers for them.

7. Knowledge Web Standardization Strategy

Standardization is of particular importance for Knowledge Web, as ontologies are meant as a shared means of communication between computers and between humans and computers. To achieve this, ontologies should be represented, exchanged and accessed in agreed-upon open standards. The Knowledge Web Joint Program of activities includes in WP1.3 a tasks about standardization. Its goal is to propose a standard of structured metadata to describe the ontology content, and to approach suitable standardization bodies (W3C, ISO, IEEE, CEN, etc.) once the KW proposal of structured metadata has been created. We have decided in KW to carry out first the proposal of the structured metadata (presented in section 4 of this document) and then to approach the standardization bodies in the framework of the standardization working group at the SDK cluster.

The standardization working group at the SDK cluster aims at coordinating the standardization activitities of the EU FP6 projects DIP (see http://dip.semanticweb.org) and SEKT (see http://sekt.semanticweb.org/) and the EU FP6 Network of Excellence (NoE) Knowledge Web (http://knowledgeweb.semanticweb.org). The Standardization Strategy described in DIP deliverable [17] is intended to be a basis for DIP activities involving the coordination of standardization group evaluation, impact analysis and engagement approaches, but can also be made a common activity for the mutual benefit of all SDK Cluster projects. However the dissemination of standardization analyses, agreement on suitable standardization strategies and the development of content proposals must be undertaken internally with the express mandate of each project’s technical project management board, and should be communicated into the SDK Cluster activity for coordination.

For carrying out the standardization activities in task 3 at WP1.3, we propose the same set of activities proposed in [17]. They are:

1. the description and categorization of each standardization group;
2. the analysis of the impact currently being generated by a standardization group; and
3. current and future Requests for Proposal (RFP)

For the analysis of each standardization group a consistent approach needs to be adopted that would support the highlighting of the following differences:

1. Membership:
    a. Active members: also highlight what members are mainly interested in ‘driving’.
    b. Passive members: where membership is either for:
        1. status quo;
        2. building an understanding from the other members; or
        3. the neutralization of the intellectual property developed.

2. Scope of the standardization group:
    a. what are the main objectives – considering the scope for impact, e.g., focusing upon American or European only interests, or providing a global or local impact
    b. is the group perceived by the market as the owner of objectives identified, e.g., referenced by industrial journals and conferences as an authority
    c. are other standardization group alliances / liaisons identified that strengthen (or threaten) the group's position
3. Group processes & timeliness:
    a. are the processes defined well so that resource commitments can be easily specified and managed
    b. is the process of harmonizing proposal inputs effective from the point of view of keeping the intensions of the inputs and keeping the participants actively involved
    c. are efforts developed and complemented in an effective manner, coinciding with other standardization dissemination events (conferences, etc). Also, are the KW results going to be timely for the standardization input call?
    d. what is the level of focus dedicated to activities so that they are seen as highly effective and acknowledged as generating true impact – i.e., some standardization groups drop proposals with inactivity or non-agreement and are subsequently seen as ineffective by the market. (This could be analyzed in a quantitative manner through counting the number of successfully completed activities over the unsuccessful ones.)
4. Standardization group impact:
    a. market acknowledgement
    b. market adoption
    c. number of competitive implementations, and what are their individual focus
    d. research and development projects that utilize the standard as a basis
    e. analysis of the similarity to existing and planned KW results should be undertaken to determine if a unique project achievement can be attained
    f. analysis of the suitability of a standardization group for the presentation of a unique KW results should also be undertaken
5. Overall a SWOT analysis (strengths, weaknesses, opportunities & threats) should be provided that may include many of the elements above so that a more complete picture can be provided – particularly useful as background analysis when presenting to either the KW EPMB or other SDK Cluster activities.

8. Conclusions

by UPM.

In this deliverable we have presented a set of structured metadata for describing ontological content of evaluated and certified ontologies. This set of metadata has been divided into three levels: core, desirable and extra. The presented metadata have also been classified depending on their features: syntactical, semantical, heuristic and pragmatic, and contextual. Our next step is to propose this set of metadata for standardization, in order to define and create ontology registries.

This deliverable have also included a survey of metadata and standards, and an identification of metadata example widely used (like Dublin Core, LOM metadata, FOAF, BibTeX, etc.).

In order to show the use of the proposed metadata, in this deliverable we have included two ontologies (the SWPATHO ontology and the OntoWeb ontology) annotated with the set of proposed metadata.

Furthermore, this deliverable have also presented a short description of the Oyster system, a peer-to-peer system for exchanging ontologies in internet; and.

Finally, we have explained the standardization activities in task 3 at WP1.3, that is, the Knowledge Web standardization strategy.

