SeLeNe - Preliminary Report: Learning Objects, Meta-Data and Standards

Kevin Keenoy
kevin@dcs.bbk.ac.uk

January 16, 2003

Abstract

We will critically examine various possible definitions of a "Learning Object" and propose a working definition to be used in the context of the SeLeNe project. A review of existing specifications for metadata associated with learning objects will be given, with a focus on their semantics and their capability to describe objects at different levels of granularity. Details of some systems that currently implement these specifications will also be examined.

Metadata specifications are currently expressed in various different ways - we will consider the feasibility of expressing the existing specifications using RDF, and look at the compatibility of different specifications expressed in this way.

The current expectations of teachers and learners with respect to computer-based learning, as well as what they would like to see these systems do in the future, will be assessed via a preliminary examination of existing studies and reports. This should give an idea of the kind of operations that any system to grow from the SeLeNe project should support if it is to be accepted by e-learning communities.

The possibilities of the personalisation of access to learning objects will be considered, including the necessary semantics of a user profile that is to be useful in helping to select relevant learning objects based on the RDF representation of the metadata associated with them.

1 What is a Learning Object?

One of the simplest views of learning objects is simply as re-usable components of courses. These components can vary in size (e.g. a diagram, a question, an exam paper, a lecture), but are generally thought of as being smaller than an entire course. Few people would disagree with this definition, but it is far too general to be of use. Problems arise when more formal definitions are attempted.

The definition used by the LTSC (Learning Technology Standards Committee) of the IEEE: "Learning Objects are any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning" (LTSC 2000), as Wiley (2000) points out, "fails to exclude any person, place, thing, or idea that has existed at anytime in the history of the universe, since any of these could be "referenced during technology supported learning"". The LON (Learning Objects Network Inc.) use an equally broad definition:
"Learning Objects are stand-alone 'chunks' of information that have value. Examples include: a chapter in a text book; an appendix in another book; a map; a graphic; an interactive

application; an online video; a wiring diagram; a simulation; and so on. A Learning Object may be any size. It could be a single pedagogical concept - e.g. the boiling point of water - or the concise instructions to install a transponder on an aircraft." (LON 2002).

However, while their definition includes every object and abstract concept ever to have existed it is clear from their usage of the term that LON generally intend to refer only to digital objects.

This ubiquity of learning objects is not a bad thing in itself - it is true that anything that has existed can be put to some educational use, so maybe we should allow the definition of learning object to be as broad as this. However, an all-encompassing definition is probably not so useful when discussing the automatic generation of meta-data and other issues of relevance to the SeLeNe project.

It may be that the ideal scenario is the availability of RDF-formatted detailed meta-data describing everything that has existed or occurred - allowing a search system to point learners in the direction of books, buildings and places that will enable them to fulfil their learning objectives. This is probably an over-ambitious vision for the time-being though, and outside the scope of SeLeNe. The learning objects described by the meta-data in a Self e-Learning Network are to be those available on the Web, so it is proposed that we use the term "Learning Object" to mean "Learning Object available on the Web".

One of the key advantages of web-based resources as educational tools is that they can be used simultaneously by many users, unlike traditional resources such as textbooks and worksheets where each learner needs a separate copy. We may want to include this sharability as a feature of the learning objects we will deal with in SeLeNe.

Maybe a working definition (which borrows from Simon and Quemada (2002)'s definition of "educational material") for use in SeLeNe could be something like:

Learning objects are electronic, sharable chunks of reusable learning content, available on the Web.

This definition includes both static and dynamic web-based objects at all levels of granularity but excludes physical objects such as textbooks and CD-ROM's, which are not sharable and cannot be stored in the kind of distributed learning-object repositories envisaged in the SeLeNe project.

2 Existing Metadata standards

2.1 The Dublin Core

The Dublin Core Metadata Initiative is a group that promotes the adoption of interoperable metadata standards across a wide range of application domains and disciplines. The Dublin Core Metadata Element Set seems to be by far the most widely accepted and used set of metadata standards for 'core' categories applicable to any internet-based content. Almost all existing learning object metadata standards use the Dublin Core as a basis and then extend it with more specialised elements. The fifteen Dublin Core elements are:

- Title - A name given to the resource.
- Creator - An entity primarily responsible for making the content of the resource.

