SITMA 2005, 19-20 August, Terengganu, Malaysia.

USING XML-BASED METADATA FOR GENERATING DYNAMIC COURSEWARE

MASHANUM OSMAN, SAZALINSYAH RAZALI

Department of Software Engineering, Kolej Universiti Teknikal Kebangsaan Malaysia[*], *Karung Berkunci 1200, Ayer Keroh, 75450 Melaka*
mashanum@kutkm.edu.my[†]

Department of Industrial Computing, Kolej Universiti Teknikal Kebangsaan Malaysia[‡], *Karung Berkunci 1200, Ayer Keroh, 75450 Melaka*
sazalinsyah@kutkm.edu.my[§]

Abstract*: The objective of this paper is to explore the mechanism between the learner and learning material in generating dynamic courseware in e-learning environment, The dynamic e-learning develop flexible and dynamic course content according to the suitability level of learner, This dynamic e-learning will be using XML-based metadata in particular. XML based metadata allows any part of a course to be modularized, saves and reused as a learning object. Each learning object is described in the object repository. Learning objects are stored in central object repository and are easily locatable on an object-by-object basis. To achieve learner adaptability, learners are allowed to specify queries that will dynamically construct courses.*

Keywords: e-learning, XML, metadata, dynamic, learner

1. Introduction

E-learning is a combination of learning services and technology to provide high value integrated learning at anytime and anyplace. Web based training is expected to grow for 95 percent per annum for the next few years (Eswari et. al). In other words, this is one of the most popular e-services.

In e-learning environment, there are basically four steps in building the course content (Dehua et. al) :

- Textbook parsing – segment the material into learning fragment.
- Presentation style defining – course compilation and presentation.
- XML validating – all the learning objects that are in the XML form must be validated.
- Courseware template matching – which are the xml files that are being used to describe the structure of the materials.

The development of electronic course material that caters all level of learners will incur high cost. For example in the IT domain, knowledge and skills change rapidly that will need the course content to keep up with the current knowledge. Listed below are reasons on why we need dynamic e-learning (Wentling et. al., 2000).

- High cost.
- Dynamic e-learning is more convenient than the traditional led training.
- Can be distributed geographically.
- Market for online course is demanding and profitable.
- Risk free environment, supports trying out new things and making mistakes.

The learner is still the most important role in the learning equation. In order to set the learning objectives, a clear understanding of the target audience is essential. This target audience for which the learning is delivered decides the appropriate model and approaches. The content constitute to only one side of the coin. A good quality content will be as good as a bad content if not delivered using efficient delivery mechanisms. The delivery mechanisms, which is a part of the technology, can range from a text based facilitated discussion to an interactive simulation exercise and could even include audio or video conferencing.

For an effective course delivery, or dynamic e-learning, a dynamic flexible course production should be produced. A flexible course production environment should build courses from annotated modular learning fragments. This will need efficient mechanisms for the generation, storage and retrieval of

huge metadata annotations. Again, to dynamically construct courses "on demand" this is how we do. The idea is to segment existing course materials (e.g., text books, animations and videos) into the so-called learning objects. Each learning objects will typically represent self-contained units, which are properly annotated with metadata. The dynamic e-learning courses are generated and requested "on demand" by assembling single fragments such as in (Hollfelder et. al) among others. This approach is similar to modularization in software engineering, where the intention is to achieve a high degree of reusable content, whereby the learning fragments can be used in the new contexts.

A learning object should support behavioral change – adapting itself to learner styles and choices. We will label such learning object as dynamic. A learning object can be defined as an entity, digital or non-digital, that may or can be used, reused or reference during any learning activity (Ip et. al., 2001). Multimedia content, instructional content, instructional software and software tools, persons and organizations are some examples of learning objects.

Again, to cater courses for different learner, we need much effort and high cost to develop the electronic courses. The main problem is to keep the course content up to date because it incurs high expenses. The best solution is as stated above that is to develop an application that have some important characteristics such as the learning content can be reused many times, learning content can be shared and allow learning content to be exchanged and shared by separate tools and systems connected via the internet. We need new approaches to develop dynamic e-learning that can generate courses automatically and adapt material to individual user needs. For example, learners log on to the system. As the learners interact with the content, results are passed back to the system. If the system allows it, the content can be generated based on learner information stored in the system. Courses will be generated automatically based on test scores, language preferences, learning style inventories, competencies, certifications, academic level, organizational roles, and other data.

