Metadata for learning objects: overview, prospects and test

Master thesis report/D- uppsats

Department of teacher education/

Uppsala Learning Lab,
Uppsala University, Sweden

by Jan Sjunnesson 26 April 2003

Supervisor: Donald Broady

Table of contents

Abstract

This thesis explores the various kinds of metadata that applies to digital learning objects. In a test six metadata tools are tried out on a digital version of a textbook for secondary schools. There is also an argument on the need to extended educational metadata, besides the ones explored and given by standardization bodies in education, knowledge management and information sciences.

Keywords: metadata, library catalogues, learning objects, Semantic Web, educational technology, knowledge representation, xml.

Acknowledgements

This thesis has been written with the kind support from Fredrik Paulsson, Donald Broady, Mikael Nilsson, Matthias Palmér, Janne Backlund, Monica Langerth Zetterman and the staff at Uppsala Learning Lab. Katarina Jandér and Jessica Lindholm at Lund university has also been helpful as well as the Netlab group there.

Parts of the work has been done undertaken within the framework of the project PADLR – a joint project on learning technology between the Learning Labs at Uppsala, Stockholm, Lower Saxony (Germany) and Stanford (USA)[1]. The thesis is dedicated to Rajasri Rao.

[1] http://www.learninglab.de/padlr/index.html

1 Introduction

Internet technology has changed the ways of learning, distribution and communication in many areas and will continue to do so. This thesis focuses on learning technology, metadata standards, tools and the future of the internet as it is shaped in its semantic content, on a new generation of web technology, the Semantic Web (see ch. 2 below),.

Many initiatives in education and technology intervene. In Sweden the government, industry, cultural and educational institutions try to foresee future changes that will evolve and bring users closer to cutting edge technology[2]. Users may be corporate staff, pupils, students, teachers or academics. Bringing all aspects into one study is hard and this thesis does not attempt to do that. Many factors are important in the development of internet based learning and resources; innovations, infrastructure, learning methods, markets, institutional responsiveness etc. This thesis concerns areas that seldom are put into one piece under one departmental heading or one subject. It deals with cutting edge web technologies, library catalogues from early 20th century, contemporary information retrieval projects and educational aspects on the new knowledge management technologies.

It is not easy to specify where this work could have been written since the area is distributed across many academic disciplines; Information and library science, ABM, computer science (AI, web technology, systems engineering, HCI), education, philosophy (epistemology), business (knowledge management) and cognitive science. Computer scientists will find the technical parts amateurish, and educationalists will perhaps not make in through them at all, bored with all codes and schemas. There is no easy way to explain this heterogeneous area at the right level, at least not for me.

The focus is on metadata standards and tools for indexing digital educational resources on the Semantic Web. A test is performed on a digital version of a textbook in philosophy for secondary school. The main question behind this thesis is to see what is needed and available to enable teachers, students and researchers to find, use and reuse digital objects captured from a book in an easy and mindful way. The technical part will be explained in section 2. This paper is an exploration of an unknown area rather than the fruit of traditional research. Overview, prospects and test are in focus. Selection of tools, methods and information management schemas has been done from practical concerns.

Choosing a regular book albeit in a digital version was done for two reasons; this book had already been used as a testded by researchers at the Swedish Royal Institute of Technology in an earlier project and a preliminary hypothesis was to continue this work with new tools and approaches. This turned out to be more complicated than

---

[2] For studies focussing on education (K-12 and higher), industry, innovation and research, see reports from IT-kommissionen at www.itkommissionen.se and in the references. One project of many in the cultural field is Nya vägar för boken at http://www.kb.se/Nvb An industrial project is stated in Aßmann (2001 a-b).

foreseen. Books have in their thesis state many qualities that digital objects not have and browsing through a handheld book seems still to me the most useful way to get an overview of its main features. This does not exclude the options that digitalization may give as recent discussions in Sweden show[3].

Another area of discussion that will be of interest when the reader has got the main points in this thesis is to what extent isolated digital learning objects can and should be placed in an educational context. Framing, constructivism, contextualization, situated learning, socio-cultural perspectives of learning etc. are all in favor of putting pieces into larger pictures, but with new tools this must not always be the case, or done by the teacher, or maybe should not be done at all.

At the end there is a discussion on extended educational metadata that is a start for an educational discussion where the future of learning technology and metadata is heading in the light of relevant theories of learning and instruction into the technology too, not just the learning objects.

1.1 Learning Objects

As already mentioned, the atom of the new learning technology in focus here is the *learning object*. This is a term that in the broadest definition designate any digital or physical object that might function as an instrument for learning, inside or outside educational systems.

What teachers, students, pupils and researchers share with one another are usually learning objects of all kinds. Making these items more available using digital representations and exchange systems would support their work and studying. Books, pictures, educational soft ware, video clips, diagrams, lesson plan, tests, laborations - anything that can be put into a course as a part or a whole lesson/ learning instance.

Terms of scale and sequence are important when defining these items, not being too small, too large and in which order (chronological, physical, logical, etc)

The mentioned broad definition of learning objects is the standard one in the most established system of learning object management, IMS-LOM which will be more considered in section 3.3. The literal definition states:

1) "Learning objects are defined here as any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning"[4]

This definition is has been criticized for being too broad and useless. A second alternative definition is proposed which gives:

[3] See Svedjedal (2000), Svedjedal (2001) and Peurell (2000).
[4] The first definition is stated by the Learning Technology Standards Committee at the consortium IEEE, see http://ltsc.ieee.org.

2) “A learning object is defined as ‘any digital resource that can be reused to support learning’” [5].

The second definition contains the concepts of reusability, being non rival (allow synchronous users) and its independence of larger systems such as courses and subject areas but leaves out the physical objects, humans, historical events and the concept of mere “referenced during” which does not include actual learning.

Here is my own definition that is more administrative than the above ones

3) A learning object is the smallest part supporting learning in a digital system of learning resources

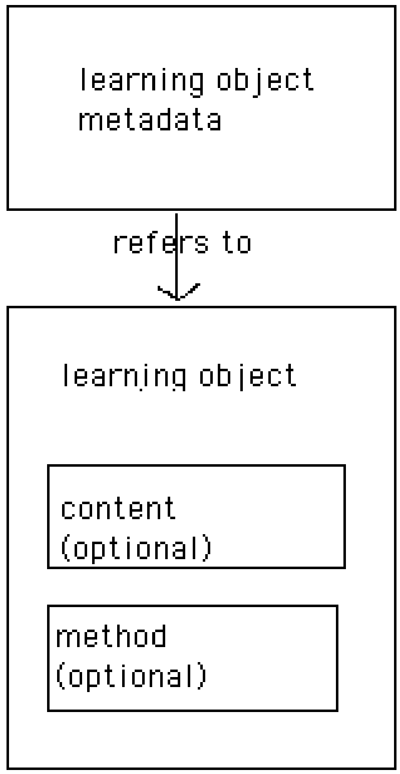
The advantage from 1) and 2) is that is specifies the learning object as an atom of a larger system. In this sense the features of learning resource management and design become more important, as will be shown in this paper.

It is not crucial to the test reported in this thesis to stick to any of these definitions though. The value of showing the various versions is to show that a discussion of learning objects is going on that is fruitful to know for anyone working with information retrieval of digital resources for education.

The learning object itself can be very barren of content and use. A digital picture for instance, or even less, an application that supports showing digital pictures one by one in a narrative way but without any pictures in it.

But this may not be a problem here since the main focus in this thesis and the current discussions has been in the *information* about the content of the object, its so-called *metadata* (see section 2.2). On the next page is a figure that shows the relations in and to a learning object, its metadata and aspects of method and content [6].

[5] Wiley (2000), p. 7.
[6] The picture is from Koper (2001). p. 4.



The metadata is here shown separated from and referring to the learning object, which itself could contain various kinds of methods and contents.

This figure and various learning object definitions may look as not much at all but the profits, uses and technology has created huge interest and high economic expectations. The value of online learning market is said to be $11.5 billion in 2003[7].

Whenever educational institutions share the same (digital or digitally represented) object, they should use the same classifications and not lock their objects in special applications that are not moveable. The commercial advantage is to generate content that is more crossover platform and by that lower costs of investment and development.[8] That is one basic idea behind the large interest, but there more.

The aim is not only to find intelligent ways to educational digital material that are designed to be used in classrooms and courses, but also to be able to use other digital material not primarily designed for educational purposes. Maps, photos, statistics and many other working materials from the world outside schools and universities would and should be more digitally available for learning purposes.

---

[7] See reference to http://wrhambrecht.com in Wiley (2000), see also Jacobsen (2001) and Downes (2001)
[8] One major institution in managing digital educational resources is IMS. For their view on current faults of learning management, see http://www.imsglobal.org/faqs/imsnewpage.cfm?number=5.

Another main idea is besides making standard scientific learning objects such as the sine wave functions available to students in a global format. There should also be opportunities for accesses to original texts that many agree are the core of human history. Works of Shakespeare, religious documents, canonical art and music, descriptions of historical events such as the Holocaust etc. form the basic material of many courses around the world.[9] Not discussing the canonical worth of these examples here, we can see that there are some basic learning advantages behind using a digital format for agreed objects.

Provided that an English class would spend many more hours on Hamlet than the average engineering program, still the engineering students would get a least the standard interpretations from a digital version whereas the language students would benefit a lot more from in-depth hermeneutic studies of the same text, but more expanded with annotations, links and other learning devices. Both student groups use the same text, but for different purposes. Engineering students would get a "Hamlet Light" – agreed – but this might be better than diving into a maze of sophisticated considerations that are provided for English students or drama historians through another navigation prepared by their teachers, all using the same text. It is all a matter of providing open learning situations where students and teachers could stop or go on in the material as they like or need to.

But the task of finding the right information is hard since the initial structure never was directed towards finding the right content or information. One is the initial decisions behind the infrastructure of the WWW technology.

"Internet has many virtues, but it – and in particular the WWW- was not designed specifically for information retrieval"[10].

This lack of qualified information retrieval support in the web is still true to some extent, but there has been an enormous development since then, which the quotes above prove. This thesis will try to cover some of that land since the quote above written in 1997 but the development is so fast that these words are already inaccurate when the thesis gets in print.

1.2. **Navigation, social and educational**

To find digital resources may not be all that problematic but how do we use them? Which to trust? Commercial web technology like the online bookstore Amazon hints to buyers that provided one has bought one book; a list of 5-10 other titles might be interesting. All done automatically by servers in the recommender systems service who know nothing but statistical inferences between similar books.

In learning situations on the web the same methods could be used but other considerations must be taken into account. Seriousness, trust and purpose are what educational authorities want to spread but that might not be as easy if material is more loosely put into open learning repositories. It will be a challenge to teachers and

[9] See Downes (2001), p.6, for discussions of "Hamlet" and "Holocaust".
[10] Quoted in Björkhem/Lindholm (2000), p. 10.

schools to build that trust when technologies proliferate that enable students to annotate all sorts of learning objects with their own intentions. E.g. “this class sucks”, “don’t download this, it’s boring “ etc. The new standard of metadata framework RDF works like that (see section 2.4 below).

Research in fields like social navigation claims those tools and information agents working with narratives and interaction are more useful than others are[11]. *Social navigation* in this sense is something that grows *dynamically*[12], like walking down a path in a forest whereas walking down a city road is not. Another feature is *personalization*, like talking to a person at a help desk at an airport, whereas reading a sign containing the same message is not[13].

Building trust is a kind of *reputation management.* A person’s recommendation let you evaluate their recommendations and determine how much trust you might want to put into them[14]. Tools will emerge that let users find what recommendations a specific person has done.

The learning environment will not be isolated in this development but needs to take advantage of dynamic open systems that not bring commercial interests into schools but rather build web rings, communities, common devices etc. of reliability. Navigation is a key word here as it points to the most direct contact with users. How to navigate in digital environment will be a civic virtue for 21st century students and citizens. The next section will show the technology needed to implement that vision.

---

11 Höök et al (2000) and Benyon/Höök (1997)

12 An example from internet is http://slashdot.com as analyzed at http://wiley.ed.usu.edu/docs/ososs.pdf

13 Dieberger et al (2000). In Recker/Wiley (2000) they distinguish between authoritative and non-authoritative metadata is a similar manner. For this discussion, see also Naeve/Nilsson/Palmér (2002) and Jacobs/ Huxley (2002).

14 Dieberger et al (2000). p 6. See http://kmr.nada.kth.se/comp/scenarios/search-scenario.HTML for a scenario where filters are uses that discern the status of the recommender, eg. professorship.

2 Semantic web

2.1. Introduction to the Semantic Web

The internet grew as we all know incredibly fast in the mid 1990s. Commercial and personal web pages flooded the otherwise still community of academics and computer scientists/ amateurs. Meaningful navigation became the main issue for the WWW-inventor Tim Berners-Lee when he in 1998 launched the vision of the *Semantic Web.*

Telling of the first part of the dream of the WWW protocol, it was becoming "a much more powerful means for collaboration between people/ . . . / a dream of people-to-people communication", in the second part of the dream,

> *"collaboration extend to computers. Machines become capable of analyzing all the data on the Web- the content, links and transactions between people and computers. A 'Semantic Web', which should make this possible, has yet to emerge but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines, leaving humans to provide the inspiration and intuition. The intelligent 'agents' people have touted for ages will finally materialize "*[15]
> .

Small digital devices will guide us to information by talking to each other, but only if we have provided them with enough guidance. That support should be lined out by humans of course but written so that machines can understand them. The Semantic Web is not a separate web but an extension of the current one in which information is given well-defined meaning, as well as better enabling computers and people to work in cooperation.

Roughly, there are two conceptual differences between the Semantic Web and the regular Web:

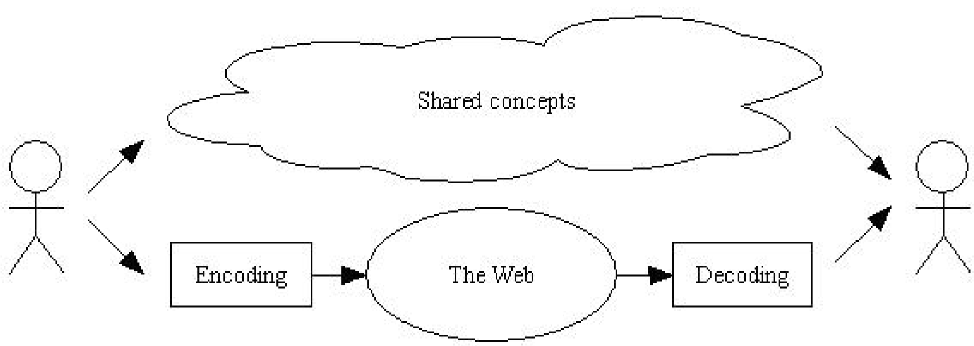
1. The Semantic Web is an information space in which information is expressed in a special machine-targeted language, whereas the Web is an information space that contains information targeted at human consumption expressed in a wide range of natural languages.
2. The Semantic Web is a web of formally and semantically interlinked data, whereas the Web is a set of informally interlinked information.