- Subject - The topic of the content of the resource.
- Description - An account of the content of the resource.
- Publisher - An entity responsible for making the resource available.
- Contributor - An entity responsible for making contributions to the content of the resource.
- Date - A date associated with an event in the life-cycle of the resource.
- Type - The nature or genre of the content of the resource.
- Format - The physical or digital manifestation of the resource.
- Identifier - An unambiguous reference to the resource within a given context.
- Source - A reference to a resource from which the present resource is derived.
- Language - A language of the intellectual content of the resource.
- Relation - A reference to a related resource.
- Coverage - The extent or scope of the content of the resource.
- Rights - Information about rights held in and over the resource.

Each of these elements is described using a set of ten attributes, six of which are common to all elements. The common elements are 'version', 'registration authority', 'language', 'obligation', 'datatype' and 'maximum occurrence'. The other four attributes, which apply to only some of the elements, are 'name', 'identifier', 'definition' and 'comment'. The elements can also be 'qualified' with additional attributes (such as encoding schemes, enumerated lists of values, or other processing clues) if wished, although this can reduce the interoperability of the metadata.

The Dublin Core pre-dates RDF, but it can be expressed using it and as of 25th October this year RDF is the recommended form of encoding for their metadata elements (see http://dublincore.org/documents/2002/07/31/dcmes-xml/ for details of the encoding).

2.2 IEEE LOM

The IEEE Learning Technology Standards Committee (LTSC) has produced a set of specifications for metadata associated with learning objects (IEEE 2002), which as of June this year is approved as an IEEE-SA standard (IEEE 1484.12.1 - 2002). The standard builds on the Dublin Core, and is based on recommendations from the ARIADNE project and IMS (see later). The LOM metadata specification forms the basis of almost all existing implementations of metadata specifications for learning objects, and should probably be the basis for metadata used in SeLeNe. The LOM specifies nine categories for metadata elements associated with learning objects, which group the data elements:

- General - information describing the object as a whole.
- Lifecycle - features relating to the history and current state of the object.

- Meta-Metadata - information about the metadata instance itself.
- Technical - technical requirements and characteristics of the learning object.
- Educational - educational and pedagogical aspects of the learning object.
- Rights - intellectual property rights and terms of use.
- Relation - define relationships between this and other learning objects.
- Annotation - comments on the educational use of the object.
- Classification - describes the learning object in relation to a classification scheme.

The Dublin Core elements (see (i), above) fit into several of these categories. None of the elements of the LOM are mandatory, which means that a LOM instance with no values for any of the elements is still said to conform to the standard. A LOM instance that extends the element set with its own elements is also said to conform to the standard. In this sense most of the existing implementations of learning object metadata specifications conform to LOM - they generally define some of the LOM elements to be mandatory, and extend this set of elements with their own 'custom' elements.

Questions of representation and encoding are not addressed by the LOM - it simply defines a conceptual structure for learning object metadata. However, RDF bindings for the LOM schema are available at http://kmr.nada.kth.se/el/ims/md-lomrdf.html.

2.3 ARIADNE & IMS

The ARIADNE project ran from 1996-2000 and, with IMS (Instructional Management Systems), produced a set of recommendations for educational metadata that helped form the basis of the IEEE LOM. People involved in the ARIADNE project have since founded the ARIADNE Foundation, which seeks to build on the achievements of the original project.

Both of these organisations now promote the use of their own metadata standards, both of which conform to the LOM standards - they take a subset of the LOM (sometimes with minor name changes), and augment these with extra elements of their own. ARIADNE specify a minimal set of mandatory elements for any learning resource along with some other optional ones, whereas the IMS specifications follow the LOM and do not specify that any fields elements must exist. The idea behind ARIADNE making some metadata elements mandatory is to address the conflict that exists between two principles they think learning object repositories should adhere to - (1) that metadata creation by learning object authors or indexers should be as easy as possible, and (2) that search for useful learning objects should be as easy as possible. The minimal set should allow for relatively good search capabilities without being too much of a burden to create.