2. Components of E-learning

Components of an e-learning framework may come from various vendors or providers. They need to be integrated in order to provide a seamless interface for end users. Several key components of an e-learning framework are discussed below:

- Learning Management System (LMS)
  An LMS provides an integrated platform for content, delivery and management of learning, as well as accessibility by range of users that may include learners, content creator and administrators. An LMS acts as the central component of an e-learning application. Figure 1 shows a model of LMS.

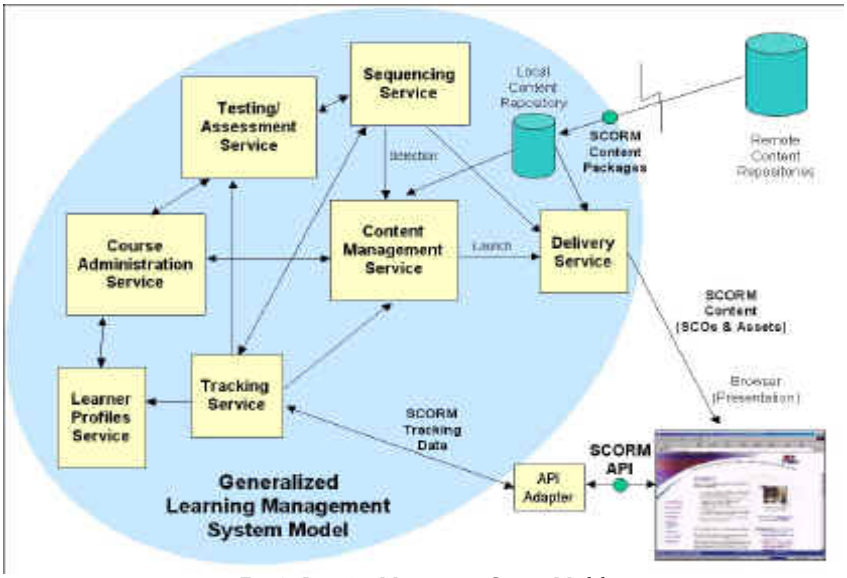

Fig. 1 : Learning Management System Model

- Content composition and integration systems
  Content integration systems include content authoring, sequencing and aggregation tools that allow content to be structured to facilitate the learning process. It allows multiple content developers and subject matter experts to share content and its components over the network. Shareable Content Object Reference Model (SCORM) is a standard specifically for the web-based learning that describe how the various components of learning system work together.

- Learning content metadata
  The e-learning framework needs to support and manage not only learning content but also metadata. The metadata allows the online learning resources to be tagged with searchable properties or attributes. The metadata need to be stored in a repository separate but linked from the learning content.

- The browser as universal interface

Users can access a web-enabled LMS that provides content integration consistently and conveniently from any location on the Internet or intranets. Web browsers will continue to evolve and support new multimedia, interactivity and collaborative technologies. Any distributed e-learning framework will naturally benefit from these advancements.

2.2 Learning Object

There are a number of academic arguments as to what comprises a learning 'object'.We define a learning object (LO) as "A computer mediated or delivered module or unit, that stands by itself, that provides a meaningful learning experience in a planned learning context" (Ip et. al., 2001). The emphasis on "learning experience" is to acknowledge the different type of interaction a learner may have with the "learning object". There is also an explicit reference to the planned learning context as opposed to the spontaneous, accidental learning that occurs when one is engaged in other activity. The learning object is being used with an intention of learning.

As what we have learned from the object oriented software paradigm, that a learning object should possess several interfaces, attributes and methods – that is it should provide several contracts for service and exhibit both state and behaviour. At a minimum, learning objects should have the following three standard interfaces (Ip et. al., 2001):

- A management Interface with an underlying Learning or Content Management System (LMS or CMS) which support authorization for access, accounting for use, learner progress tracking, etc.
- A Learning Interface, which present an appropriate learning interface to learner combine with the rest of the course to support atomic learning experience.
- An Instructional Design Interface, which supports discovery of the resource, customization, and assembly of learning objects into courses suitable for different requirements of the learners.

2.3 Learning Object Metadata

The term learning object metadata are very confusing to most people. The term "meta" comes from a Greek word. It is means "alongside, with, after, next". In Latin and English the word means something transcendental, or beyond nature. Metadata can be defined as "data about data", it is descriptive information about resources for the purposes of finding, managing and using these resources more effectively. Metadata is widely used in libraries system. For example, the library catalog, it contains a set of metadata records with elements that describe a book or other library item: author, title, date of creation or publication, subject coverage, and the call number specifying location of the item on the shelf. It enables the readers to find books or other material that they want easily and quickly.