[15] Berners-Lee (1999), p. 157 –8. See also Berners-Lee (2001), Berners – Lee (2002), Palmer (2001), Gustavsson (2001), Aßmann (2001), Iselid (2001), Fensel et al (2003), http://home.swipnet.se/semanticweb and www.semanticweb.org

Tim Berners-Lee has forecast a situation where people would carry personal digital devices connected to the web through semantics in order to make their daily lives easier[16]. In his telling example, a man gets call from his sister who says that their mother is sick and needs to see a specialist soon. The sister tells her device (or "Semantic web agent") to find a doctor with certain skills, but also to find lots of other information on the planning and scheduling of not only her but also her mother's, her brother's and the available doctor's agendas who would interact to find times for all three in the next two weeks. This could not be done with the web as it works today.

The so –called *metadata tags* that mark up the common web pages written in HTML (the code language that presents content on the web as texts, pictures etc) were not designed to be aware of (meta)data such as time, location, specialization in medicine, not to speak of making several agendas from different people exchange information. HTML can however support some metadata structures but the semantic content is not sufficiently rich enough.

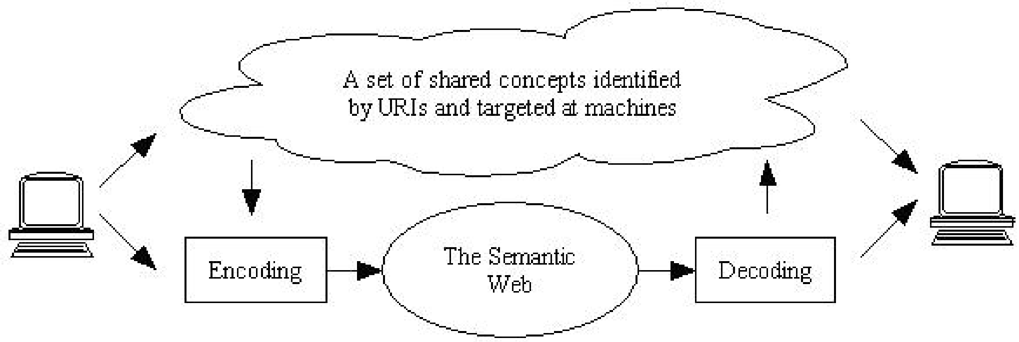
Below it is shown how concepts are being shared in information exchange over the web. In fact these concepts could be stated in grammar, vocabulary, or HTML. In a normal non-computerized setting humans also use shared concepts such as grammar, logic, emotional and bodily sign languages. In Figure 1, this is exchange is shown over the web. Encoding HTML by a web designer and decoding is when a user views the content presented along with the shared concepts, in this case the normal web code language HTML.



[17]

In the next picture, this information of shared concepts is shown in detail and pointing out that several more machine understandable ways must cooperate.

[16] Berners-Lee (2001)
[17] Both figures are taken from Benny Gustavsson's semantic web project available at http://home.swipnet.se/semanticweb/200108/diff_use.HTML. In Gustavsson (2001) the logical theory of the Semantic Web language is presented in detail.



Machines need to be able to speak to each other in this vision by using well-defined digital measures such as URI. This acronym reads “Universal Resource Identifiers. A subset to URI are called URLs (Universal Resource Locators – eg. a web address beginning with http://…). Targeting at machines means to identify a URI that represents some kind of knowledge. Earlier versions of knowledge representation systems did not use URI, which can also be a unique number that could point down to any part of an object, digital or otherwise. A disadvantage is that not all material has URI, books for instance. The targeting process must be machine-readable but structured so that humans can benefit from them.

As we shall se later library catalogues are a kind of knowledge representation that all can understand and relate to. Index cards with titles and classifications represent the information of contents and other features of eg. a book, but does not tell exactly on what shelf the book is located. It gives a number or letter series but does not specify the shelf. The library could be torn down and rebuilt only if the classification remained. New bookshelves, new samples but same content.

The idea of Semantic Web is that it could work like a gigantic container of resources that speaks to each others and not relating to the geographical or local hosts and servers but the content structure. This idea aim at separating presentation from content, which means that the same content will be presented in various ways.

To do this, we have to have some means that tells what the document or object is about, rather than how it will be represented and with what software program. This *metadata* are the backbone of much of what the new web technology is all about as well as it is in the library catalogue. Putting the metadata into the right places yes, but by whom? Who pays? And why? - are all critical points of the Semantic Web vision[18], but luckily we do not take the full discussion here on its possibility in practice. Librarians should in any case be of great resource here.

[18] http://www.hyperorg.com/backissues/joho-jun26-02.html#semantic

2.2 Metadata

A major ambition with the new technology in the Semantic Web vision is to mark learning resources with metadata. Metadata is usually defined as "data about data". A library index card is common metadata item, just as any information that informs a user, machine or human, about an object[19]. In this thesis the focus in on machine-readable metadata. This digital metadata can be attached to a file associated with the (learning) object or put directly inside the computer code of the object. Just like a library card and a book cover are different representations of the same metadata of the book, e.g. its title, author, edition and so on.

There are as shown several kinds of digital metadata. Here are the main classifications:

Internal metadata – built on mark up language, often written in the code of the object
External metadata – built on mark up language, stored externally but attached to object
Structual metadata – systems for learning design, content packaging, software applications etc

In this paper all kinds will be considered but the emphasis is on external metadata.

Reasons for using metadata are simplicity, flexibility, interoperability and standardization. With these features a user should find her way through archives, catalogues etc., providers should make little effort to file it properly and if everybody works with the same standards, catalogues should interact automatically. What metadata should enable users to do is to find, identify, select and acquire the desired objects, if possible according to the international library organization IFLAs requirements[20]. Many similar organizations seem to agree on what metadata is all about and it is of course useful if everybody agrees on some standards. But there are situations when metadata standards are not efficient. Since there is no "metadata police" on the web it is possible to work in various standards simultaneously (see section 3.7).

Organizations calling for strict adherence to standards are not alone in the endeavor of cataloguing since some computer technicians for instance want their material and ideas to be shared in new ways. Many of these new ideas seem to come from experiments with and discussions of peer-to-peer technology, open source/Linux and everyday use of file sharing, communication in web communities/mailing lists etc.

---

[19] http://www.w3c.org/DesignIssues/Metadata.HTML and http://www.w3.org/Metadata/Activity.HTML . For a good overview of metadata and web technology from a library perspective, see Björkhem/Lindholm (2000). The next paragraph relies on their discussions at p. 13 and 22.
[20] See Besser (2002), Frohman (2000), Frohman (2001) and Arms (2000) for critical discussions of the web as a digital library.

According to the Knowledge Management Research group at the Swedish Royal Institute of Technology there are misconceptions about metadata[21]. They state the misguided metadata conceptions as

- metadata is objective data about data
- metadata for a resource a produced only once
- metadata must have a logically defined semantics
- metadata can be described be metadata documents
- metadata is the digital version of library indexing systems[22]

These criticisms are relevant to this thesis since it will explore some tools that definitely does not support preformed semantics and objectivity. The traditional metadata community, e.g. libraries, is still working with metadata that rely on these conceptions. But with the growing e-learning technologies there will probably be a blend of objective metadata (bibliographical data, copyright issues etc) and subjective annotations and conceptual navigation through learning resources.

2.2.1 Metadata and HTML

Metadata are already a part of web technology since HTML may use metadata in its header (mostly but also elsewhere) that tells something about the content of the web site[23]:

Below is a sample from metadata entries of a common web page written in HTML. It is from the Swedish government's web. After the tag <meta> we can see the metadata contents. This information is indexed by search engines and is necessary but not at all sufficient in order to structure information retrieval better

<title>The Swedish Government</title>
<meta name="description"
content="This website describes the swedish government and its policies. The latest press releases, webcasts, publications, reports and bills from the swedish government.">

<meta name="keywords"
content="sweden, swedish government, ministry, prime minister, minister, policy, legislation, news, reports, publications, press">

<meta name="robots" content="index, follow">
<meta name="dc.creator.address" content="webbred@regeringen.se">
<meta name="formgivning" content="tommy.sundstrom@crosscom.se">
<meta name="GENERATOR" content="Microsoft FrontPage 4.0">
<meta name="ProgId" content="FrontPage.Editor.Document">

This information after the meta tags may look all right but HTML is not a good way of expressing metadata for many reasons. The code may be unstructured but that is

---

21 http://kmr.nada.kth.se
22 Nilsson/Palmér/Naeve (2002). See also Recker/Wiley (2000).
23 www.sweden.gov.se

not a problem since browsers are very permissive. The more important problem with metadata is that the tags do not have precise descriptive contents[24]. One very useful way to express content accurately is used here and that is the part of the meta name that begins with *"dc"* which stands for Dublin Core (a widespread object specification, see section 3.2).

There is a difference between *octopus* as a culinary dish, a sea creature and octopus as a metaphor for power. Nowhere in an HTML document is it possible to divide these various meanings of the word *octopus*. A search through MSN search engine gives these hits:

- **Octopus, Äldrevård och Demensfrågor**
- **MSN Encarta - Octopus**
- **About Octopi**
- **Monsters of the Deep - Blue Blood**
- **Did You Known That an Octopus Has 3 Hearts?**
- **Class Cephalopoda**
- **Tokyo Food Page**
- **Seafood Recipe Archive**
- **Octopus Newsgroup Postings**

For schools with little time, few computers and divided subjects are this result not satisfying. Metadata must be ordered in a more content oriented way.

The need for sorting web pages and other digital objects out from one another has produced many new metadata specifications and standards. A metadata specification informs the user about the characteristic features about the object that is relevant for its purpose. The specifications are agreed upon within its community, scientific, commercial or otherwise granted by representative bodies.

There are specifications that are relevant to education and learning resources management. In section 3 are the most important ones presented but there are more initiatives still[25]. But first we will take a look at the technologies XML and RDF that supports metadata sharing.

2.3 XML

Keeping track of documents according to content rather than graphic presentation or technology has been the aim of the markup language SGML since 1980s. SGML (Standardized General Markup Language) is important as a forerunner of XML (Extended Markup Languare). Infact, XMl is subset of SGML designed for the web. SGML defines in a meta vocabulary the permitted elements, signs and syntax for any

[24] See Wallberg (2000) for disadvantages with HTML in metadata retrieval. Also Kronman/Parnefjord (1999), p. 3. See also http://www.lub.lu.se/svenskmetadata/
[25] SCORM - http://www.adlnet.org, EML- http://eml.ou.nl, LMML – http://www.lmml.de and SIF – http://www.siia.net/sif

markup language, whether used internally in databases or in information exchanges. HTML is for instance written with the SGML specification of a markup language[26].

Neither SGML nor HTML is able to separate content and presentation in a manner that satisfies both order and appearance. It was inevitable to create another markup language, XML that prescribes grammar and vocabulary for metadata that includes the web. The metadata that are written with reference to XML does not include the presentation of the content, but contains a framework of metadata that enable the metadata structure to tell what the content of the information is about.

XML can be written for any document structure, which is why it well used in the area of learning objects of all kinds, weather audial, visual, graphic, textual or physical. If the information written in XML is well defined according to XML syntax and logic (being non permissible to mistakes), a machine will be able to draw out content and conclusions that are intelligible to humans and computers. Important for XML editing is that the document being marked with XML tags, has open and closed tags where each tag can have a certain value or attribute. The vocabulary and the combinations are not fixed, hence its widespread use and flexibility, but there are two most widespread applications of XML[27] .

One may say that XML technology is too wide to specify the contents of all sorts of areas. XL specifies how a meta language should be written in its syntax and structure but not the contents and values etc. For this purpose there has emerged within professional and representative bodies markup languages e.g. MathML for mathematics, ChemML for chemistry, SMIL for multimedia etc. These are written in XML syntax but the tags, elements, attributes, values etc of the tags are relevant only to certain areas. Other domain related specifications might use XML but also other techniques.

Specific educational markup languages are in the making and this thesis will explore some of them that are independent of markup technique. The difficulty of keeping track of education is that several factors are in motion; the technology that will be useful to educational institutions, information retrieval systems, library indexing and so forth. The very general area of learning and education, where almost everything could be an object of study and research, makes it different in relation to other more stable areas. The human, emotional and social, not forgetting national and regional, aspects of educational settings are also factors that are hard to express in logical sequences that computers like.

A more forceful technology may enable users to do just that, RDF.

---

[26] See Östlund/ Hermundstad (2001) for an XML introduction in Swedish. For latest accurate information see the W3C consortium at http://www.w3c.org/XML.
[27] Two techniques exist for developing the logical structure of an XML document, Document Type Definitions (DTD) and XML Schema. See latest news at http://www.w3c.org

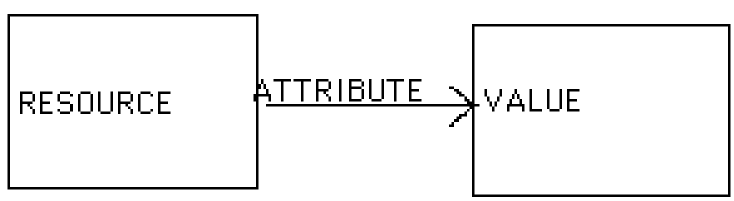
2.4 RDF

Resource Description Framework (RDF) in a framework for metadata standards. It is a foundation for processing metadata and supports interoperability between applications that exchange machine-readable information. Hence RDF is very important for the intelligent "agents" that Berners- Lee dreamed of.

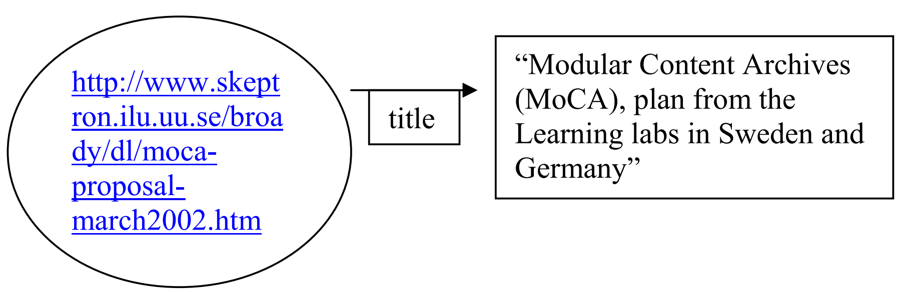
RDF consists of three parts[28]:

1. Syntax (XML or other) represented in graphs or relations
2. Specified semantic specifications (such as the global bibliographical metadata standard Dublin Core, considered below at section 3.2)
3. A unique RDF data model which consists of three entities, Resource- Attribute – Value

In this model there is a unique identifiable resource, that has certain attributes, which has certain values. The resource can be a web site, a book etc, but must be represented by a URI (Uniform Resource Identifier). Below is a the model shown empty



If this model is filled with information another triple appear:



---

[28] See Kronman/ Parnefjord (1999) for an introduction in Swedish. For latest accurate information see http://www.w3.org/RDF/. See also Miller (1998), Bray (2001), Decker et al (2000) and Miller (2002). Development of access to educational resources by RDF is done , among others, by the Edutella team at http://edutella.jxta.org/ (see Nilsson (2001b) for an overview)

This information can be varied so that the middle term title can appear as a value in a another scheme and so on. It is not specified what contents should be in each, just their roles as resources, values or attributes. They may all be web sites for instance. But the flexibility is structured so those standard metadata vocabularies are used for attributes. What a title is is specified in the Dublin Core metadata specification (see section 3.2) which is written with an identifier that says which standard is used.