The IMS specifications have RDF bindings, but they currently recommend an XML-based implementation. The ARIADNE foundation provides tools for authoring metadata conforming to their specifications, but do not seem to give implementation details from their website. Their authoring and querying user interfaces are probably worth looking at in the context of SeLeNe, and they can be downloaded free from the website.

2.4 CEN/ISSS LT

The European Committee for Standardisation (CEN) Information Society Standardisation System (ISSS) Learning Technologies (LT) workshop is currently involved in work on the internationalisation of the IEEE LOM specifications. Their aim is for versions of the LOM in all EU languages - German, French, Spanish, Italian and Catalan versions already exist. They are also interested in the identification of alternative versions of learning objects in other languages. They agreed in April 2002 to set up an e-Learning Technology Standards observatory - the closing date for applications is November 2002, so it will be worth keeping an eye on their activity as work on SeLeNe progresses.

3 E-Learning - User Expectations and Requirements

There seems to be much anecdotal mention of user expectations for e-learning systems (e.g. "our system exceeds users' expectations", "most e-learning systems fail to meet user expectations", "e-learning gives low learner satisfaction") but little in the way of published user requirement studies. Perhaps many of the studies that have been carried out are jealously guarded commercial secrets. The many people offering guidelines about how to make great e-learning systems seem to be self-professed gurus, with little or no evidence to back up their claims that "this is how to satisfy users of an e-learning system". Completion rates for purely electronic courses are generally low though, with 15-20% being a respectable figure. This would seem to suggest that the systems often fail to meet user requirements somewhere along the way.

Many of the suggestions about how to satisfy users of an e-learning system are to do with the final presentation of learning material, and have a lot in common with principles of good web-design in general. These include things like:

- Keep the number of 'clicks' needed for navigation to a minimum.
- Have versions of content that are suitable for users with any connection speed.
- Keep scrolling to a minimum.
- Have a consistent user interface, and include a 'help' button.

However, issues of the final presentation of material really fall outside of the scope of SeLeNe - 'good' learning objects will conform to good presentation principles, but the SeLeNe project itself is not about the production of learning objects.

There are some issues that apply to all teaching and learning, which are not very well addressed by current e-learning systems:

- Learners have different learning styles - different people build, process and store knowledge in different ways. This means that different people will relate to a particular learning resource in different ways. Human instructors can learn which style of presentation suits which learner and adjust their mode of presentation accordingly. Current e-learning systems don't really allow for this at all, and will present the same sequence of learning modules to every user of the system. Some systems do allow course designers to specify which materials should be presented to which users at which time, but

even where such facilities exist they are rarely used. This is due to the time, effort and expense involved in producing multiple pieces of content presenting the same material in different ways, and then setting up the system to deliver it only to the right students.

A solution would be to select materials for users with different learning styles automatically, thus removing the burden from the course designer - if learning objects are marked-up with sufficiently detailed pedagogic metadata (e.g. "this learning object is particularly useful for helping visual learners grasp such-and-such a concept"), and learners have profiles that match up with this (e.g. "this person learns things best when they are presented visually"), then a search utility over a learning object repository should be able to match users' learning styles as well as the content they require.

- Learners have different backgrounds and previous experience - so different learners may need to focus on different material to achieve the same eventual learning objective - e.g. one learner might need to gain knowledge of statistical techniques, another medical practice, and another recent political history, for each of them to reach the objective "be able to critically assess problems facing the health service today".

There are also issues specific to searching for electronic educational resources:

- Learners at different levels of attainment or following different curricula may use the same search terms, but are looking for very different material, e.g. a Secondary school pupil searching for 'atomic structure' will require different information from a Chemistry Masters student searching for the same thing. This problem arises less in a traditional learning environment because the library (and hence textbooks) available to students at different levels are different. In our example each student would visit their respective libraries, look for a science/chemistry textbook, check for 'atomic structure' in the index and find material at the required level. With a distributed electronic repository of learning objects materials at all levels would be returned by a search for 'atomic structure' (give it a go yourself - http://www.google.com/search?q=atomic+structure - the top result is about the 3D structure of inorganic chemicals, the second for US High School student pages on chemistry, the third a history of the atom from the Ancient Greeks to present day).