In the e-learning environment, the resources refer to the learning object. Metadata describes characteristics about content objects used for learning. It facilitates searching, management and linking granules of content. Metadata enables learners, authors and others to search, retrieve and assemble reusable learning objects. Every learning object should be annotated according to the metadata schema properly to ensure that the system can retrieve the learning object base on the specific needs of a course (Hollfelder et. al).

There are issues of interoperability and reusability. Interoperability means that the content from multiple sources working equally well with different learning systems. Whereby reusability is where the content developed in one context being transferable to another context. Interoperability and reusability are the most important characteristic of learning object to generate an automatic electronic course. Without them, every time a course or an interactive electronic training manual needs to be updated, the material must be rewritten and the process of developing high-quality content cannot be achieved. We need a global standard learning object metadata so that we can mix and match content from multiple sources.

A few years ago, there has been proposed to create a global standard for learning object. The main purpose of the proposed project are :

- To enable learner or instructor to search, evaluate, acquire and utilize learning objects.
- To ensure the reusability of learning objects.
- To ensure the interoperability of learning objects.
- To enable computer agents to automatically and dynamically compose personalized lessons for an individual learner.
- To enable education, training and learning organizations, both government, public and private to express educational content and performance standards in a standardized format that is independent of the content itself.
- To provide researchers with standards that support the collection and sharing of the data
- To develop interchangeable content that can be reused, assembled, and disassembled quickly and easily.

On the 6th December 2000, the Learning Technology Standards Committee Learning Objects Metadata (LTSC-LOM) Working Group of the IEEE (Institute of Electrical and Electronics Engineers) and the Dublin Core Metadata Initiative (DCMI) announcing their joint commitment to develop interoperable metadata for learning, education and training[10]. The joint Memorandum of Understanding (MoU) is signed by officers representing the LOM Working Group and DCMI. The document, regarding the IEEE standard P1484.12, is co-signed by representatives of concurring projects: ARIADNE (Alliance of Remote Instructional Authoring and Distribution Networks for Europe), EdNA (Education Network Australia), GEM (Gateway to Educational Materials), and the IMS Global Learning Consortium. The MOU signifies that a core set of specifications and accredited standards is emerging to support an interoperable infrastructure for worldwide e-learning.

3. SCORM

SCORM is an acronym for Shareable Content Object Reference Model, a content descriptive format that was born out of the US federal government's Advanced Distributed Learning (ADL) initiative. The federal government spends millions of dollars each year to develop e-learning content, including online courses, courses distributed on CD's and intranets. In the 1990's the government recognized that it was difficult to reuse this content.

The Department of Defense, for example, found that the various branches of the military had developed e-learning content on similar topics, such as management and acquisition rules. Even though those courses essentially covered the same content, it was nearly impossible to share e-content between military branches because they were developed without a common standard, and they were not designed for reuse in other courses.

The government also realized the benefits of an international standard for e-content on the training industry. A common international standard for sharing learning content would stimulate an international learning economy, similar to the economy that is developing around the Internet. If standards allow for reusing learning content developed for one course, then learning content will become a commodity.

As a result, in 1997, the Department of Defense and the White House Office of Science and Technology Policy launched the Advanced Distributed Learning (ADL) initiative (Jones, 2002). Its primary goal is to develop a learning economy by providing access to high-quality education and training material, easily tailored to individual learner needs and available whenever and wherever needed. To accomplish this goal, the ADL consolidated emerging e-learning specifications from the major international standards groups into a single specification, referred to as SCORM.

Simply stated, SCORM. is a set of specifications for developing, packaging and delivering high quality education and training materials whenever and wherever they are needed. SCORM. - compliant courses leverage course development by ensuring that compliant courses are ***RAID*:**

- Reusable**:** easily modified and used by different development tools,
- Accessible**:** can be searched and made available as needed by both learners and content developers,
- Interoperable**:** operates across a wide variety of hardware, operating systems and web browsers, and
- Durable**:** does not require significant modifications with new versions of system software.

Although this is a government initiative, it is be wrong to think that this cannot be used in industry and academia. In fact the SCORM specifications are a composite of several specifications developed by international standards organizations, including the IEEE, IMS, AICC and ARIADNE. New versions of SCORM are now released every 3-6 months by the ADL (Jones, 2002). Each new release incorporates recent changes and expansions of existing international specification. This process is likely to continue for years to come.

