WebALT Metadata = LOM + CCD

Jouni Karhima, Juha Nurmonen, and Matti Pauna

University of Helsinki, Department of Mathematics and Statistics
P.O. Box 68 (Gustaf Hällströmin katu 2b)
FI-00014 University of Helsinki
{Jouni.Karhima, Juha.Nurmonen, Matti.Pauna}@Helsinki.Fi[1]

Abstract. We explain the chosen technologies for describing and storing the multilingual mathematical e-learning content produced in the WebALT project. We use the IEEE standard LOM (Learning Object Metadata) for encoding metadata. For classifying contents we use CCD (Course Content Dictionary) which is taxonomy describing mathematical topics and their relations. We present an implementation of these technologies, called WALTER (WebALT E-resources Repository), which is a distributed repository holding contents, their metadata, and the classification structure.

1. Introduction

Metadata and classification technologies are important topics in e-learning area, because they enable better searching and sharing of digital materials. Metadata is machine understandable information to describe a resource, e.g. its author or what topic it is about. Furthermore, the format of the resource and possible system requirements can be given helping to display the content appropriately. Often metadata attributes of a resource are obtained from controlled and formally specified vocabularies which set a common framework of the concepts in a given subject matter. There are several alternative techniques in this field and in this paper we discuss some of them from the viewpoint of the Web Advanced Learning Technologies (WebALT) project [1] whose objective is to create an XML database of multilingual mathematical e-learning resources.

Classification technologies, or taxonomies, are used to classify digital contents, usually collected into a repository. Particularly, in mathematics teaching several taxonomies are used, each created for different purposes. The American Mathematical Society has developed the AMS Mathematics Subject Classification [2] which is usually used for defining the exact subject area of research publications. OpenCyc [3] is a project where ontologies for several fields, including mathematics, are being developed. LivingTaxonomy [4] is an open-source project for defining taxonomies for education in several disciplines.

Taxonomies are used in online repositories for learning objects to classify and store contents. For example, Merlot [5] has for several years collected large amount of educational materials with its own taxonomy. Similar project is the Campus of Alberta Repository of Educational Objects (CAREO) [6] which uses the standard LOM metadata (see later) but it only supports key words rather than a real taxonomy. A large repository

[1] This work is sponsored by the eContent project EDC-22253-WEBALT

collecting educational resources in the United States is Gateway to Educational Materials (GEM) [7]. The European project LeActiveMath [8] is developing an online repository of learning materials specifically for mathematics education. In many of these cases the problem is that in a repository the taxonomy is pre-defined and users have little possibilities to modify it. For example, the iLumina project observed that "How to best implement or extend standardized vocabularies and taxonomies is an open question" [9].

WebALT has chosen as a starting point for its *extendible* taxonomy the mathematics taxonomy of LivingTaxonomy since it is quite comprehensive and specifically designed to be used in teaching. We have used it as a basis of our own taxonomy called Course Content Dictionary (CCD) which was already introduced in the Helsinki Learning System project [10] and also discussed in [11]. With the CCD it is possible to define and edit the classification with topics and their relations in several languages. Our system is flexible since it is possible to enlarge taxonomies, or even add alternative taxonomies, if needed. In the WebALT system the taxonomy is integrated to the WebALT E-resources Repository (WALTER) in which it is possible to classify e-learning contents according to the CCD.

Classification information for a piece of content is written into its metadata. We use the Learning Object Metadata (LOM) [12] standard for metadata because it is very comprehensive and specifically designed for educational purposes.

In this paper we first discuss mathematical digital learning materials and their format, especially for exercises. Secondly we examine metadata standards and technologies. We shall describe our choices regarding metadata information, e.g. how and what information fields are used. Next the CCD is presented with some comparisons with existing taxonomies. Finally we discuss the WALTER repository and its implementation.

2. WebALT e-learning contents

The WebALT project started January 2005 with funding from the European Union eContent program and it uses existing standards for representing mathematics on the web and existing linguistic technologies to produce a database of language-independent mathematical problems. The final product of the project will be a showcase consisting of a variety of learning resources for mathematics education stored into a repository, in particular multi-lingual exercises.

The WebALT repository can contain many types of learning materials, such as exercises, lecture notes, and animations. Mathematical animations and interactive calculators are particularly useful for the learners as they help to experiment with and visualize mathematical concepts. Animations are implemented usually with Java applets, Maplets (applets enhanced with Maple) and WebMathematica (web pages with Mathematica interactions).