An advantage is that RDF enables people to name resources in ways that uses both flexibility and standardization but extends by far the contemporary use of internet and information retrieval. Several metadatamodels can be used for the same learning object, which is a great advantage of RDF. A vast area of metadata use for learning opens up as the comment below shows.

"The most fundamental benefit of RDF compared to other metadata approaches is that using RDF, you can say anything about anything. Anyone can make RDF statements about any identifiable resource" [29].

These statements could include descriptions, certifications, annotations, version trackings, reuse etc. Combining these statement for educational purposes four areas appear important[30]

- Intelligent software agents finding relevant information
- Personal annotations of any learning resource
- Collaborative and distributed authoring and course construction
- Reuse of learning material

These four new and dynamic areas of digitally supported education and learning have to be known to schools and higher education institutions, but today very few teachers and education managers know of these features of recent web technology. This concerns both XML and RDF. The Swedish Agency for Education is trying though to get across with some more popular information on the educational uses of these technologies to the school community[31].

2.4 Ontologies

Ontology is a philosophical term that designates what exists and started in ancient philosophy, but has lately been used in computer taxonomies to specify the logical structure of a controlled vocabulary. Not only syntax must be specified as with XML and semantic triples as with RDF, but also what rules must follow in a certain area.

---

[29] Nilsson (2001a), p. 3.
[30] Op cit, p. 7
[31] http://www.skolutveckling.se/skolnet/smultron/infostruktur/

Ontology in computer science deals with classes, slots and individuals. Classes are collections of objects that may also be subclasses to other, such as books may be to learning objects, or humans to mammals or living creatures[32].

Much of it is simple logic. To be a person is to be a part of the class of mammals.

Class-def person
Subclass – **of mammals**

Slots specify binary relations or subrelations.

The slot son –of could be defined as

Slot-of has son
Subslot of has child

where value or individual of the subslot above is a mother or a father.

Slots can have constraints. One may require that the value or individual filling in a has son-slot must be a male

slot-constraint – has son
value-type male

What slots of which classes must be represented and how will they interact with others? Are they optional? An example, if something is a learning object, is age of audience required? In order to achieve machined reasoning ontology languages are very important. 19th century philosophical research by Gottlob Frege and C.A. Peirce built the now classical First Order logic. It is the traditional kind of (symbolic) logic that build all languages formally on five basic existential primitives which combines to make logical conclusions[33].

Table of the five semantic primitives[34]

Primitive	**Informal Meaning**	**English Example**
Existence	Something exists.	*There is a dog.*
Coreference	Something is the same as something.	*The dog is my pet.*
Relation	Something is related to something.	*The dog has fleas*
Conjunction	A and B.	*The dog is running and barking*

[32] Examples taken from Bechhofer et al (2000)
[33] A primitive is a category of an ontology that cannot be defined in terms of other categories in the same ontology.
[34] Example taken from Sowa (2000a). For more on ontologies and knowledge taxonomy, see Frängsmyr ed (2001), Maedeche (2002), Sowa (2000b), http://www.jfsowa.com, http://www.ontoprise.de and http://www.ontoknowledge.org

Negation	Not A.	*The dog is not sleeping.*

Ontology works as an underlying structure in knowledge- and learning management technologies and is a key means to achieve the Semantic Web vision. The inferences that was also developed by Aristotle in his syllogistic still work in computer algorithms for building knowledge and represent it.

Ontology engineers will have to work with domain experts in order to develop functionality for users of all kinds. In education this means that educational taxonomic ontologies have to merge with commercial and other kinds that may intervene in education (standards, administrations etc).

The practical applications of XML, RDF and ontologies will be shown in section 4, but now we will consider some metadata specifications that are relevant to education and information retrieval.

3 Specifications

3.1 Introduction

Metadata technology requires established specifications. They bring content closer to users than XML, RDF and ontologies do. They describe the syntax needed (XML or other), scope of use (K- 12 education for instance) and set up a logical structure with elements, attributes and relations that are domain specific. The domains are education, research and learning here, but also library catalogues will be considered since they rely on very stable and much older standards than those of web technology, not to mention Semantic Web technology which is still in development.

Below some specifications will be brought up that almost all later will be used with metadata tagging tools. These are the most used specifications but there are others.

3.2 Dublin Core

The specification Dublin Core (DC) is mainly bibliographical but is useful in many other areas too where documents such as books, articles etc are used[35]. It is by far the most widespread indexing on web documents and is also used when indexing all kinds of items. In the HTML example above two lines were written in bold, namely

```
<meta name="dc.creator.address"  content="webbred@regeringen.se">
```

The tag element

"dc.creator.adress"

is written with reference to the Dublin Core standard (i.e. "dc") and the value for it is "webbred@regeringen.se">

This enables the Swedish government to be reached by search engines in well-structured and internationally standardized manner. But since there are no rules in metadata tagging in HTML, the government could invent their own ways of expressing metadata content.

[35] http://dublincore.org/ . For Swedish references on DC see Hedberg 1999a-b and http://www.kb.se/Bus/Metadata/dc/default.htm .

Dublin Core (DC) consists of 15 elements:

Title (Titel)	Contributor (Medarbetare)	Source (Källa)
Creator (Upphovsman)	Date (Datum)	Language (Språk)
Subject (Ämne)	Format (Format)	Relation (Relation)
Description (Beskrivning)	Type (Typ av resurs)	Coverage (Täckning)
Publisher (Utgivare)	Identifier (Identifikator)	Rights (Rättigheter)

These are the most basic information a user would want to know about an object. Some are easy to fill in. Others like "coverage", which is supposed to indicate a chronology, jurisdiction or geography, can be harder to fill in.

The element *Description* is where teachers would find information on the educational features of a document, eg. for recommended age of target groups, but otherwise DC is too general to the school community. For light browsing of documents it fills its purpose though for students and teachers looking for materials in depth it might be too superficial.

An important feature of DC, and other metadata specifications too, is the possibility to qualify an element more. In the example above

"dc.creator.adress"

the element "creator" was extended to give information of the address to the creator, not only the name. This can also be done in areas where educational aspects not will be enough expressed in the 15 element.

A working group within the DC organization tries to find qualifiers that fulfill educational needs. However, the future seem also to bring new elements, such as *audience* ("category of user for whom the resource is intended") and *standard (*a reference to the education or training standard with which the resource is intended") [36]. In July 2002 educational *level*, replacing the more vague proposed term *audience,* was approved as an optional new element for educational resources but other proposed elements are still in the pipeline

DC is important as a base for other specifications such as IEEE (see section 3.4) but first we will consider other more strictly catalogues for library.

[36] http://dublincore.org/groups/education/ . Sutton/Mason (2001) and the proposal from Dublin Core Education Working Group (2000).

3.3 Library catalogues

Five years ago the library community did not look upon the web as of much use

"In short, the Net is not a digital library. But if it is to continue to grow and thrive as a new means of communication, something very much like traditional library services will be needed to organize, access and preserve networked information" [37]

Much web development since the quote above from 1997 works towards a digital library, even if some librarians have an uneasy feeling that the Semantic Web vision will falter[38]. Bibliographical metadata such as DC is new but of course libraries have used library index cards for centuries. The idea of mentioning them here is that they will be used when taxonomy is built on the tested book.

Concerning metadata technology, the value of library catalogues lies foremost in giving values to catalogue entries. Here we will look at two international classification catalogues (LC and Dewey) and a Swedish (SAB).

3.3.1 LC

The American *Library of Congress* (LC) has an established and detailed vocabulary for indexing books and other documents. Since 30 years LC also has made a computer based catalogue system, the MARC posts (Machine Readable Cataloguing)[39] which is used all over the world, including Sweden. This format is used in information retrieval of all kinds, but will not be considered here even though it is an important part of digital learning resources.

For purposes of testing tools for digital metadata in this thesis though, the index *Library of Congress Subject Headings* (LCSH) [40] is more important as well as the next classification.

3.3.2 DDC

The Dewey Decimal Classification (DDC) is another international standard for bibliographic classification using numbers from 0 to 9 to index various subjects. It is maintained by Library of Congress[41]

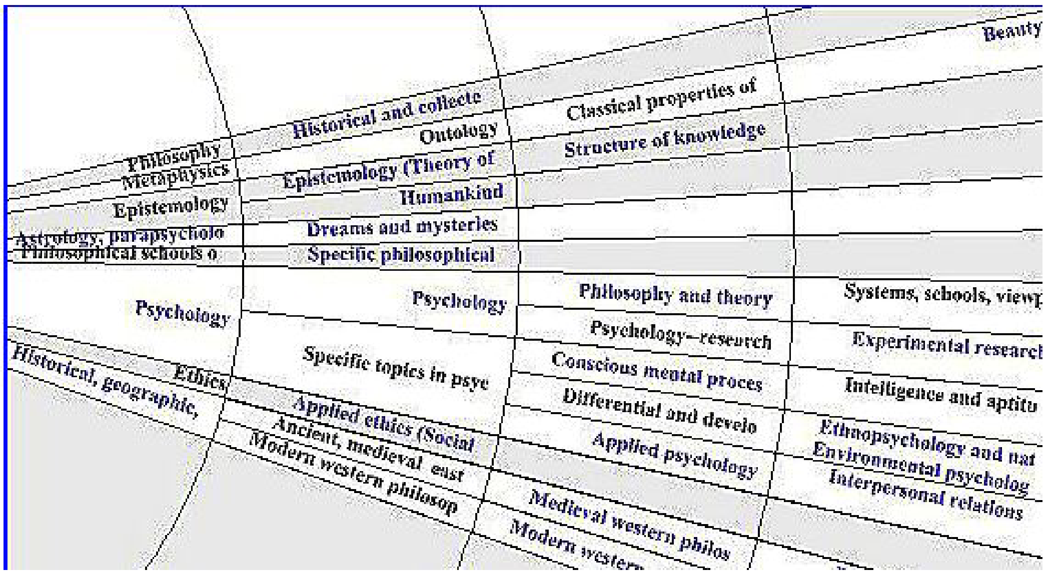
It is a system that works well towards the mapping to SAB and other library catalogues, although it is not as spread as LC. However, its recent use in the EU funded project Renardus aimed to develop a web-based subject gateway to European libraries and databases makes it interesting for educational purposes, albeit only in higher education and research.

---

[37] Quoted in Björkhem/Lindholm (2000), p. 18. See also http://www.cni.org
[38] Lu/ Dong/ Fotouhi (2002) and Brooks (2002).
[39] http://www.loc.gov/marc/marc.HTML . LC is expressed in XML and RDF, see http://www.loc.gov/standards/marcXML//
[40] http://www.kb.se/bus/lcsh.htm
[41] http://www.oclc.org/dewey/

Renardus points to resources that are stored in a local library or datasbase using a graphic interface where the DDC behind a user-friendly interface.

Here is a sample of philosophy:



Clicking the fields result in a presentation of the digital resources available at the collaborating academic institutions and libraries, indexed with DDC into ever more fine-grained numbers [42].

Dewey is useful to know but will not explicitly be mentioned in the test, although it would have been just as easy to work with that classification.

3.3.3 SAB

Swedish public and scientific libraries use a domestic classification code, SAB[43] (acronym for the *Svenska Allmänna Biblioteksföreningen, Swedish Public Library Association)*. The system uses letters where A is for library issues, B for general and cultural issues, C for religion and so forth until X which is music. For each letter there is subsections, e.g. *Eabpu* is the educational methodology for teaching with computers (where *E* stands for education, *ab* methodology and *Pu* for computers (and *Pu* itself is for computers under *P* which is for industry, technology and communications)).

[42] http://www.renadus.org. See Koch/Neuroth/Day (2001), Jacobs/Huxley(2002), Neuroth/Koch (2001) and Koch (2000). Other subject gateways are http://www.thegateway.org/, http://www.rdn.ac.uk and http://www.kb.se/kvalitet.htm. For technical aspects, see http://www.imesh.org/toolkit/

[43] For an overview in Swedish, see http://www.bibl.liu.se/sab/huvtswe.shtm

However this classification is being mapped towards LCSH in a recent Swedish project[44]. In the project new subject headings will be stored on a database that will ensure quick reference when needed. This service is of course available to schools, important since all school and university libraries in Sweden are built on SAB.

We leave the library classifications here for specifications committed to educational resources and management.

3.4 IEEE -LOM

Two international consortia, *Instructional Management Systems* (IMS[45]) and *Institute of Electrical and Electronics Engineers* (IEEE) with its group *Learning Technology Standards Committee* (LTSC[46]), have agreed on a metadata specification that regards education, learning and educational management, the IEEE-LOM standard. LOM stands for *Learning Object Metadata*.

The specification contains over 70 elements that are grouped into nine categories:

a) The *General* category groups the general information that describes the learning object as a whole.
b) The *Lifecycle* category groups the features related to the history and current state of this learning object and those who have affected this learning object during its evolution.
c) The *Meta-metadata* category groups information about this metadata record itself (rather than the learning object that this record describes).
d) The *Technical* category groups the technical requirements and characteristics of the learning object.
e) The *Educational* category groups the educational and pedagogic characteristics of the learning object.
f) The *Rights* category groups the intellectual property rights and conditions of use for the learning object.
g) The *Relation* category groups features that define the relationship between this learning object and other targeted learning objects.
h) The *Annotation* category provides comments on the educational use of the learning object and information on when and by whom the comments were created.
i) The *Classification* category describes where this learning object falls within a particular classification system[47]

The IEEE- LOM specification is the most widely used by commercial and educational content providers and institutions, even if not all elements are used for all objects, an impossible task. IMS itself is a commercial consortium with its own IMS standards

---

[44] http://www.amnesord.kb.se/
[45] http://www.imsglobal.org
[46] http://ltsc.ieee.org/
[47] http://ltsc.ieee.org/doc/wg12/LOM_WD6-1_1_without_tracking.doc

that implements the IEEE- LOM standard and makes it compatible but it is reversible in the sense that all IMS standards are compatible with the IEEE-LOM standards.