- Even learners at the same level have different objectives in seeking learning material - e.g. "I want to cram enough to be able to pass an exam" as opposed to "I really want to deepen my understanding of this topic".

In terms of user expectations and requirements all of these issues can be summarised as "I want the system to give me exactly what I need, when I need it, even if I don't know exactly what I need myself". This is a tall order, but hopefully SeLeNe can begin to address some of these issues through the creation of user profiles to aid personalisation of search and navigation through a space of learning objects and their associated metadata.

4 User Profiles and Personalisation

There are some existing standards for user profiles, some of which may be useful as a basis for the user profiles used by SeLeNe.

4.1 vCard

The vCard schema for personal (and business) information covers the 'basics' of personal information, and has an existing RDF expression (http://www.dstc.edu.au/Research/Projects/rdf/draft-iannella-vcard-rdf-00.txt). It holds the kind of information usually found on a business card - name, address, date of birth, e-mail address, etc. This is obviously nowhere near enough information to do any useful personalisation of access to learning objects, but as it is a standards-based specification it may be worth using as a basis for our profiles.

4.2 IEEE LTSC Personal and Private Information (PAPI) draft standard

The PAPI standard (http://ltsc.ieee.org/wg2/papi_learner_07_main.pdf) gives a specification for both the syntax and semantics of a 'Learner Model'. This can characterise a teacher or learner, and holds information on learning styles, existing skills and abilities, etc. as well as basic personal information. It allows definition of elements at many levels of granularity. This specification is definitely worth looking at in more detail in the context of SeLeNe. It may have all the elements we will want in a user profile to help personalise access to learning objects. Details of the possible encodings of the PAPI data need to be explored (specifically, the possibility of using RDF).

4.3 eduPerson

EduPerson is a scheme used by US universities to enable transfer of information about people involved in higher education (both staff and students). It holds little information over and above vCard - it has some additional attributes such as affiliation, description, entitlement and preferred language. SeLeNe will need to hold some of this information in its profiles, but as eduPerson is primarily a US innovation that adds little to the vCard standard we probably don't need to worry about it too much.

4.4 IMS Learner Information Package (LIP)

A LIP (http://www.imsproject.org/profiles/lipbest01a.html) stores data about learners in eleven categories. The data stored is supposed to aid "recording and managing a learning-related history, engaging a learner in a learning experience and discovering learning opportunities for learners". This specification is also worth looking at in the context of SeLeNe.

4.5 Universal Learning Format

ULF (http://www.saba.com/standards/ulf/Pdf/ulfOverview.pdf) is a specification developed by Saba Software, based on Dublin Core, vCard and other educational metadata standards to describe both learning content and learners themselves (so it could have been mentioned in section 2 on existing metadata standards too). Formats are defined for competencies, profiles and certification. RDF is used for resource description and discovery - although it is a proprietary specification it may be worth looking at in the context of SeLeNe.

4.6 Personalistion Conclusion

However personal profiles are stored, we will need to identify the information that will be most useful in personalising learner's search of and access to learning objects. This will certainly

include details of things such as preferred learning styles, current levels of attainment, learning goals, interests, locality information, languages, and learning history. User profiles should adapt and expand over time as users interact with the system, and as learning objectives become skills that have been gained.

5 References

IEEE (2002), Draft Standard for Learning Object Metadata [On-Line]. http://ltsc.ieee.org/doc/wg12/LOM_1484_12_1_v1_Final_Draft.pdf

LON (2002), http://www.learningobjectsnetwork.com/Concepts.htm

LTSC (2000), Learning technology standards committee website [On-line]. http://ltsc.ieee.org/

Simon, B & Quemada, J (2002), "A Reflection of Metadata Standards Based on Reference Scenarios" [On-Line]: http://www.wu-wien.ac.at/usr/wi/bsimon/publikationen/SimonQuemada-ReflectionOnMetadataStandards.pdf

Wiley, D. A. (2000), "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.* Retrieved Nov 01 2002 from the World Wide Web: http://reusability.org/read/chapters/wiley.doc