For lecture notes a variety of formats are commonly used ranging from web pages to PowerPoint slide shows and PDF documents. All of these require metadata too and can be stored in the repository. We discuss here some of the chosen formats and technologies for exercises, since they are the primary target for producing learning materials in the WebALT project.

2.1 Exercises

Digital exercises enable automatic testing and practicing of students' mathematical skills. Often mathematical exercises involve multiple steps to achieve the final result so that a student has to master several skills. This has been implemented with *problem trees* (see figure 1) in the Helsinki Learning System project [10] in which some members of the WebALT project took part. From our experience this style of exercises where the student is prompted the next exercise depending on the answer to the previous one is highly desirable.

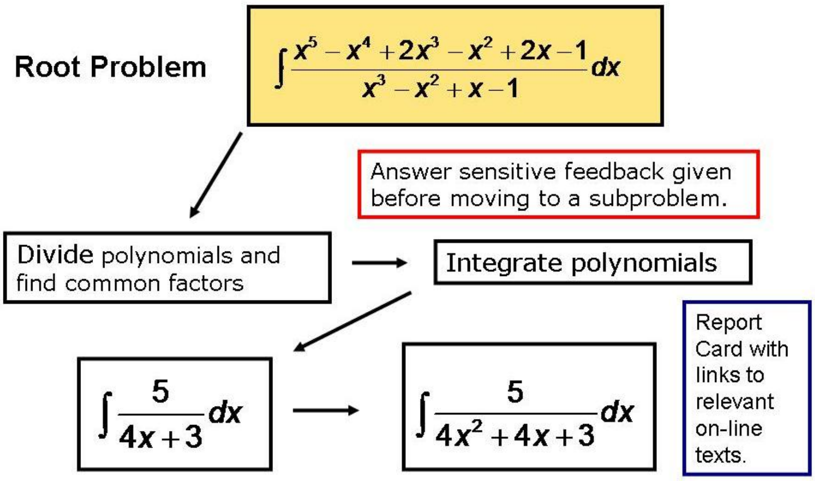

Fig. 1. A problem tree structure displaying feedbacks between steps. Also the system can produce a report card after the whole exercise showing the steps and giving links to relevant topics in the lecture notes (with the help of the CCD).

Problem trees are interesting question types since they can analyze students' skills and give specific feedback during steps. For example, an integration involving partial fraction decomposition can be first given to the student. If she cannot solve it, the system can ask whether she can do the partial fraction decomposition to see if the student had a problem there. So in this way, a rather challenging exercise can be broken into steps with hints. Notice that the individual problems occurring in the nodes of the problem tree could be reused if they were stored into a repository separately. Therefore having fine grained classification helping to find relevant sub problems for a given topic is an important requirement.

The individual problems occurring in the steps of a problem tree can be of many types, ranging from simple multiple choice problems to open problems where students answer with a formula and to graphical problem types, e.g. where the student has to select some points in a diagram. Many systems exist where students can rehearse these problems or even take exams. Moodle [13] is a popular open-source virtual learning environment with only little support for advanced mathematical question types and limited only to multiple choice because it lacks support for computer algebra system back-ends. We would also like to mention AiM (for Assessment in Mathematics) [14] and STACK (for System for Teaching and Assessment using a Computer algebra Kernel) [15] which are computer aided assessment systems designed for mathematics with algebraic evaluation of students' answer by using an external computer algebra systems. There are also commercial systems supporting assessment, most notably MapleTA, which allows for a wide range of problem types and is powered by the Maple system. However, in all of these it is impossible or difficult to implement problem trees.

The generic computer format for mathematics exercises is XML based using MathML or OpenMath [16] to encode mathematical formulas, making them easier to be utilized and displayed in the browser. Moreover, as exercise formats are roughly equivalent in their expressivity, using XML transformations between them is feasible, and therefore one does not have to commit to one format. The most used and implemented digital encoding formats for exercises are Question and Test Interoperability (QTI) [17], which is a standard format for exercises in any subject, MathQTI [18], which is a mathematics extension of QTI, LeActiveMath exercise language [19], and EDU [20], which is a text format not encoded in XML and used in MapleTA. AiM uses Maple packages to produce questions written in AiM syntax which can include Maple commands. STACK, which is a continuation of AiM, has similar language but encoded in XML.