The specification is expressed both in XML and recently also in RDF[48]. This will extend the IEEE-LOM initiative further due the current large interest in RDF as a forceful Semantic Web technology.

3.5 EML

Educational Modelling Language (OU EML[49]) from Open University of Netherlands goes beyond IEEE-LOM and all other specifications. EML applies educational models and theories of learning by making them explicit on learning objects and by providing pedagogical roles for the users. The ambition is to provide a pedagogical framework for learning objects, not just a repository with no information on how to use the learning objects in specific learning situations, alone or in groups, with teacher/tutor or without. The metadata are being external and more related to describing the overall structure and not specific internal contents

The three main ideas of EML are:

1. Classification of learning objects in a semantic network derived from a pedagogical meta-model
2. Building a containing framework expressing the relationships between the classified learning objects.
3. Definition of the structure for the content and behavior of the different types of learning objects.

Well-known theories of instruction and learning will be used when tagging the resources. EML states three types of theories:

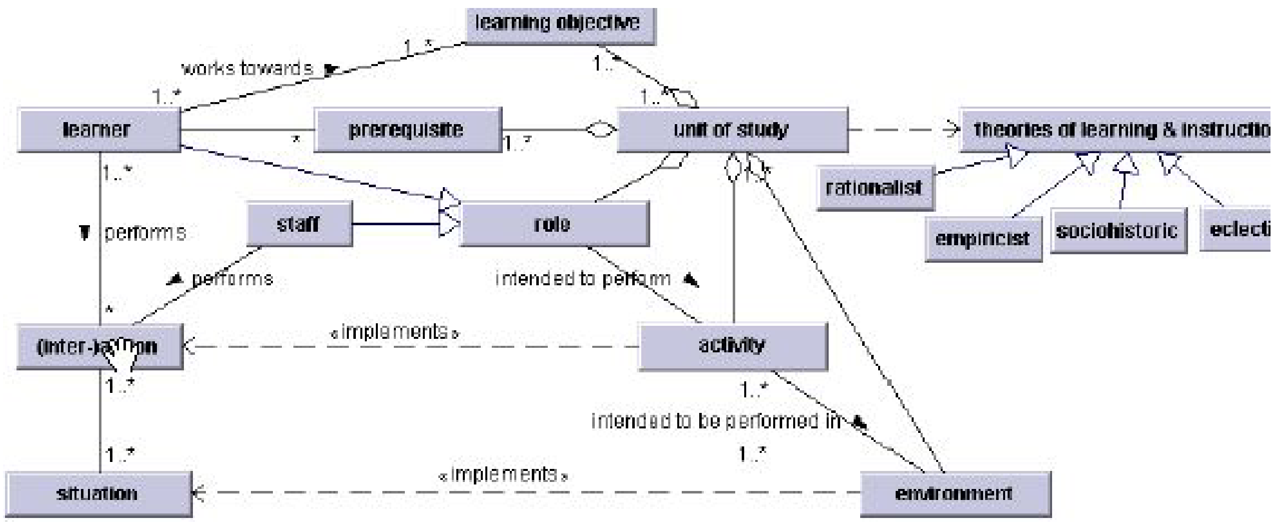
1. Empiricist (behaviorist)
2. Rationalist (cognitivist and constructivist)
3. Pragmatist- sociohistoricist (situated)

These types are distilled in a process that needs to be discussed seriously by educationalists, but leave them here as the EML group proposes them.

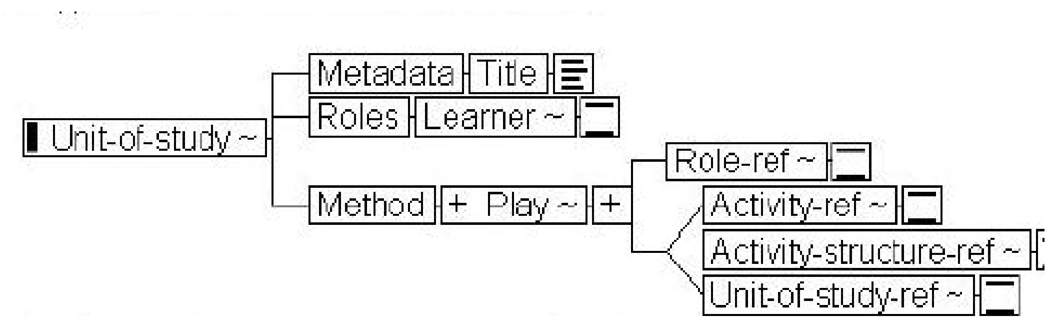
The various roles staff, students and groups may play in these theories are to be defined by the EML model.

The overall integrated model is viewed below[50]:

48 http://kmr.nada.kth.se/el/ims/metadata.HTML
49 http://eml.ou.nl and Sloep (2000)
50 Koper (2001) p. 14 and next at p. 22.



EML works as a pedagogical framework for metadata editing of all kinds of educational resources. It can point to the real sources without containing them. The smallest unit of study with no content is then the framework as seen here:



EML has recently been chosen by the IMS as providing a base for the newly established Learning Design specification[51] and seems to provide the educational context that some claim is lacking in current metadata specifications of learning objects. The discussions are just about to begin and have been neglected in the discussions on metadata that tend to emphasize technology and economy [52].

A main question for ideas such as EML is to what degree they lock users to their roles, as in a computer game where only there are certain characters. Another important area to explore critically is who is doing the tagging and marking of a certain pedagogical framework. If it is the subject specialists they may not be as

[51] http://www.cetis.ac.uk/content/20021008012855 , see also further CETIS news in 2002 that support EML as a coming educational metadata standard. The DC working group on education will also present new metadata for "pedagogy" in spring 2003. See note 38.
[52] Recker/Wiley (2000) and http://wiley.ed.usu.edu/writings2.pl, especially Wiley (2000b).

attentive to the pedagogical limits and drawbacks as educationalists could to be. The whole idea of sequencing in learning design is crucial here, as maintained by learning designer David Wiley. Essentially it concerns lining out a sequence of learning objects in a given fashion by the content designer. The discussion has just begun and we will return when considering some similar ideas in section 5.

3.6 TEI

Important and rich texts have always been subjected to detailed analysis. The Text Encoding Initiative (TEI[53]) was founded in 1987 to develop guidelines for encoding machine-readable texts of interest in the humanities and social sciences. Now TEI is available in XML. TEI uses over 400 different tags in order to show metadata for every paragraph, line, word and down to letters.

Here an example of an un-coded piece from the novel *Jane Eyre* by Charlotte Brontë:

> Chapter 38, page 474
> 'Have you, miss? Well, for sure!'
> A short time after she pursued, 'I seed you go out with the master, but I didn't know you were gone to church to be wed'; and she basted away. John, when I turned to him, was grinning from ear to ear.
> 'I telled Mary how it would be,' he said: 'I knew what Mr Ed-ward' (John was an old servant, and had known his master when he was the cadet of the house, therefore he often gave him his Christian name) -- 'I knew what Mr Edward would do; and I was certain he would not wait long either: and he's done right, for aught I know. I wish you joy, miss!' and he politely pulled his forelock.

The coded version:

```
<pb n='474'/>
<div1 type="chapter" n='38'>
/…/
p><q>Have you, miss? Well, for sure!</q></p
<p>A short time after she pursued, <q>I seed you go out with
the master, but I didn't know you were gone to church to be
wed</q>; and she basted away.  John, when I turned to him,
was grinning from ear to ear.  <q>I telled Mary how it would
be,</q> he said: <q>I knew what Mr Edward</q> (John was an
old servant, and had known his master when he was the cadet
of the house, therefore he often gave him his Christian
name) &mdash; <q>I knew what Mr Edward would do; and I was
certain he would not wait long either: and he's done right,
for aught I know.  I wish you joy, miss!</q> and he politely
pulled his forelock.</p>
```

The tags above, ie. "div1 type="chapter"n, pb n='474', p, q would be categorized in an editor that lets the user view and edit them in a more flexible way that with traditional HTML. A researcher might want to highlight all instances when a certain character in a novel speaks etc. If a viewer would like to only quotes an editor would highlight everything marked between <q> and </q>.

---

53 http://www.tei-c.org/

There is a light version that uses fewer tags, TEI Lite, which also will be used in this thesis. Materials being marked with TEI are primarily historical, linguistic, literary[54] or bibliographical.

The TEI metadata structure is embedded in the code and not very easy to get a grip of. Another disadvantage is that TEI marked text must be transforemd with style sheets shown with browsers. Editing a text with HTML gives a lot less alternatives but it can be viewed with all browsers. When finished though, TEI is useful when browsing with appropriate tools through large text archives, selecting the preferred elements. Other uses of TEI requires specific software.

There are other easier textual specifications that might be useful when coding texts on more technical objects, eg. DocBook[55].

3.7 Application profiles

All above-mentioned specifications are based on the idea of standards and adherence to common practices. In reality however, implementers do not have to, and do not, obey. There are too many considerations that are more important that loyalty to one standard when designing metadata applications. A solution is to choose between some given standards and combine them into what you need. The procedure and result is called "Application profiles"[56]. It is not entirely new. In the library community even "local fields" and various applications of MARC can be fitted to similar purposes.

In applying this idea to learning objects, one could use DC as basic metadata for the features of the objects, and then go on with relevant element form of existing specifications such as IEEE-LOM etc, drawing on their established spaces for names. In fact, the DC working group on education is working such an application profile[57] . Issues to explore further are semantic overlapping and parallelism, metadata syntax (RDF or XML).

This was the last of metadata specifications and technical aspects. Over to the test.

---

[54] See Wallberg (2000) for a comparision between TEI and HTML in coding the texts by CJL Almqvist and H Ibsen
[55] http://www.docbook.org
[56] See Heery/Patel (2000), Baker et al (2001), http://www.ariadne.ac.uk/ and http://www.schemas-forum.org/
[57] See note 38.

4 Test of digital editing of school text book

4.1. Overview of the test

The Swedish publishing house Ekelunds has generously submitted the preprint files to a course book in philosophy for Swedish secondary school, entitled *Filosofisk tanke* by Jan S. Andersson and Nils-Göran Mattsson (FT in the text below hereafter).

The files were in Quark Xpress (Mac version) and contained all necessary data for printing the book, but none about the metadata that could be useful to teachers and students if the book was made available in digital form. In the converting process to XML and XHTML formats over Indesign the original files became very loaded with additional useless code that severed the editing considerably.

However, the original idea of this whole thesis was to explore how using metadata editing tools that an interested could enhance the digital version but not completely computer skilled teacher would use. This turned out much more complicated than foreseen. The exploration of tools is therefore in focus in this section, not in the whole thesis.

Here is an outline of the available contents of the book that were tested with various metadata editing tools:

- A title page with necessary bibliographical information
- a table of contents
- a hyperlink to a chapter on 18th century enlightenment philosophy
- A preface from the authors with guidelines
- A chapter on the enlightenment philosophy with a hyperlink on a definition of the philosophy of George Berkeley
- Web link to philosophy for Swedish students
- Web link to Berkeley's philosophy

The digital material was put on a digital web portfolio in order to get the URLs. But the ambition to test navigation and construct a full exploration of the book and its digital resources were left undeveloped due to my lack of coding skills and the vast material.

4.2. Tools

The tools used to edit the texts and links were almost all in a development state and a part of the Semantic Web community with one exception, the commercially available metadata editor XML Spy. Some edited just texts and codes while others used graphics. There are more tools available but these six were selected due to their reputation among my colleagues.

The questions that were addressed in the test:

- what are the metadata editing functions ?

- what does the interface look like?
- how are files stored?
- what metadata specifications are being used?
- what would a full-fledged scenario with a total digital edition of FT and its digital resources look like? Pros and cons?

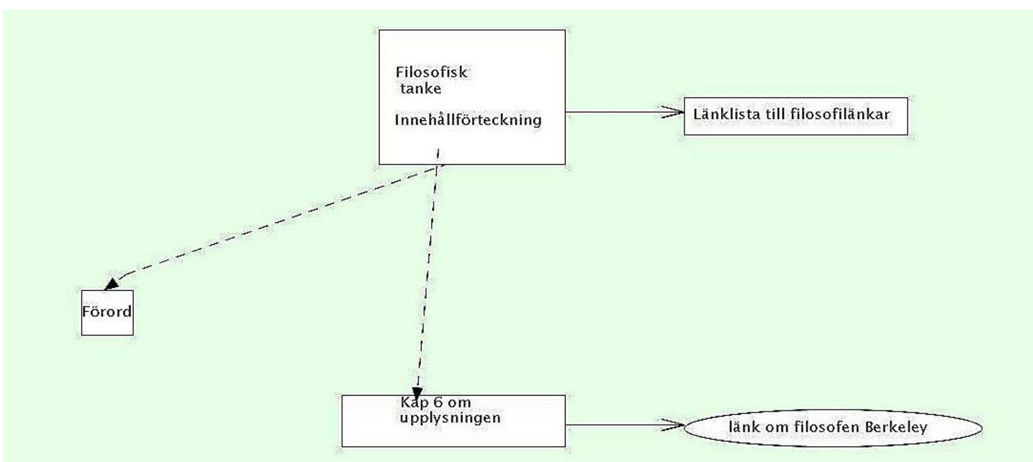
Not all questions were possible to answer for each tool but the main features are given and compared in the summary (section 4.9).

The first tool tried was the experimental conceptual browser Conzilla.

4.3. Conzilla

This conceptual browser written in Java is used to explore relations between concepts and items in a fashion that reminds of traditional mind maps. Conzilla uses the visual language UML (Unified Modelling Language) in order to express relations, classifications, associations etc[58].