9. References

1. Cathro, Warwick. (1997). "Metadata: an overview". Standards Australia Seminar, "Matching Discovery and Recovery". August 1997. http://www.nla.gov.au/nla/staffpaper/cathro3.html
2. "The Value of Metadata". http://www.fgdc.gov/publications/documents/Metadata/metabroc.html
3. Gilliland-Swetland, Anne. (2000). "Setting the Stage". In: Murtha Baca ed.: *Introduction to Metadata, Pathway to Digital Information*. Getty Information Institute. http://www.getty.edu/research/conducting_research/standards/introMetadata/2_articles/index.html
4. "What is Metadata?" Metadata Education Project. http://www.sdvc.uwyo.edu/Metadata/what.html
5. The Metadata Workbook On-line v.19991028. A Guide for Metadata Creation (10/99) by the Louisiana Geog. Info. Council for Metadata training workshops (LAGIC) http://lagic.lsu.edu/Metadataworkbook/
6. Schweitzer, Peter N. "Frequently Asked Questions (FAQ) on FGDC Metadata". (USGS). http://geology.usgs.gov/tools/Metadata/tools/doc/faq.html
7. "Metadata 101: a two minute introduction to Metadata" The Wyoming Geographic Information Sciences Center. Metadata resources at the University of Wyoming. http://www.sdvc.uwyo.edu/clearinghouse/metainfo.html
8. Berners-Lee. Tim. "Metadata Architecture". January 6, 1997. Status: personal view, but corresponds generally to the W3C architecture for Metadata. http://www.w3.org/DesignIssues/Metadata.html
9. Taylor, Chris. (2003). "An Introduction to Metadata". 29 July 2003. http://www.library.uq.edu.au/iad/ctmeta4.html
10. "Metadata". Curriculum Materials Information Services. The Department of Education and Training. Government of Western Australian. http://www.eddept.wa.edu.au/cmis/cat/meta.htm
11. Lassila, Ora. "Introduction to RDF Metadata". W3C NOTE 1997-11-13. http://www.w3.org/TR/NOTE-rdf-simple-intro-971113.html
12. "Why is Metadata important?" Metadata Education Project. http://www.sdvc.uwyo.edu/Metadata/why.html
13. Hart, David; Phillips Hugh. Metadata Primer "A 'How To' Guide on Metadata Implementation". National States Geographic Information Council. http://www.lic.wisc.edu/Metadata/metaprim.htm
14. "Using and implementing Metadata" Metadata Education Project. http://www.sdvc.uwyo.edu/Metadata/m_implement.html
15. Joint Workshop on Metadata Registries, Workshop Report, Draft 1.6, 1997. http://www.lbl.gov/~olken/EPA/Workshop/report.html
16. Karvounarakis, Greg; Christophides, Vassilis; Plexousakis, Dimitris. "Querying Semistructured (Meta)Data and Schemas on the Web: The case of RDF & RDFS". Department of Computer Science. University of Crete.
17. K. Elms, E. Dorner. "Deliverable 7.3. DIP standardization Activity". June 2004. (FP6 – 504083).
18. J. Broekstra, M. Ehig, P. Haase, F. van Harmelen, A. Kampman, M. Sabou, R. Siebes, S. Staab, H. Stuckenschmidt, C. Tempich. "A metadata model for semantics-based peer-to-peer systes". In proceedings of the WWW'03 WS on Semantics in Peer-to-Peer and GRid computing (2003).
19. P. Haase, J. Broekstra, M. Ehrig, M. Menken, P. Mika, M. Olko, M. Plechawski, P. Pyszlak, B. Schnizler, R. Siebes, S. Staab, C. Tempich. (2004). "Bibster-A semanticss-Based bibliographic Peer-to-Peer system". Proceedings of the third international semantic web conference. Springer. 2004. Pages: 122-136.

20. Jarrar M. (2005) “A principled methodology for Ontology Engineering”. Submitted. PhD thesis, Free university of Brussels (VUB).

21. Jarrar M., Demey J., Meersman R. (2003). “On Using Conceptual Data Modeling for Ontology Engineering”. In Aberer K., March S., and Spaccapietra S., (eds): Journal on Data Semantics, Special issue on “Best papers from the ER/ODBASE/COOPIS 2002 Conferences”, Vol. 2800, pp.:185-207, LNCS, Springer, ISBN: 3-540-20407-5, October 2003.

22. Halpin, T. (2001). “Information Modeling and Relational Databases”. 3rd edn, Morgan-Kaufmann.

23. Cathro, Warwick. “Metadata: an overview”. Standards Australia Seminar, "Matching Discovery and Recovery" August 1997. http://www.nla.gov.au/nla/staffpaper/cathro3.html

24. Dodds, Leigh. “An Introduction to FOAF”. February 04, 2004. http://www.xml.com/pub/a/2004/02/04/foaf.html

25. FOAF Vocabulary Specification. Namespace Document 2 Sept 2004 - FOAF Galway Edition. http://xmlns.com/foaf/0.1/

26. LaTeX: A Document Preparation System by Leslie Lamport, 1986, Addison-Wesley.