For exercises, WebALT favors the MathQTI format because it is based on OpenMath and QTI standards, from which it inherits the support for an extensive number of question types. In QTI an author can define random template variables producing many different problems from the same template. MathQTI extends QTI by allowing template variable definitions and students' answer processing encoded in OpenMath, enabling more sophisticated algorithmic problem types. As the WebALT exercises will be multilingual, some OpenMath encoding is also needed to express the linguistic attributes of the problem text, see [21] for details. Simple multiple choice problems in mathematics do not necessarily need the sophisticated facilities provided by MathQTI and OpenMath, so some of the WebALT problems could be imported into systems only supporting QTI, such as Moodle. This was also a reason to choose MathQTI since it is upwards compatible with QTI.

3. Metadata

Metadata is commonly understood to be a very important technique to add machine understandable information to describe contents. It enables users to classify, search, and share resources whether they are digital or non-digital. In almost every university, digital material is made available to students by the professors and seldom is the material re-used outside the university or even outside the specific class. One major reason is that the online lectures and problems are not described using metadata and thus not made available in a common framework shared by other potential users.

The production of metadata is gradually being regarded as a key issue, at institutional or national levels, and attention is paid to metadata standards. While digital learning resources develop extensively, the availability of good metadata becomes crucial as soon as those resources are to be shared.

3.1 Metadata standards

Dublin Core [22] is perhaps the most well-known metadata standard. It is designed to describe general publications and as such is inadequate for learning resources. Metadata standards particularly developed for education are for example Learning Object Metadata (LOM) [12] and Dublin Core Educational [23] which defines a set of education-specific elements, element qualifiers, and value qualifiers (controlled vocabularies) to be used with the Dublin Core. The abovementioned GEM repository project also has its own metadata which is an extension of Dublin Core.

The Learning Object Metadata, created by the IEEE, defines attributes that are needed to adequately describe a learning object. LOM is quite comprehensive and certainly more extensive than Dublin Core. A learning object is defined as an entity which can be used and referred to during technology supported learning. This entity can be digital or non-digital allowing in particular metadata descriptions of printed materials. The metadata descriptions of learning objects include content type, learning objectives, software tools, persons and organizations, among others. Learning object metadata can then be used for searching and automatic processing of learning objects and for their classification. A more recent application is metadata harvesting systems, such as Open Archives Initiative [24], providing automatic retrieval of contents from registered metadata repositories.

LOM is an information model and IEEE also provides an XML binding for it defined by an XML schema. The IMS Global Learning Consortium has also developed an XML schema (see [25]) for LOM which is the one the WebALT project uses.

The WebALT repository will be the interface to the database collecting and organizing contents, for instance, the mathematical problems in MathQTI format. LOM metadata is attached externally to every piece of content giving also relevant taxonomy classification information about it.

Although LOM metadata is already specifically designed for educational purposes, it is still necessary to further adapt it to our more restricted context consisting of only mathematics related content. LOM has more than 70 fields distributed into nine categories to cover all potential usage and hence there are also many of them not very relevant for our purposes. Applying and restricting metadata has been regarded as an important issue in the projects that deal with metadata and repositories. For example, the iLumina project has studied the filling and use of metadata in its repository and on basis of it, identified a subset of LOM elements that fit their purposes, see [9] for details. Also the ULI (University teaching network for computer science) project restricted LOM into a subset of 15 fields to be used in their project, see [26]. Therefore it is of also our concern to eliminate these in advance to avoid frustrating the metadata authors and to encourage them to fill in the more important parts as the value of metadata is only in those fields that are carefully filled in. As a consequence and to further emphasize this we have also tried to separate the most important fields from those that give additional information and we try to bring them in front when presenting the metadata input form.

3.2 WebALT's restriction of LOM

We present and motivate here some of the choices we made restricting the LOM elements. On the other hand, some information specific to our project is hard to express with the fields

LOM currently provides. Examples of these are the question type (e.g. multiple choice, formula, etc.), whether the question text is in multilingual or natural language form, and whether the question has feedback and hints. Let us go through some of the LOM categories and analyze their fields in view of WebALT's needs.