Here the items from FT are represented by boxes and arrows that in themselves mean different things (associations, classifications etc):

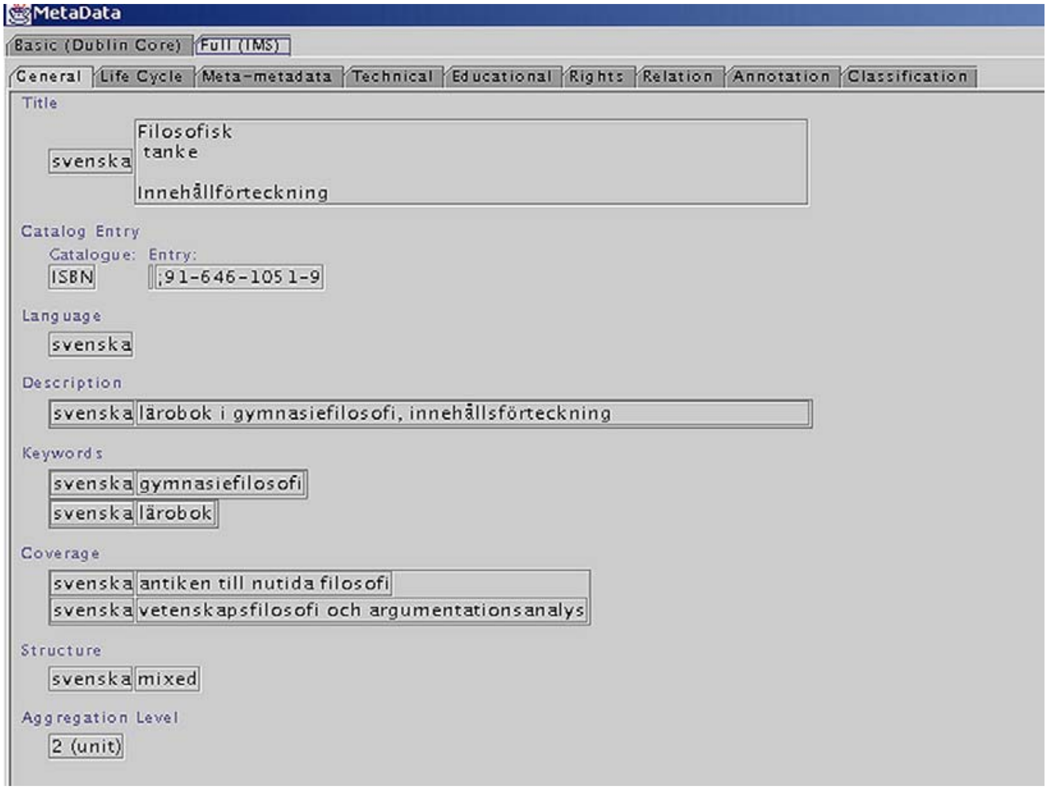


Conzilla uses smart functions for navigation but here the metadata functions are in focus. It has a metadata editor, Imse Vimse[59], built in its interface and uses it in when the function "Component editor" is selected. Two alternatives are available, "Basic" with Dublin Core entries and "FullÓ with IMS entries.

Here is a screenshot of the metadata-editing interface where entries are filled according to the file of the table of contents from FT:

---

[58] Naeve 2001a, Naeve 2001b, Nilsson 2002 and Nilsson/Palmér 1999 and http://www.conzilla.org
[59] ImseVimse will be considered separately in next section



These data can be accessed in a separate XML file, stored automatically on the hard disc. The metadata tags automatically created are in bold fonts:

```
<?XML version="1.0" encoding="ISO-8859-1" standalone="yes"?>
<Component VERSION="1.0">
  <MetaData>
    <record>
    <metametadata>
        <catalogentry>
        <catalogue>isbn</catalogue>
        <entry>
            <langstring>;91-646-1051-9</langstring>
          </entry>
        </catalogentry>
        <metadatascheme>IMS Lom 6.1</metadatascheme>
        <language>sv</language>
    </metametadata>
    <general>
        <title>
          <langstring lang="sv">Filosofisk
 tanke
Innehållförteckning</langstring>
    </title>
    <catalogentry>
        <catalogue>ISBN</catalogue>
        <entry>
            <langstring>;91-646-1051-9</langstring>
          </entry>
    </catalogentry>
    <language>sv</language>
```

```
    <description>
        <langstring lang=”sv”>lärobok i gymnasiefilosofi,
innehållsförteckning</langstring>
    </description>
    <keywords>
        <langstring lang=”sv”>gymnasiefilosofi</langstring>
****** (code ommitted)
  </MetaData>
  <Neuron TYPEURI=”urn:path:/org/conzilla/builtin/types/UML/ClassD/concept”/>
</Component>
```

All boxes and lines can be edited with metadata. That means that a student may follow all kinds of advices by following not only the arrows but also the metadata that is added to them. For instance, an arrow representing association points to a box representing a suggested web link. All items can have metadata.

Conzilla is still under development and it is not anything that a regular teacher would use for editing digital material. But the combined features of metadata editing and conceptual browsing are quite an extraordinary idea that has many advantages.

However, the metadata editor preselects only DC and IMS specifications and has no other options, which fits well with the educational community. The tool must be extended for the future though. Conzilla is available as an open source software tool and as such free to use.

Scenario with Conzilla:

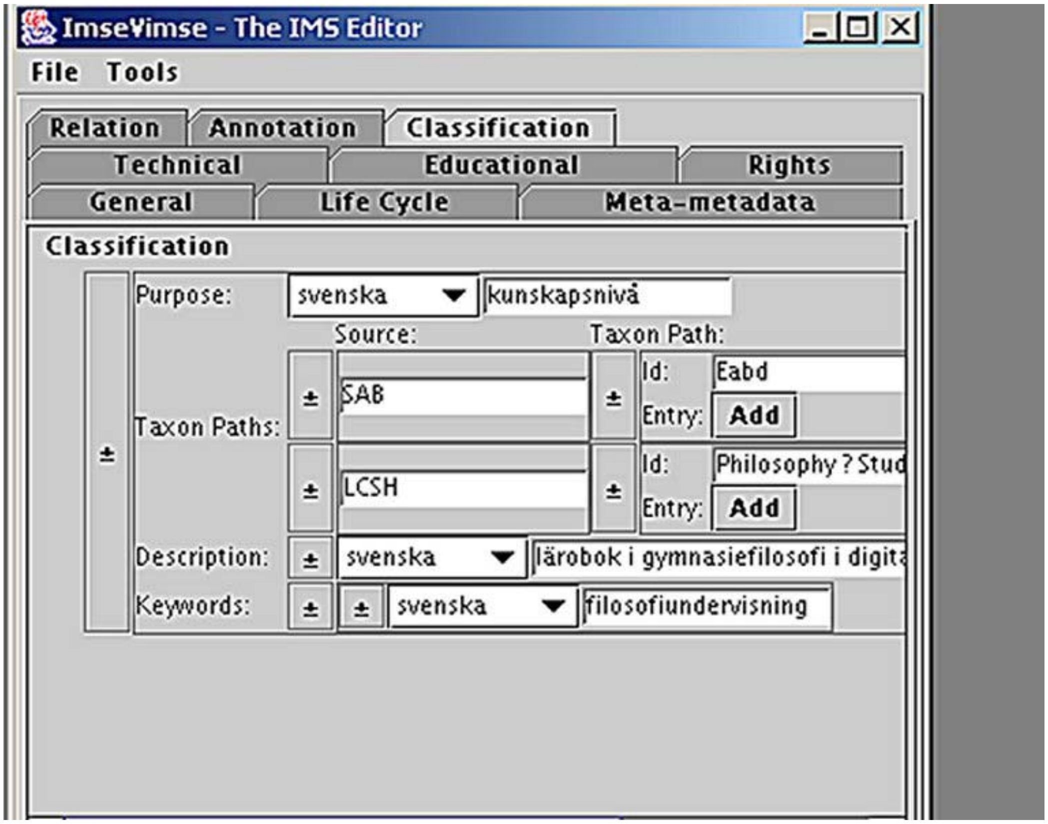
Philosophy with its reliance on logic, propositions and interferences is an ideal subject to is taught using a conceptual graphic tool such as Conzilla. The book's historical chapters would easily fit into the systematic chapters where possible. Areas such as philosophy of science would also benefit from using Conzilla. It is not hard to envision students browsing with Conzilla through the book, along with a digital and graphic version of its teaching guidance manual.

The separate metadata files would have to be used a Content Packaging editor in order to be used (see section 4.9). Since they are stored outside the document much work would be needed to fit them in the right slots in an editor. Also, all pointing into part of chapters using for example Xpointers[60] (or HTML anchors) from a conceptual representation may be an obstacle when dealing with such large amount of text, 580 pages (including the manual).

4.4 ImseVimse

As mentioned this metadata Java editor is built in Conzilla but functions also well alone. It edits only IMS metadata. This is a screen shot from the interface with data from FT.

[60] For X Pointer information, see http://www.w3c.org/XML/Linking



ImseVimse generates an XML file as well as Conzilla, but has also a function under "Tools" for previewing the file. Languages may also be selected[61]. ImseVimse is more user friendly than Conzilla due to its basic functions. There is not much to do that can go wrong. It is a nice little tool that fulfills its purpose well. ImseVimse is available as an open source software tool and as such free to use.

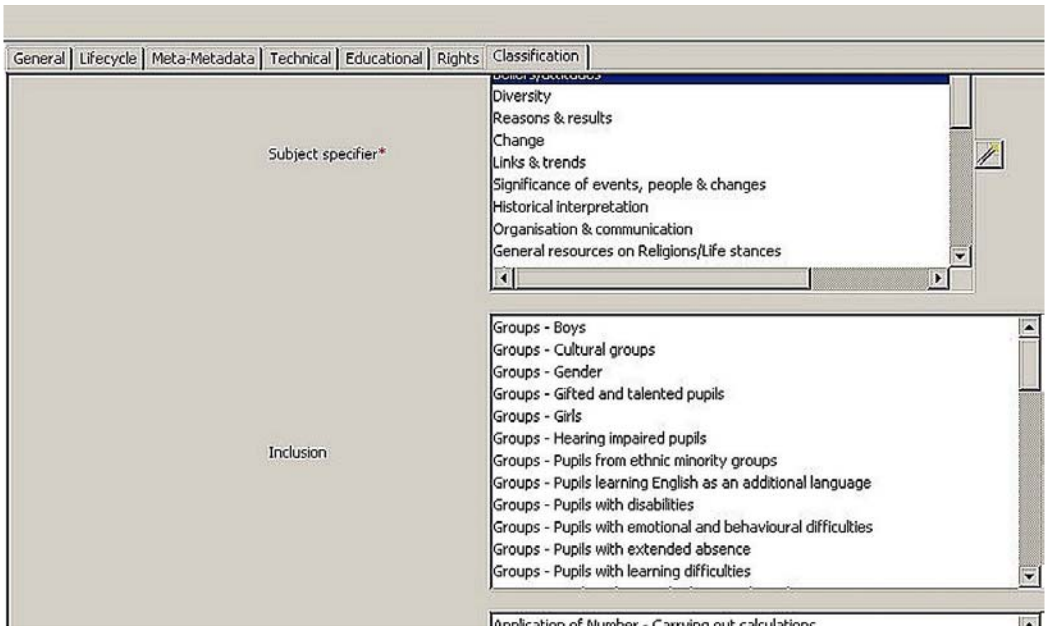
Scenario with ImseVimse:

The tool would produce as many metadata files in XML using IMS specifications as possible but nothing else. This is on the other hand something useful when the contents is edited in a content packaging editor. Much work is needed when a full-fledged digital version of FT exists in XML files. ImseVimse is maybe best used when editing chapters but not down to the details.

[61] For more information, see http://kmr.nada.kth.se/imsevimse/index.HTML.

4.5 Tagging Tool

The British initiative Curriculum Online[62] has launched a RDF editor called "Tagging Tool" written in Java. It is interesting even if it seems domestic in its taxonomy of students. The specification blends IEEE –LOM and DC with national classified slots like "For pupils with extended absence ", "For gifted and talented students" etc in the Inclusion class. Here is a screenshot of the classification-editing mode:



The created RDF files can be viewed but not edited with Tagging Tool. They are to be saved and assessed by another tool later. Tagging Tool creates also templates that can be reused. This function is very useful when dealing with similar learning object such as chapters and paragraphs in FT.

It edits only web-based resources, stored on ftp or http sites. One must be a member of the Curriculum Online community in order to use its ftp/http features. In that way the tagged learning object gets stored in their database. A nice feature is its ability to append resources. That is, two individual files containing data being combined. For example, two people might be inputting metadata simultaneously at different locations. Their work can be combined into one file for use in the search engine. . Tagging Tool is free to use after registration at the Curriculum Online website.

---

[62] For the Tagging tool see http://metadata.ngfl.gov.uk/taggingtool/ , for the Curriculum Online project see http://www.dfes.gov.uk/curriculumonline/index.sHTML and for metadata specifications, see http://www.dfes.gov.uk/curriculumonline/docs/COL%20vocabularies%20v1_5.pdf

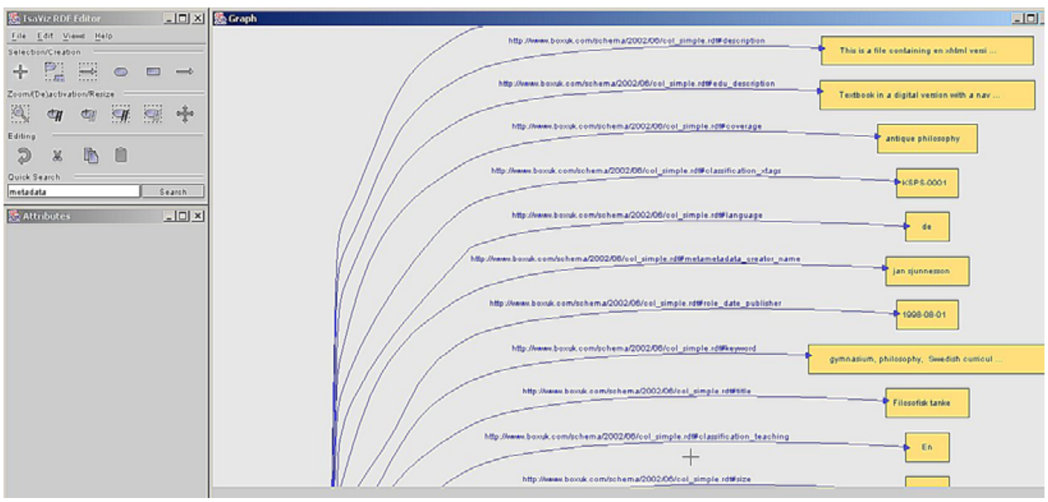
Scenario with Tagging Tool:

As with ImseVimse this tool would create a stock of files, these in RDF though, that would have to be edited and packed together with some Content Packaging tool (see section 4.9.).
FT is together with useful philosophy and study guidance links well suited for a full-fledged digitalization with Tagging Tool. The many alternatives for classification elements and vocabularies (although philosophy does not have its own tag yet) makes it an effective instrument for tagging all kinds of related resources. For lesser ambitions than classifying FT and its related digital resources ImseVimse will do with far less functions

4.6 IsaViz

IsaViz is a visual tool for browsing and authoring RDF models represented as graphs. It is written in Java but needs also Java Virtual Machine software. All available from the W3C consortium who developed the tool.[63]

Here is a screen shot of the RDF file that was created with Tagging Tool, representing elements and nodes in the Curriculum Online specification mentioned above.