Title, Description, Keywords
These are the common fields telling about the content of the resource. Title will give the first impression and description a further glance for someone who is browsing through resources one by one. Keywords will make sure that all closely related words, that are not put in the title or description, will be explicitly mentioned and the resource can be found using keyword searches.
Author, Editors, and other contributors
The author of a resource is a very important piece of information. Not only we want to record it for copyright purposes, but also many users, searching for resources, are likely to eventually identify their favorite authors. Other contributors may also gain regard for only contributing quality content. Inside LOM, author information is encoded in vCard format.
Classification
A great effort has been made in WebALT to produce a high quality classification discipline in view also of its potential use with problem trees. The value in classification is that it creates a whole new method of searching, browsing, to compete or cooperate with keyword search. This method, being less creative and more automatic, is fundamentally different from keyword searching and may result better hits with less effort when well done. More about classifications follows in the next section.
Rights, Cost and Copyrights
This is important metadata, but it may not be up to common metadata or resource editor to provide this information. It is likely that these kinds of fields will be filled in according to strict publisher policies and will even have precompleted values.
Difficulty, Context, and other educational information
Difficulty, from very easy to very difficult, context, school or higher education, and other such fields are significantly trickier to fill in than any of the previous ones and WebALT intends not to force authors to do that. However, we might expect that when browsing a Single Variable Integration topic in the classification, the very easy exercises demonstrate the usage of very basic tools of integration and the very difficult ones present complex problems with no clear-cut solution methods written in textbooks. For this kind of fields it is necessary to create clearly written guidelines to ensure general compatibility.
Technical information, free annotations, etc
Some technical information is always needed for automatic handling and in some cases also for searching. This includes data format, physical size and possibly other things depending on final system implementation. A chance for free annotations is also provided for any notes, observations etc. that have no more suitable field for them.

4. Classification

WebALT course materials will be archived and searched primarily by means of a taxonomic index based on a classification of mathematics topics defined by the Course Content

Dictionary (CCD). The CCD was already introduced as a prototype in the Helsinki Learning System (HLS) [10] and is now implemented more completely in the WebALT project. In HLS the CCD helped to catalogue the single step problems occurring in problem trees for reuse. Since the CCD is a technique for classification, it is also applicable to other disciplines than mathematics. In the WebALT system the CCD is tightly integrated within the metadata editor and repository, where it is used in the user interface, to allow the user to link resources to topics in a tree representation of the taxonomy, as displayed in the next figure.

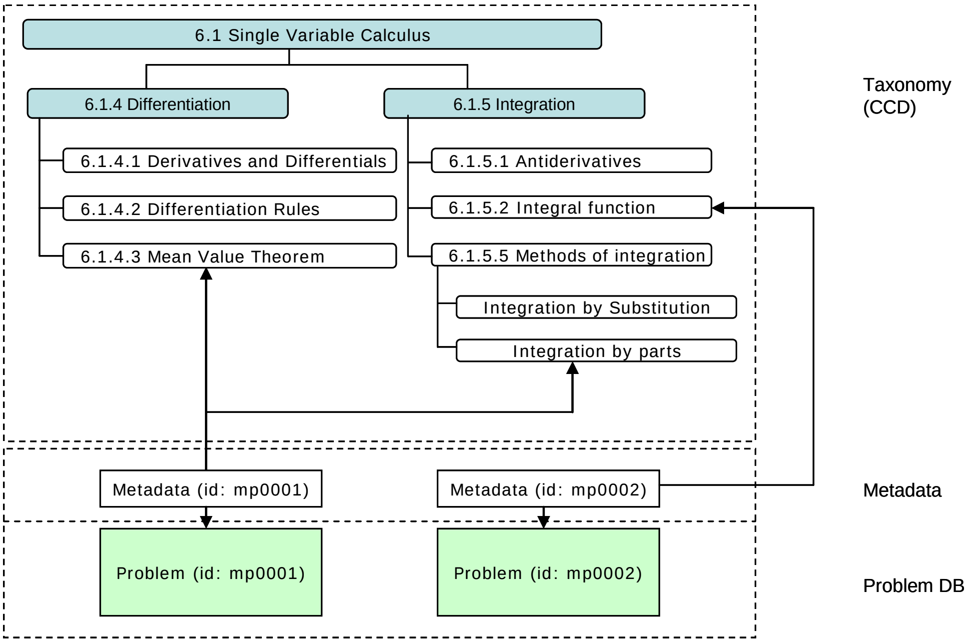

Fig. 2. WebALT repository (WALTER) structure. Metadata is the binding link between resources and the taxonomy.

With the CCD an author can classify materials and share them with other instructors who can browse through the classification to search for suitable content for their courses. It sets a unique framework for authors to develop course materials and therefore it also enables collaboration to produce a basic mathematics course, as each author can concentrate on specific topics. This makes also the lifetime of the material longer, since it is properly stored and classified for others to find and reuse.

The CCD defines a common hierarchy of topics for a course in a way similar to a table of contents in a standard textbook. But on the other hand, the CCD should be understood like taxonomy, so it just describes a collection of topics and their relations, not as a strict table of contents to be followed precisely in a specific course. Topics need not be covered in any specific order or some topics could be omitted in the lecture notes of a specific course. For this reason, the CCD must be quite general so that it covers many topics of a subject making it easy to find a natural place for resources inside it.

The CCD allows constant extending and adapting its current state, which makes it more than static taxonomy. For instance, editors might want to make it more fine grained to allow more detailed classification for their materials preventing the situation that huge amount of resources are classified under a too general topic. For this reason, the taxonomy can be edited, and if a teacher wants to use another, new taxonomies can be created to the repository. It should be noted though that having multiple taxonomies for the same course makes searching and sharing more difficult. For instance, it would be a problem if all the teachers of a course created their own private taxonomies, because then sharing is not likely to happen.

Therefore some policies and guidelines on updating the CCD have to be set. For instance, it should not be allowed to remove a topic if there are contents stored under it or if it has subtopics. On the other hand, one should be able to introduce a new topic if there are enough contents to be classified under it. As a minimum requirement, we have defined user rights levels so that users who are for example allowed to upload contents are not necessarily allowed to change the classification.

4.1 Taxonomies

There are several taxonomies used in mathematics teaching, each created for different purposes. These include American Mathematical Society's Mathematics Subject Classification [2], Merlot [5], Mathematics Content Dictionary (MCD) [27], and OpenCyc [3]. Each of the mentioned can be used as a basis for the CCD. WebALT has chosen as a basis, to expand upon in its CCD, the LivingTaxonomy.

Other taxonomies discussed here are hard to fit into educational purposes. Some of them are too detailed and are therefore suitable only for specifically designed lecture courses. These include AMS Mathematics Subject Classification, which is designed for research content, not for teaching purposes. Also a Finnish taxonomy called Mathematics Content Dictionary (MCD) has its background in a certain polytechnic training program, and therefore is only applicable for similar programs in engineering mathematics. Other taxonomies, in turn, are too general and each topic would cover all too much material. These include Merlot and also OpenCyc.

4.1.1 Living taxonomy

LivingTaxonomy is a quite comprehensive and clear taxonomy. The LivingTaxonomy Project (see [12]) currently has nine taxonomies for various disciplines and they aim at, among other things, creating global, open source, standards-based taxonomies for digital education. In mathematics the LivingTaxonomy Project relies on the Core Subject Taxonomy for Mathematical Sciences Education [28], which is an approved taxonomy for use in classifying digital resources in mathematics education by the Mathematical Sciences Conference Group on Digital Educational Resources. This taxonomy can have several levels, depending naturally on topic. The next figure shows the main topics of this taxonomy and how the main topic Calculus is expanded into subtopics.

1. Numbers and Computation
2. Logic and Foundations
3. Algebra and Number Theory
4. Discrete Mathematics
5. Geometry and Topology
6. Calculus
7. Analysis
8. Differential and Difference Equations
9. Statistics and Probability
10. Applied Mathematics
11. Mathematics History

1. Single Variable
2. Several Variables
3. Advanced Calculus
4. Tensor Calculus
5. Calculus of Variations
6. Operational Calculus

1. Functions
2. Limits
3. Continuity
4. Differentiation
5. Integration
6. Series

Fig. 3. Topic hierarchy of LivingTaxonomy for Single Variable Calculus

We find that the classification criteria used in LivingTaxonomy are adequate and reasonably detailed. On the other hand, it is still flexible for high school and university level teaching, and it also supports several tables of contents used in standard textbooks for mathematics courses.