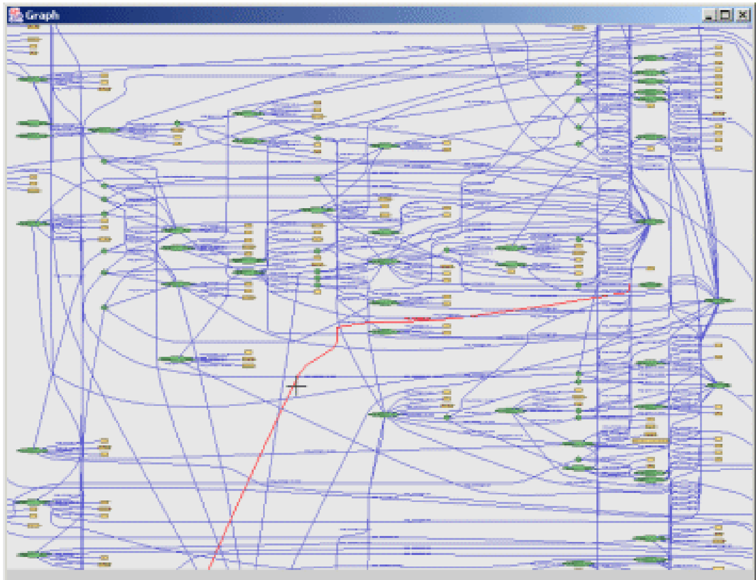


The idea is to get visual image of the structure of the files and in that mode also focus on them with graphic tools and edit. The various nodes, elements and namespaces appear to the left in an editor and can be changed. The editor can color the elements and nodes according to preferred functions. The RDF code can be viewed separately but not edited raw.

IsaViz's full potential is viewed in this second screenshot that does not involve any files from FT, but when all of its files are marked up they may appear something like this from a far in IsaViz:

---

63 http://www.w3.org/2001/11/IsaViz/ . The W3C consortium is the WWW standard organisation that governs and leads the web by developing common protocols that promote its evolution and ensure its interoperability.



Any specification can of course be used in IsaViz.

Scenario with IsaViz:

The graphical representation of FT with its resources in RDF files with IzaViz might be the simplest way to represent a digitized book. However, IsaViz is not a conceptual tool like Conzilla where a student actually can browse around for information using both actual files and web links. In IsaViz the RDF code is structured in an easy viewed way but this is for the editors and producers of the information, not the users.

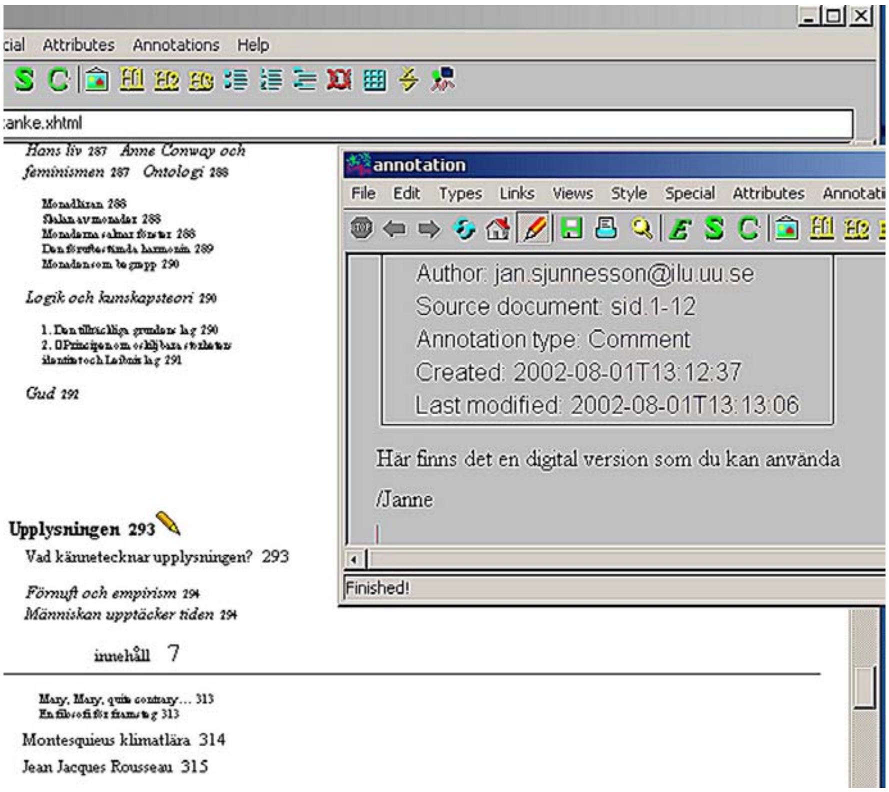
4.7 Annotea

This tool is quite different from the others. It is an annotation tool that does not use any metadata specifications but lets the user annote whatever with what comments he/she wants. Annotations mean comments, notes, explanations, or other types of external remarks that can be attached to any web document or a selected part of the document without actually needing to touch the document. When the user gets the document he or she can also load the annotations attached to it from a selected annotation server or several servers and see what his peer group thinks.

Annotea[64] is free software from the W3C written in Java and integrated in the experimental web browser Amaya, also from W3C[65]. Annotea uses RDF in order to make RDF- triples and can also parse the files showing all RDF/XML syntax errors

[64] http://www.w3.org/2001/Annotea/. See also Kahan et al (2001). Other annotation tools are OntoMat Annotazier and OntoAnnotate http://annotation.semanticweb.org . For annotation on subject gateways, see http://www.imesh.org/toolkit/work/components/annotation

Using Annotea was complicated since the annotations were supposed to be stored at a server account at W3C, but this did not work easily. A file from FT was annotated anyway though and is here shown with annotation. Do not miss the little yellow pen that indicates the annotation, which lets the annotation popping up when marked:



The technology behind Annotea is Xpointer. It allows for traversals of a document tree and choice of its internal parts based on various properties, such as element types, attribute values, character content, and relative position.

Amaya and Annotea are experimental tools that do not seem to be functional at the time of test (July 2002). Storing the annotations is not easy whether locally or on a server. The idea is that when a document with annotations is viewed the annotations should be loaded too, but this has not been the case yet.

Scenario with Annotea:

To use Annotea annotating FT when it is finished would be an ideal instrument for teaching with its ability to get a dialogue with students concerning the comments, done by students as well as teachers. Maybe even examination could be done with Annotea if each student had an account with passwords and was obliged to answer to

[65] http://www.w3.org/Amaya . Annotea works also on the browser Mozilla, but this is not tested here, see http://www.mozilla.org, and http://annozilla.mozdev.org/ . There is a version that should work on Explorer but it is rather unfinished at http://www.jibbering.com/snufkin/

the teacher's questions. All annotations would then be stored somewhere for the teacher to access and validate.

But Annotea would have to be combined with other editors if many resources are being marked, using DC and/or IMS-LOM specifications.

4.8. XML Spy

XML Spy[66] is commercial software that is not related to the ambitions of the Semantic Web community as the other tools. It edits XML documents internally and is a professional tool. The code below is edited inside the document with XML Spy but could be done with any other editor. This code shows the tags in another way than the other tools with externally stored metadata tags.

The code comes from previous work done by Hans Melkersson and Kjeld Jensen at Swedish Royal Institute of Technology who coded (except images) the whole chapter 6 on the Enlightenment with XML using the TEI-Lite as specification, which is not applicable to TEI itself [67].

```
<?XML version="1.0" encoding="iso-8859-1" standalone="yes"?>
<?cocoon-process type="xslt"?>
<!DOCTYPE TEI.2 SYSTEM "www.tei-c.org/Lite/DTD/teixlite.Dtd">
<?XML-stylesheet href="Fil-tank-xlite.xsl" type="text/xsl"?>
<div1 type="chapter" n="6"> <head type="rubrik">Upplysningen</head> <div2 type="avsnitt"> <head
type="underrubrik">

        Vad kännetecknar upplysningen?</head>
<p> Storbritannien och Frankrike var tillsammans den huvudsakliga drivkraften
bakom upplysningen. Tidsmässigt brukar upplysningstiden förläggas mellan den
engelska &quot;ärorika&quot; revolutionen 1689 och den franska revolutionen
1789, men redan under 1600-talet eller ännu tidigare fanns det en grogrund för
de idéer som kom att bli karakteristiska för 1700-talet. <list> <head>Typiska
upplysningsföreteelser är tex:</head> <item>Mekanistisk uppfattning om universum
men i slutskedet även organiska världsbilder</item> <item>Människans upptäckt
av tiden</item> <item>Skepticism och deism</item> <item>Befrielse från fördomar
och vidskepelse</item> <item>Empirism</item> <item>Determinism</item>
    <item>Liberalism</item>
<item>Vetenskapliga upptäckter och deras spridande genom vetenskapsakademier,
tidskrifter och Encyklopedin</item> <item>Den tidiga upplysningens tro på framsteg
för samtiden, vilken den sena upplysningen ersätter med tron på framsteg för
framtiden</item> <item>Lycka eller välfärd som målsättning för samhällslivet
Tanken att naturen är till för människans &quot;nytta&quot;.</item> </list>
</p>
<pb/> <div3> <head>Förnuft och empirism</head>
<p>Upplysningen kännetecknas bl a av ett kraftigt avståndstagande från de traditionella
 auktoriteterna och en stark tro på förnuftet som vetenskaplig metod. Denna metod
```

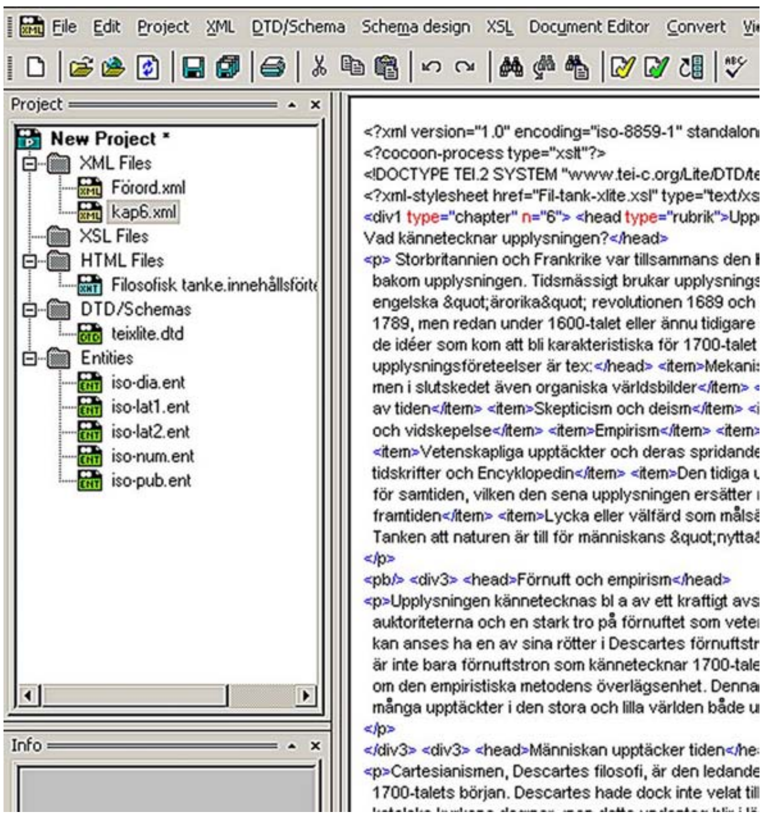
66 http://www.XMLspy.com

67 http://www.tei-c.org/Lite/ . For more content design in the humanties done at CID, see http://cid.nada.kth.se/il/id/default.HTML

kan anses ha en av sina rötter i Descartes förnuftstro (rationalism). Men det
är inte bara förnuftstron som kännetecknar 1700-talet, utan många är också övertygade
om den empiristiska metodens överlägsenhet. Denna hade nämligen lett fram till
många upptäckter i den stora och lilla världen både under 1600-talet och 1700-talet.
</p>
</div3> <div3> <head>Människan upptäcker tiden</head>

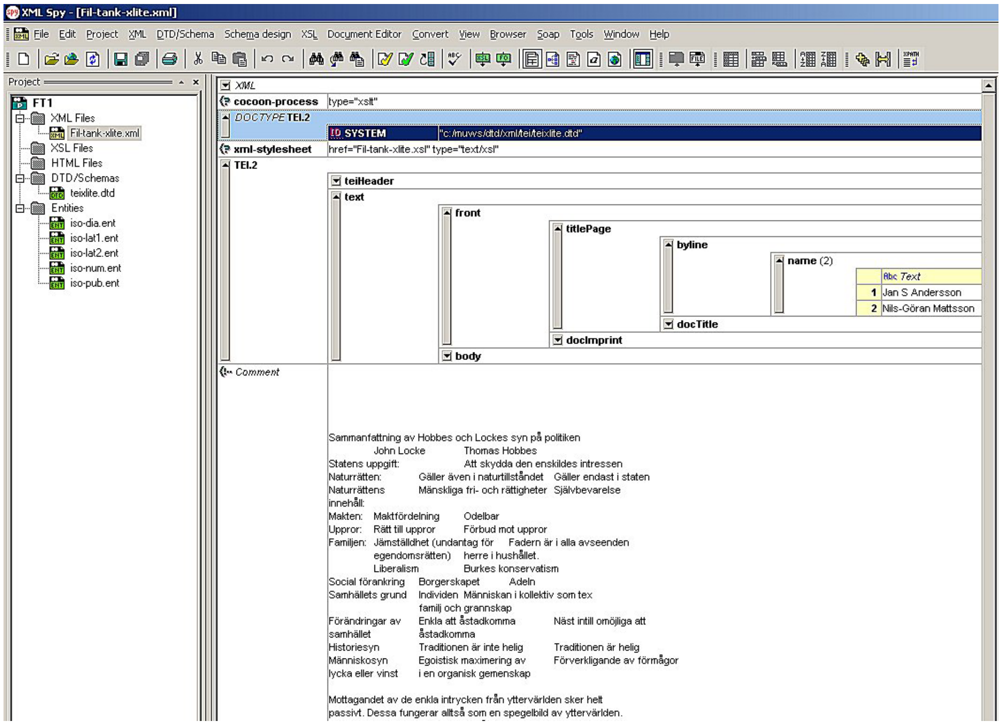
As can be seen here, the tags being used are those that divides paragraphs, lists items, headlines etc. The specification TEI is designed as mentioned for archived material that needs special attention to fonts, textual details, diacritics etc.
Of course any other specification can be used with XML Spy, but here the work is already done with TEI so that is why this specification is being discussed.

This screenshot shows the TEI - specification, files and special files for entities (codes that supports unknown letters like the Swedish å, ä, ö):



XML Spy has features that extend far beyond the previous free tools. One can list the elements and attributes in separate windows and the tool suggest what items may be used.