4.1.2 Extending taxonomy

As mentioned earlier, the CCD allows extensions and adapting. The LivingTaxonomy is a very good taxonomy for general purposes but still too general to be used in a single course. Our experience showed that in many cases it is difficult to find sufficiently narrow topics for materials e.g. in a Calculus university course. For that reason, we wish to extend parts of LivingTaxonomy so that each subtopic covers a reasonable amount of material. This has been implemented for Single Variable Calculus part. To mention one subtopic that has been extended, the subtopic Integration of Single Variable Calculus (see above) looks like follows:

- 6.1.5 Integration:
  - 6.1.5.1 Antiderivatives
  - 6.1.5.2 Integral function
  - 6.1.5.3 Riemann Sums and Definite Integrals
  - 6.1.5.4 Fundamental Theorem of Calculus
  - 6.1.5.5 Methods of integration
  - 6.1.5.6 Improper integrals
  - 6.1.5.7 Numerical Integration
  - 6.1.5.8 Areas and volumes
  - 6.1.5.9 Further applications of integration

4.2 Representation of the CCD

Conceptually, the CCD is a tree as a data structure, where the relation is the subtopic relation. For encoding this structure with XML, WebALT has developed an XML schema under the namespace `http://www.webalt.net/ccd`. In this data structure, which is summarized in table 1, `Topic` is the main element and each topic also allows adding subtopics, related topics, related subtopics and prerequisites into it. The `subtopic` relation is naturally the most important (we have added for performance reasons also the inverse `supertopic` relation to each topic). Some of the data fields describe metadata for the topic, such as author, although topics have very little metadata as they are not really learning objects by themselves. For the metadata of topics we use LOM fields where applicable. Most important of these is the `description` field, where one can input e.g. mathematical definitions and examples. Moreover, the title and description of a topic can be given in several languages enabling the creation of a common taxonomy in a multicultural framework.

Table 1. CCD Data structure

`Title`	Title for the topic as specified in LOM
`Identifier`	Id of the topic
`Description`	Short description of the subject covered by the topic as specified in LOM
`Path`	A hierarchical path to the topic, e.g. 6.1.1
`Contribute`	Author information as specified in LOM
`Origin`	Source of the taxonomy, e.g. LivingTaxonomy
`Subtopic`	List of subtopics of this topic
`Prerequisite`	Topics that are required to understand this topic
`Related Topic`	Cross references to other topics in the CCD
`Related Subtopic`	Cross references to other subtopics of this topic in different branches in the CCD
`Supertopic`	Link to the supertopic

5. Repository implementation

The WebALT project has developed a repository implementation called WebALT E-resource Repository (WALTER), see [29]. It is a platform to store content and edit its metadata which also allows modifying the classification (CCD) for users authorized to do so. WALTER indexes its content by the CCD classification allowing one to browse the hierarchy to search, display, download, and upload resources.

WALTER is a web application running in a Java2EE environment. As the underlying database it uses an open source native XML database called eXist [30], which was chosen because it stores and indexes XML files and can query them using the XQuery language, which is similar to relational query languages, and also allows one to modify XML data with the XUpdate language. The database can also store binary files, so in particular it can contain multimedia e-learning contents. Furthermore, eXist is compatible with the XML:DB standard [31] which is a common interface for communicating with the database to perform standard database procedures, such as XQuery queries and XUpdate modifications.

WALTER is designed to be a *distributed* metadata repository in the sense that one can add metadata records for contents that are stored in remote databases, as shown in figure 4.

This enables the sharing and retrieval of relevant course materials from different locations across the web. It also means that if a new publisher wants to use WALTER for publication, he/she only needs to install a local database, which then can be connected to the centrally managed world-wide system. The different systems communicate with the XML:DB protocol so that each database can be of any compliant implementation such as Apache Xindice and the commercial X-Hive. This assures a real ease-of-use in adopting the system and also a great chance for interoperability between different publishers.