Here a screenshot that shows the contents of the file as represented by its XML tags:



XML Spy has many editing functions, too many to mention here[68]. Important is that the tool validates the XML file that must be correct in all details and also checks its well formedness towards the chosen specification, stored locally or on a URL. XML Spy is therefore also an XML parser.

Another important function is the possibilities to view files in text, browser and schema editing mode. Although XML Spy is a commercial tool it still does not work as easy as one could hope for. One has still to know a lot about encoding and finding errors.

Scenario with XML Spy:

XML Spy would probably be good to edit all of FT into its paragraphs and elements but would generate an enormous amount of code inside the file. Of course the file could be divided but all files containing the code inside them would be quite large

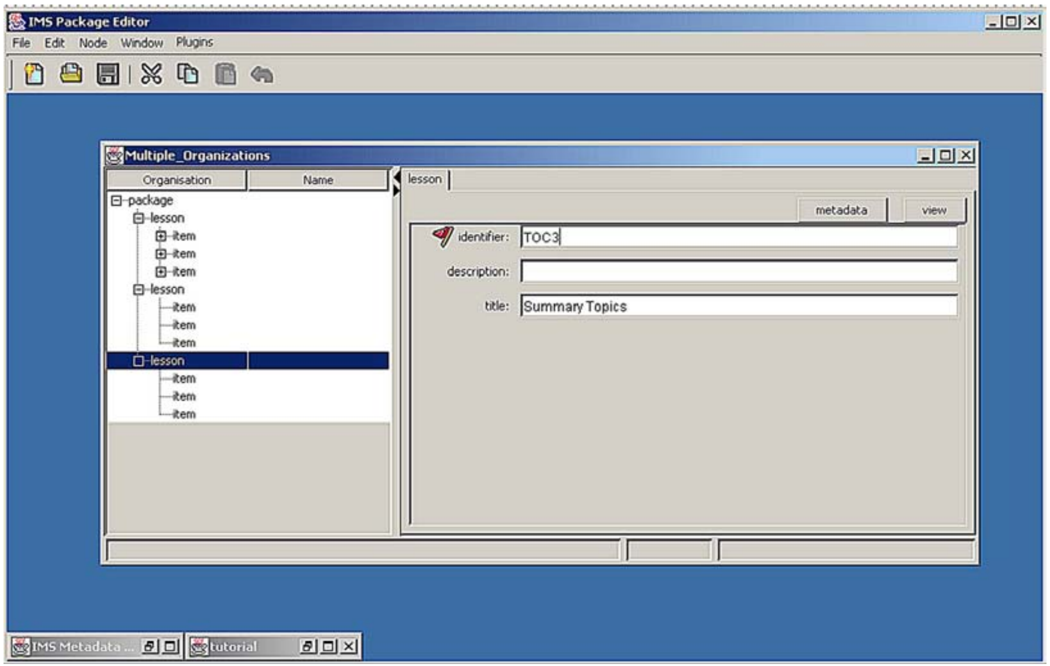
For all of FT it would take enormous effort to encode all its paragraphs and textual features. An advantage is done is that all elements of a certain kind would show up in a window if one chooses to view them.

[68] Östlund /Hermundstad (2001) has many examples of XML Spy functions.

4.9. Content packaging

None of the tools will be enough when it comes to present digital material from a bundle of files if they are to be exported into a larger learning management system. With more tools that let the metadata tagged learning objects be fitted into classes, lessons aggregate all kinds of learning items. To achieve this they need to be packed in a specific way using a Content Packaging editor. A sample of such a tool is shown below[69]

This editor, EC-Pac, uses besides the metadata specifications IMS and DC also Canadian Core and UK National Learning Network. Here is a screenshot from the interface with its mode of aggregate lessons etc:



Content Packaging has its own specifications done by the IMS and seems to be invaluable for large scale and commercial applications of learning management software. The EC-Pac editor supports Macromedia's Dreamweaver, Microsoft's LRN Toolkit, Blackboard, WebCT and many more.

4.10 Summary of the test

Testing tools was interesting but more complicated than foreseen. Without supervision from kind programmers at Uppsala Learning Lab the whole project could not even be started. An ordinary Swedish teacher at any level would have to get the same support, which is at this time unlikely. Maybe computer engineers or

[69] See http://www.met.ed.ac.uk/pac-man/editor/navig.HTML

programmers working at an office for learning material in a county or a national office would be the right people in the near future to work with this kind of metadata editing. Content providers too of course would make the effort to put in metadata if they saw an advantage and some do it already. But the question here was to see how an ordinary teacher would handle the tools and it seems like a long way to go yet.

All tools except XML spy edits metadata outside the file. A disadvantage with this is the problem of pointing down in a text. For instance, the chapter on the Enlightment has a section on George Berkeley. With these tools one must make HTML – anchors to select this section or, which Annotea does automatically, using Xpointer.

For an unexperienced teacher interested in metadata technology, the simple ImseVimse would probably be the best solution to tag metadata to small learning objects, leaving whole course books aside. But then again XML Spy with its parser functions, large editing features is the best if someone wants to work on the whole piece.

What struck me when trying out the tools is that there are so much many stages and code languages to go through and learn to present a digital version, even when one has produced a well-formed XML code[70].

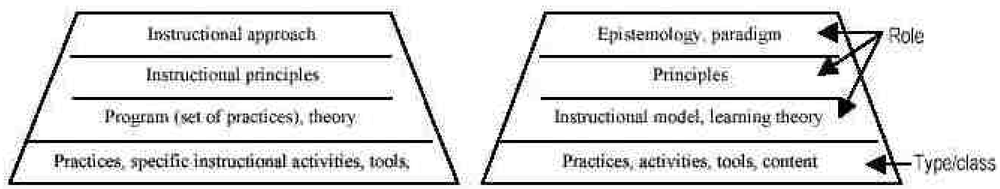
The whole idea of fitting objects into larger sequences of learning instances, e.g. lessons, classes, preparatory test, is crucial to the discussions on what theories and idea of learning and education one has. Within the EML community it would be natural to use the tools together with Content Packaging, but there are other views too, which will be seen in the next section.

[70] In order to present the XML files, one has to learn CSS or XSL. To make URI nodes down in the text, one has to know Xpointer and Xpath

5 Extended educational metadata

As mentioned in section 1.1 and 3.4. , there has been lately an interest metadata editing for purposes of learning and instructional design (Wiley 2000a, Recker/Wiley 2000, EML etc). The used tools were not tested for any of those purposes but some of them could handle these data easily, e.g. Conzilla, Annotea. Content Packaging is also crucial here. But what learning design methods to use and for what purpose?

In Germany researchers at the Learning Lab Lower Saxony has contributed to the discussion around the uses of instructional and learning design on learning objects and metadata[71]. The sheer inability in the IMS- LOM system to contain metadata on learning design is acknowledged to be too superficial for students in order to choose between digital resources. In a hierarchy of four pedagogical dimensions and its abstraction layers the IMS-LOM entities is said to address only the lowest, most practical, detailed level. Decisions on the top level affect all level from above. The German concept of Bildung is referred as a paradigm for teaching as a reflective practice.



Didaktische Dimensionen nach Merrill Vorschlag für die Integration von Rollen

A model with "instructional roles" is then proposed that enable users to choose among two different educational theoretical approaches, i.e. Problembased Learning and Ausubel's cognitivism[72]. With supporting metadata schemes each for "Roles according to Ausubel" and "Roles according to PBL" an extended LOM meta-model is lined out. The learning objects may stay the same but roles differ if made explicit.

Another approach from the German team concerns not the roles users may choose when navigating among metadata schemas, but the skills and competencies acquired using the learning object. This is the tradition in knowledge management where staff is required to upgrade themselves to certain levels, diplomas and standards. In another thesis[73] the L3S group asked the teaching staff when they demanded of successful students using the Open Learning Repository (OLR).

The work by L3S is not finished yet but aim to propose new metadata standards to annotate learning environments designed for competency – based learning. An evaluation of the third version of OLR system gives that it is not always the case that

---

[71] http://www.learninglab.de . The papers are Allert/Dhraief/Nejdl (2001), Allert/Richter (2001a and 2001 b), Allert et al (2002), Allert/ Qu/Nejdl(2002*)* and Allert/Richter/Nejdl (2002). An example of learning environment though not with theories of learning explicit but with user profiles is WHURLE, see Brailsford et al (2002)
[72] Allert/Dhraief/Nejdl (2001)
[73] Allert/ Richter(2002b)

use, which is proposed by the designers, is actually understood and used in an appropriate way by the users[74].

There are many ideas in digital learning management that seems to be discussed with few educationalists and researchers in pedagogy being present. The German team opens up for academics to contribute in this area. Analyzing the process from making a learning object in a context to distributing it in an isolated manner to someone unpacking it in another context is not new in educational research. Many similar studies has been done of the production of textbooks and their use in school contexts.

A new area opens up here by investigating to what extent it is possible to re-contextualize learning objects and actually get knowledge, not just information. A fear I share with David Wiley is that de-contextualized learning objects will just become something like clip art objects, silly readymade pictures to invitations, a kind *of Clip Art Instruction* (CAI) instead of a desire*d Computer Assembled Instruction*[75].

Wiley and L3S point to weaknesses of current metadata learning systems that in part can be resolved without breaking up the old IEEE- LOM system. With enough emphasis on architectural thinking and time for development the visions of an extended metadata structure for learning objects and resources will evolve.

[74] Allert et al (2002). IMS' Learning Information Package (LIP) is also moving towards providing metadata of goals, competencies, skills etc, see http://www.imsproject.org/profiles/lipbest01a.html
For critical views on knowledge management in academia, see Fuller (2000), Fuller (2001) and Day (2001)
[75] Wiley (1999), p. 6..

6 Summary

Looking back on the road traveled it seems like a huge task to summarize all these loose ends and points made in various discussions. Nevertheless, a main point has been to show that the area of digital learning objects and their management are too important to be left to computer industries and science. My hope is that this thesis has shown new research tasks to educationalists, teachers and students that need their knowledge and participation in order to make progress, academically and in delivering technical solutions.

Another point was to collect as much information as possible on the state-of-the-art in educational technology concerning metadata and the Semantic Web. If this point as been too fragmented it is in part due to my lack of putting it all together, in another part due to the very speed innovations and initiatives proliferate. Nobody, at least not me but neither Tim Berners-Lee even nor some top executive of some consortium is able to foresee the future, not to speak of lead it. It seems like the initial idea of the internet, dating back in early 1990s when information wanted and was free, is behind this next step in internet technology. Even though some major corporate players are in the game the future of learning technology is more in the hands of inventors and users. The Semantic Web vision along with peer-to-peer and open source technology are reviving these open and free ideas, cracking up some institutions including schools and classrooms with new exciting technology. Sharing that vision has been one main point in this thesis.

The test that was performed became a lesser part of the thesis than predicted but some lessons will be learned from it. One lesson is that such a project I started, trying out tools when editing a digital book, is not a one man's job, it is for team of programmers and educational specialists. In fact, the idea is still available to try a full-scale version of my preliminary test. Another lesson is that many tools, coding procedures and technical questions are not yet solved with the new Semantic Web technology. Beta-versions abound, and beta- beta. Also, all mentioned metadata standards are structured in so many layers and for so many purposes that they need attention to from so many areas that very few people actually can cover.

A lesson learned late in the work in this thesis was when I spoke with the director of the Swedish Association of Educational Publishers, Jerker Fransson. His concern was that question of copyright must be cleared out before any digtitalized textbooks would become available to schools and students. His attention was more of using digital management to protect the objects in an efficient way, maybe with a DOI (Digital Object Identifier[76]). Also, the time and money it takes marking learning objects with metadata need a larger market than a Swedish speaking audience does. Stina Degerstedt, librarian at the Royal Library, Stockholm, was also concerned with copyright but the library had agreements with some publishers in order to catalogue electronic material in a safe way.

[76] Digital Object Identification (2002)

These late considerations does not leave me without hope for a time when learning objects are free, easy assessed and well- tagged with lots of useful information. It just may take more time than a summer month.

7 References

Allert, Heidrun/ Dhraief, Hadhami/ Nejdl, Wolfgang (2001) *How are Learning Objects used in Learning Processes. Instructional Roles of Learning Objects in LOM* , Learning Lab Lower Saxony. (Hannover: University of Hannover).

Allert, Heidrun/ Richter, Christoph (2002a) *Learning Object and the Semantic Web. Explicitly Modelling Instructional Theories and Paradigms.* Learning Lab Lower Saxony. (Hannover: University of Hannover)

Allert, Heidrun/ Richter, Christoph (2002b) (*Re-) designing an Educational setting . Trails of Competency in an Open Learning Repository.* Learning Lab Lower Saxony. (Hannover: University of Hannover)

Allert, Heidrun et al (2002) *Instructional Models and Scenarios for an Open Learning Reposity.*. Learning Lab Lower Saxony. (Hannover: University of Hannover)

Allert, Hedrun/Richter, Christop/ Nejdl, Wolfgang (2002*) Context specific metadata for learning.* Learning Lab Lower Saxony. Hannover: University of Hannover)

Allert, Heidrun / Qu, Changtao/ Nejdl, Wolfgang (2002*): Ein theoretischer Ansatz zur Rolle der Didaktik in Metadaten* . Learning Lab Lower Saxony. Hannover: University of Hannover)

Almerud, Peter (2000) “Elektronisk publicering och distribution av litteratur – ett förslag”. Final report in *Nya vägar för boken*. (Stockholm: Royal Library) . Available at http://www.kb.se/Nvb/forslag.htm (03-05-06)

Andersson, Jan S / Mattsson, Nils Göran (1998**)** *Filosofisk tanke.* (Solna: Ekelunds) (abbreviated FT)

Andersson, Jan S / Mattsson, Nils Göran (2001) *Filosofisk tanke. Lärar- och elevhandledning* **(**Solna: Ekelunds)

Aßmann Uwe (2001a) *Countdown for a new Industrial sector. Why Sweden should enter the Semantic Web movement now* . (Linköping: Dept for Software Engineering, Linköping University.)

Aßmann Uwe (2001b) *The Swedish Semantic Web initiative* (Linköping: Dept for Software Engineering, Linköping University.)

Baker, Thomas et al (2001) ”What term does your metadata use ? ” *in Journal of Digital Information,* vol 2 (2) . Available at http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Baker (03-05-06)

Bechhofer, Sean et al (2000) *An informal description of Standard OIL and Instance OIL.* Available at http://www.ontoknowledge.org/oil (03-05-06)

Benyon, David./Höök, Kristina. (1997) "Navigation in Information Spaces: supporting the individual" in *Human-Computer Interaction*: INTERACT'97, S. Howard, J. Hammond & G. Lindgaard (editors), pp. 39 - 46, (Oxford: Chapman & Hall)

Besser, Howard (2002) "The Next Stage: Moving from Isolated Digital Collections to Interoperable Digital Libraries", in *First Monday,* vol 7 (6), Available at: http://firstmonday.org/issues/issue7_6/besser/index.HTML (03-05-06)

Berners-Lee, Tim (1999) *Weaving the Web* (New York: Harper-Collins).

Berners-Lee, Tim et al (2001) "The Semantic Web" in *Scientific American*, May 17. Available at http://www.semanticweb.org/resources.html#publications (03-05-06)

Berners-Lee, Tim (2002) *"The World Wide Web – Past, present and future".,* Japan Prize Commemorative Lecture 2002. Available at: www.w3c.org/2002/04/Japan/Lecture.HTML (03-05-06)

Björkhem, Miriam/ Lindholm, Jessica (2000) *Metadata för det digitala biblioteket.* Maigsteruppsats. (Lund: Biblioteks och informationsvetenskap , Lunds universitet). Available at http://www.kult.lu.se/bivil/fulltext00/2000-7.pdf (03-05-06)

Brailsford, T.J. et al (2002) "*Autonavigation, Links and Narrative in an Adaptive Web-Based Integrated Learning Environment".* Proceedings from the WWW2002 conference, Hawaii. Available at http://www2002.org/ (03-05-06)

Bray, Tim (2001) "What is RDF?". Available at http://www.xml.com/pub/a/2001/01/24/rdf.html (03-05-06)

Broady, Donald (2001) "Digitala arkiv och portföljer", in IT-kommissionen (2001)

Brooks, T.A (2002) "The Semantic Web, universialist ambitiion and some lessons from librarianship", in *Information Research* no 7 , vol 4

Business Week Online (2001) *"A library to end all libraries"*. July 23

Dahlström, Mats, "En nod att stilla bedja om " in *HumanIT*, no 3, 1998, (Borås: Bibliotekshögskolan, Högskolan i Borås). Available at http://www.hb.se/bhs/ith/3-98/md.htm (03-05-06)

Day, Ronald (2001) "Totality and Representation: A History of Knowledge Management Through European Documentation, Critical Modernity, and Post-Fordism", *Journal of the American Society for Information Science and Technology*, no 9. vol 52

Decker, Stefan et al (2000*) The Semantic Web – on the respective roles of XML and RDF* (Palo Alto: Dept of computer science, Stanford university.)

Dhraief, Hadhami/ Nejdl, Wolfgang/ Wolf, Boris/ Wolpers, Martin (2001) *Open Learning Repositories and Metadata Modelling*. Learning Lab Lower Saxony (L3S). Hannover: University of Hannover)

Dieberger, Andreas et al (2000) *Social Navigation: Techniques for building more usable systems ,* in *interactions*, Nov-Dec issue, ACM, 2000

Digital Object Identification (2002) *The DOI handbook* (Oxford: International DOI Foundation, Inc). Available at http://www.doi.org/hb.html (03-05-06)

Dolphin, Ian/Miller, Paul (2002) "Learning Objects and the Information environment " in *Ariadne* no 32. Availble at http://www.ariadne.ac.uk/issue32/iconex/ (03-05-06)

Downes, Stephen (2001). "Learning Objects: Resources For Distance Education Worldwide" in *International Review of Research in Open and Distance Learning:* Vol 2, no 1

Dublin Core Education Working Group (2000) *"DC- education proposal"*. Issued by http://dublincore.org/groups/education

Fensel, Dieter et al (eds) (2003) *Spinning the Semantic Web* (Cambride: MIT press)

Frohmann , B (2000) "Cyber- ethics: Bodies or bytes ?" *International Information and Library Review*, August 2000.

Frohmann, B. (2001). "Discourse and documentation: some implications for pedagogy and research". *Journal of Education for Library and Information Science*, vol 42 no 1

Frängsmyr, Tore , ed (2001) *The structure of knowledge since the renaissance* (Berkeley: University of California at Berkeley press).

Fuller, Steve (2001) "KNOWLEDGE R.I.P ? Resurrecting knowledge requires rediscovering the university", *Journal of Critical Postmodern Organization Science*, vol 1, no 1

Fuller, Steve (2000) *A critical guide to knowledge society newspeak*. Proceedings from the conference "Knowledge management: Concepts and controversies" at University of Warwick, feb 2000. Available at http://bprc.warwick.ac.uk/thesiss.HTML (03-05-06)

Gustavsson, Benny (2001a) "Vad är den Semantiska Webben ." Available at http://home.swipnet.se/semanticweb/swe/200110/intro_kort.html (03-05-06)

Gustavsson, Benny (2001b) *On the Semantic Web language.* Reports from the MSI. No 01065 (Växjö; School of Mathematics and Systems Engineering, Växjö University)

Handschuh; Siegfried/ Staab, Steffen (2001*) Authoring and Annotations of web pages in CREAM.* (AIFB, University of Karlsruhe, Learning Lab Lower Saxony, University of Hannover, and Ontoprise GmbH, Karlsruhe.)

Hedberg, Sten (1999a) *Metadata- Kataloginformation på Internet. Report to Nya vägar för boken (*Stockholm: Royal library).
Available at http://www.kb.se/Nvb/Katalog/kat1.htm (03-05-06)

Hedberg, Sten (1999b) ”Dublin Core i Sverige” in *HumanIT* no 4 1999 (Borås: Bibliotekshögskolan, Högskolan i Borås) Available at http://www.hb.se/bhs/ith/4-99/sh.htm (03-05-06)

Heery, Rachel / Wagner , Harry, "A Metadata Registry for the Semantic Web" in *D-Lib Magazine* May 2002 Vol 8 (5). Available at http://www.dlib.org/dlib/may02/wagner/05wagner.HTML (03-05-06)

Heery, Rachel/ Patel, Manjula (2000) "Application profiles" in *Ariadne* no 25. Available at http://www.ariadne.ac.uk/issue25/app-profiles/ (03-05-06)

Höök, Kristina et al.(2000). "Individual Differences in Social Navigation". In *Proceedings of the Workshop on Social Navigation: a design approach? (CHI2000)*, Hague, Holland, April. ACM.

Iselid, Lars (2001) "Den semantiska webben", in *Datormagazin* No 8. (Stockholm)

IT-Kommissionen (2001a) *IT i skolan. Rapport nr 45.*
(Stockholm: IT- kommissionen,) Available at http://www.itkommissionen.se (03-05-06)

IT-Kommissionen (2001b) *Vårt digitala tjänstesamhälle – vision 2011* + (Stockholm: IT-kommissionen) Available at http://www.itkommissionen.se (03-05-06)

IT-Kommissionen (2002a) *IT- samhällets mjuka infrastruktur. Rapport nr 46.*
(Stockholm: IT- kommissionen,) Available at http://www.itkommissionen.se (03-05-06)

IT-Kommissionen (2002b) *Breddtjänster – ett nytt skede i IT- politiken.* Delbetänkande. SOU 2002:51 (Stockhom: Regeringskansliet) Available at http://naring.regeringen.se (03-05-06)

IT-Kommissionen (2002c) *Hänger högskolan med i Internet ?. Rapport nr 47.*
(Stockholm: IT- kommissionen,) Available at http://www.itkommissionen.se (03-05-06)

Jacobsen, P. (2001). "Reusable learning objects - What does the future hold?" in *E-learning Magazine*.

Kahan, José et al (2001) *Annotea: An Open RDF instrastructure for shared web annotations.* Proceedings of the WWW10 International Conference, May 2001, Hong Kong. Available at http://www.w3.org/2001/Annotea/Thesiss/www10/annotea-www10.HTML (03-05-06)

Kunze, Tobias/ Brase, Jan/ Nejdl, Wolfgang (2002) *Editing Learning Object Metadata. Schema driven input of RDF metadata with the OLR3 – editor* (Hannover: Learning Lab Lower Saxony and KBS institute ,University of Hannover)

Koch, Traugott/ Neuroth, Heike/Day, Michael (2001) *Renardus: Cross-browsing European subject gateways via a common classification system* (DDC). Available at http://www.lub.lu.se/netlab/staff/koch.html (03-05-06)

Koch, Traugott (2000) "Quality- controlled subject gateways" *in Online information review.* Vol 24, no 1. Available at http://www.lub.lu.se/netlab/staff/koch.html (03-05-06)

Koper, Rob (2001) *Modelling units of study from a pedagogical perspective* (Heerlen: Open University of Netherlands). Available at http://eml.ou.nl/introduction/docs/ped-metamodel.pdf (03-05-06)

Lannerö, Per (2002) *Att utnyttja IT effektivt - Vitbok om infostruktur* (Stockholm: Skolverket) Available at http://www.skolutveckling.se/skolnet/smultron/infostruktur/ (03-05-06)

Liedman, Sven-Eric (2001*) Ett oändligt äventyr* (Stockholm: Bonniers).

Lu, Shiyong/ Dong, Ming/ Fotouhi, Farshad (2002) "The Semantic Web: Opportunities, challenges" in *Information Research*, no 7, vol 4.

Maedeche, Alexander (2002) *Ontology learning for the Semantic Web* (Dordrecht: Kluwer)

Miller, Eric (2001) "An introduction to RDF" in *D- Lib Magazine* May 1998. Available at http://www.dlib.org/dlib/may98/miller/05miller.html (03-05-06)

Miller, Libby (2002) *RDF query by example.* (Bristol: Institute for Learning and Research Technology , Bristol University). Available at http://ilrt.org/discovery/2002/04/query/ (03-05-06)

Naeve, Ambjörn. (2001*), The Concept Browser - a new form of Knowledge Management Tool,* Proceedings of the 2 nd European Web-based Learning Environments Conference (WBLE 2001, Lund, Sweden). Available at http://kmr.nada.kth.se (03-05-06)

Naeve, Ambjörn./Nilsson, Mikael/ Palmér, Matthias. (2001), *The Conceptual Web - our Research Vision,* Proceedings of the First Semantic Web Working Symposium, Stanford University, USA. Available at http://kmr.nada.kth.se (03-05-06)

Naeve, Ambjörn/ Nilsson, Mikael /Palmér, Matthias. (2002) Semantic Web Metadata for e-Learning - Some Architectural Guidelines. Proceedings of the 11th World Wide Web Conference (*WWW2002*, Hawaii, USA). Available at http://kmr.nada.kth.se (03-05-06)

Neuroth, Heike./Koch Traugott (2001) *Metadata mapping and applications profiles.* Proceedings of DC 2001 conference (Tokyo: National institute of informatics) Available at http://www.lub.lu.se/netlab/staff/koch.html (03-05-06)

Nilsson, Mikael. (2001a), "*The Semantic Web: How RDF will change learning technology standards* ", Centre for Educational Technology Interoperability Standards (CETIS). Available at http://www.cetis.ac.uk/content/20010927172953/viewArticle (03-05-06)

Nilsson, Mikael (2001b)**"**The next wave: CETIS interviews Mikael Nilsson about the Edutella project", Centre for Educational Technology Interoperability Standards (CETIS). Available at http://www.cetis.ac.uk/content/20010927163232/ (03-05-06)

Nilsson, Mikael. (2002), *Geometric Algebra with Conzilla - Building a Conceptual Web of Mathematics.* Master Thesis in Mathematics (Stockholm: Department of Mathematics, Royal Institute of Technology). Available at http://kmr.nada.kth.se

(03-05-06)

Nilsson, Mikael./Palmér, Matthias. (1999**),** *Conzilla - Towards a concept browser ,* CID-53, TRITA-NA-D9911, (Stockholm: Department of Numerical Analysis and Computer Science, Royal Institute of Technology). Available at http://kmr.nada.kth.se (03-05-06)

Palmer, Sean (2001) *The semantic web. An introduction.* Available at http://infomesh.net/2001/swintro (03-05-06)

Peurell, Eric (2000) *Users and producers on line* (Stockholm: Royal library)

Recker, M.M./Wiley, D.A. (2000). "An non-authoritative educational metadata ontology for filtering and recommending learning objects" (Logan: Dept of Instructional Technology, Utah State University). Available at http://wiley.ed.usu.edu/docs/non-authoritative.pdf (03-05-06)

Sloep, Peter (2000) *Learning technology standardization*. (Open University, Netherlands). Issued by http://eml.ou.nl

Sowa, John (2000a) *Ontology, metadata and semiotics.* Available at http://users.bestweb.net/~sowa/peirce/ontometa.htm (03-05-06)

Sowa, John (2000b) *Knowledge represention .Logical, Philosophical, and Computational Foundations* (Pacific Grove: Brooks Cole Publishing)

Svedjedal, Johan (2000) *The literary web. Literature and publishing in the age of digital production.* (Stockholm: Royal Library)

Svedjedal, Johan (2001) *Den sista boken* (Stockholm: W & W)

Sörlin, Sverker (2002) "Cultivating the places of knowledge" in *Studies in philosophy and education (forthcoming).* Available at http://www.sister.nu/pdf/wp_9.pdf

(03-05-06)

Truedsson, Lars (2001) “Nya metoder krävs för bättre sökning” *in Klassrum Direkt*.no 3 (Stockholm: Skolverket)

Wallberg, Charlotte (2000) *Uppmärkning av skönlitterära verk : en jämförelse mellan användningen av Text Encoding Initiative och Hypertext Markup Language*. (Borås : Bibliotekshögskolan, Högskolan i Borås). Available at http://www.hb.se/bhs/slutversioner/2000/00-35.pdf (03-05-06)

Whitley, E. A. (2000) *“Tacit and Explicit Knowledge: Conceptual Confusion around the Commodification of Knowledge”*. Conference proceedings from “Knowledge Management: Concepts and controversies” feb 2000, BPRC, University of Warwick, England. Available at http://bprc.warwick.ac.uk/papers.html (03-05-06)

Wiley, David (1999). *Learning objects and the new CAI: So what do I do with a learning object ?* (Provo: Dept of Instructional psychology and technology, Brigham Young university) Available at http://wiley.ed.usu.edu/docs/instruct-arch.pdf

(03-05-06)

Wiley, David (2000a). “Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy “ in D. A. Wiley (Ed.), *The Instructional Use of Learning Objects: Online Version*. Available at http://reusability.org/read/chapters/wiley.doc (03-05-06)

Wiley, D. A. (2000 b) . *Learning object design and sequencing theory.* Dissertation (Provo: Brigham Young university) Available at http://wiley.ed.usu.edu/writings2.pl (03-05-06)

Åström, Petter (1999) *XML.* (Stockholm: Docendo)

Östlund , A/Hermundstad, H. (2001) *XML, DTD:er och XSL*. (Stockholm: Docendo)

Personal contacts:

Stina Degerstedt, Royal library, Stockholm
Jerker Fransson, Swedish association of educational publishers, Stockholm
Sten Hedberg, Uppsala university library, Uppsala
Per Lannerö, Metamatrix development & consulting, Stockholm
Net lab, Lund university library, Lund