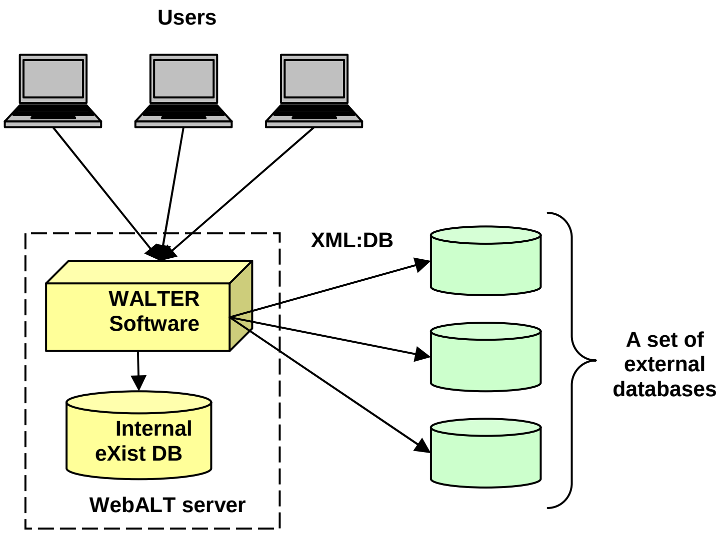

Fig. 4. WALTER software is located in the WebALT server with its internal database that stores for example the CCD classification, users and the locations of external resource databases. New external resource databases can be added simply and WALTER connects them behind the scenes with XML:DB protocol so that a common user may not even notice the distributed system.

WALTER includes a *metadata editor* for editing LOM records such as they are described in section 3.2. In the editor the metadata fields are separated into two groups based on their importance. An author creating metadata will first see the most vital fields and has an option to also view and edit the others. In the end there will at least in some contexts also be a third group of metadata fields which an author is not authorized to edit at all. Ways to encourage authors to fill in as many fields as possible are under consideration, one already implemented is the possibility to store personal *metadata templates*. As resources created by one author often share a lot of common information, and therefore it may be reasonable for him/her to fill in them carefully once and then use the same template in the future.

The searching capabilities of the repository offer an advanced way to use the specified metadata and classification either separately or combined. A user may perform a keyword search to cover the whole repository, or alternatively, first browse into a classification topic and then search only for resources that have this topic occurring in their classification path. It is also be possible to specify in which metadata fields the keywords should occur. After a search has been carried out, it is essential to offer the results in a convenient way showing the relevant parts of the metadata in a compact form, so that the user can immediately pick the items that seem most useful for his/her purposes.

6. Conclusions

The WebALT project combines existing technologies such as OpenMath, metadata, and taxonomies to produce a repository for multilingual mathematical learning materials. WebALT is committed to supporting emerging standards, of which most interesting are the Sharable Content Object Reference Model (SCORM) [32] and ontologies, in the e-learning arena and hence it has adopted existing standards for describing the content of mathematical exercises, such as OpenMath for the mathematical fragments, QTI (extended with MathQTI) for questions and interactivity, and LOM for the metadata content. The CCD technology combined with the WALTER repository enables classifying, storing, searching, and sharing these materials.

The future development of the WebALT project will take into account semantic web technologies because they naturally fit with metadata and taxonomies. In particular, implementing the CCD using ontologies seems feasible. There are some projects already which are developing metadata using semantic web technologies, of which most useful in this setting is the Resource Description Framework (RDF) [33]. Let us mention in particular the SCAM project [34] which builds a framework for storing and handling RDF encoded metadata and closely related one, the SHAME metadata editing framework [35].

As noted earlier, there exist several standards for encoding metadata, such as LOM and Dublin Core. One important issue is how to combine metadata information of a resource described using different standards. One answer to this problem is again provided by RDF by which one can give attributes to a resource using triples of the form (resource, relation, value). The idea is that the parts of this triple are well-defined in the way that a computer can analyze its purpose, since typically the relation and the values come from well-defined ontologies and vocabularies. For instance, in [26] it is showed how to combine Dublin Core and LOM metadata attributes with RDF records. Doing so requires binding LOM to RDF, which is provided in [36]. It is also possible to have the metadata of a resource distributed across the internet since these RDF triples can co-exist in different metadata repositories describing various aspects of the resource.

As mentioned, SCORM and similar content delivery standards are important technologies gaining support from e-learning platforms. SCORM is a comprehensive set of standards for e-learning technologies, specifying also behavior of content, supported partly by many virtual learning environments, for instance, WebCT and Moodle. Therefore it is natural to add support for importing and, most importantly, exporting SCORM packages from the WALTER repository.

A future goal also for the project is making the repository accessible through web services technologies, such as SOAP messages. This would enable better distribution with systems not necessarily using exactly the same metadata format and database accessing protocol (i.e. XML:DB) as WALTER does. At the moment some components of the WebALT system are connected via web services, which enable the running of different services (e.g. language generation service) in different servers.

7. Acknowledgements

We would like to thank the referee and especially Olga Caprotti for suggestions and corrections of this paper.

References

1. *WebALT, http://www.webalt.net*.
2. *AMS Mathematics Subject Classification, http://www.ams.org/msc/*.
3. *OpenCyc, http://www.opencyc.org/*.
4. *The Living Taxonomy Project, http://www.livingtaxonomy.org/*.
5. *Merlot, http://www.merlot.org/home/SubjectCatIndex.po?discipline=Mathematics*.
6. *CAREO, http://careo.ucalgary.ca/cgi-bin/WebObjects/CAREO.woa*.
7. *GEM, http://www.thegateway.org/*.
8. *LeActiveMath, http://www.leactivemath.org/*.
9. Heath, B.P., McArthur, David J., McClelland, Marilyn K., Vetter, Ronald J., *Metadata Lessons from the iLumina Digital Library.* Communications of the ACM, 2005. **48**(7): p. 68-74.
10. *Helsinki Learning System, http://mark.math.helsinki.fi/HLS/*.
11. Karhima, J., Nurmonen, J., Pauna, M., *"Course Content Dictionary for sharing on-line educational materials", http://mathstore.ac.uk/articles/maths-caa-series/aug2005/*. 2005, Maths CAA Series.
12. *IEEE Learning Object Metadata, http://ltsc.ieee.org/wg12/*.
13. *Moodle, http://moodle.org/*.
14. *AiM, http://maths.york.ac.uk/moodle/aiminfo/*.
15. *STACK, http://www.stack.bham.ac.uk/*.
16. *OpenMath, http://www.openmath.org/*.
17. *IMS Question and Test Interoperability Specification, http://www.imsglobal.org/question/*.
18. *Mathematical Question and Test Interoperability, http://www.maths.ed.ac.uk/Math-QTI/*.
19. Cohen, A.M., Cuypers, H., Barreiro, R., *MathDox: Mathematical Documents on the Web, http://www.win.tue.nl/~hansc/mathdox3.pdf*. 2005.
20. *Brownstone EDUCampus, http://www.brownstone.net/products/edu/*.
21. Strotmann, A., Ng'ang'a, W., Caprotti, O., *Multilingual Access to Mathematical Exercise Problems.* Electronic Proceedings of the Internet Accessible Mathematical Computation Workshop, 2005.
22. *Dublin Core Metadata Initiative, http://dublincore.org/*.
23. *Dublin Core Education Working Group, http://dublincore.org/groups/education/*.
24. *Open Archives Initiative, http://www.openarchives.org/*.
25. *IMS Metadata Schema, http://www.imsglobal.org/xsd/imsmd_v1p2p2.xsd*.
26. Brase, J., Nejdl, W., *Ontologies and Metadata for eLearning*. 2004, in Handbook on Ontologies (Staab, S., and Studer, R. eds.).
27. *Mathematics Content Dictionary, http://matta.hut.fi/mattafi/materiaalipankki/*.

28. *Core Subject Taxonomy for Mathematical Sciences Education,* http://people.uncw.edu/hermanr/MathTax/.
29. *WALTER - Webalt E-resource Repository,* http://www.webalt.net/WCR/.
30. *eXist Open Source Native XML Database,* http://exist-db.org/.
31. *XML:DB Initiative,* http://xmldb-org.sourceforge.net/index.html.
32. *SCORM 2004,* http://www.adlnet.org/scorm/index.cfm.
33. *Resource Description Framework (RDF),* http://www.w3.org/RDF/.
34. Palmér, M., Naeve, A., Paulsson, F., *The SCAM Framework: Helping Semantic Web Applications to Store and Access Metadata*. 2004, European Semantic Web Symposium 2004.
35. *SHAME: Standardized Hyper Adaptable Metadata Editor,* http://kmr.nada.kth.se/shame/.
36. Nilsson, M., Palmér, M., Brase, J., *The LOM RDF Binding - Principles and Implementation*. 2003, Proceedings of the Third Annual ARIADNE conference.