3.1 *SCORM Components*

Scorm can be described in many ways. The Advanced Distributed Learning Co-Laboratories refer to this as the Shareable Object Reference Model (SCORM). SCORM is described in terms of the following three components:

- Content packaging
- Runtime communications and
- Course metadata

Content packaging refers to the packaging of all resources needed to deliver a course into a single zip file. The format for this file is described by the SCORM aggregate model, which is based upon the IMS Content Packaging Specification, version 1.1.2. The zip file contains not only the course files, it also contains an XML file, referred to as the imsmanifest file, describing the course contents and content sequencing.

The runtime communications in a SCORM-conformant course are conducted using two elements:

- Runtime commands for communicating student information to and from the LMS, and
- Student metadata for storing information on individual students.

Course metadata are data packaged with a course when it is archived in a SCORM. Repository. These data allow a course author, or student, search a learning repository containing hundreds of lessons and courses and to identify the learning content they want to use or view. For example, the course title, description, keywords, etc. are all considered course metadata.

4. XML

Reusable learning object represent an alternative approach in developing dynamic e-learning. Course can be composed dynamically base on individual user needs. All the learning object should annotated appropriately with metadata so that system can find, retrieve and assemble the learning object to the specific need of a course. To describe the learning object metadata, we can use IEEE LOM as a reference. The LOM standard is the first standard for learning content to be released by an accredited standards organization. This standard ensures that LOM will have the high degree of semantic interoperability and learning object can be reused and exchange between different LMS. Before this, online learning material sources are annotated with various metadata vocabularies. Problems occur when certain work group wants to use the current learning object in their project. They have to transform the metadata description to the project specific-scheme.

Having small units of learning object will increase reusability because small learning object can be composed more flexible. For example, we can divide a textbook into each paragraph and generating metadata according to LOM standard. A set of learning object can be integrated to compose a course.

SCORM uses IEEE LOM, the SCORM XML file is a <LOM> tagged metadata. XML provide several benefits over HTML or other data formats in an e-learning framework. There are a few factors why XML is the most appropriate language in e-learning (Eswari et. al):

- The most attractive benefit is its simplicity. XML can enable business data and learning content to be served, received and processed on the web as easily as HTML
- XML is extensible, platform-independent and it supports internationalization and localization
- XML allows learning content to be labeled in detail, making it possible to compose electronic course based on a learner's needs. This will allow for more accurate searches.
- XML extends the advantages of interoperability by integrating content and tracking learner progress across several different LMS.
- XML allows for more interactive content than HTML.
- XML provide flexibility, different style can be applied on the same data base on user's need because it allow separation of style from content (Eswari et. al).

If we look through the example below, the learning object has been described base on the 9 categories of metadata structure. Data elements for each category describe the learning object. Table below show the category and data element that used to describe the learning object:

Table 1: Data elements used to describe the learning object

LOM Category	Data element
<general>	Title, catalog entry, language, description, keywords
<lifecycle>	Version, status, contribute.
<meta-metadata>	Metadata scheme
<technical>	Format
<educational>	Description
<rights>	Cost, copyright, description.
<relation>	-
<annotation>	Description
<classification>	Purpose, taxon path, description, keywords.

4. Proposed Design Methodology

It is beneficial to look into other similar framework used by other provider. This can help in understanding more of the approaches taken by various e-learning provider currently available. In our approach, we choose to implement our system using the XML-based metadata. The vast amount of metadata that we have to deal with need for efficient storage and query mechanisms. This can be achieved by using XML technology. XML technology has already been successfully used to furnish solutions for mission-critical data exchange, publishing and software development. (Eswari et. al.,). XML technology enables us to develop application specific markup languages that better describe data and can be exchanged from one system to another system.

4.1 *Existing Conceptual Framework*

Some of the frameworks are conceptual as to the description of the working of the various e-learning provider in reports and papers. Three e-learning frameworks, which are related to the proposed framework in this paper, are given below

- Teachware On Demand .
- Network-Training Collaboration in Europe and China (NCEC).
- Dynamic Courseware Generation on the WWW (DCG).

4.2 *Teachware On Demand*

The project "Teachware on Demand" develops infrastructures and tools for creating, updating and distributing electronically available teaching content. In the TOD approach, a learning fragment is a self contained, modular piece of course material. These fragments are annotated to their schema which is project-specific extension of the standardized LOM. The metadata annotations are encoded in XML. For course compilation and presentation, TOD is generated semi-automatically by selecting (based on query) appropriate course fragments and by structuring them into a training course, which is a composition of fragments. The course structure is built dynamically on the corresponding pre-knowledge conditions of fragments to be selected. They are using bottom-up approach, where the course structure for example a table of content is static and given a priori. The complete course is then presented to the learner who has access to a web portal or to a company's knowledge base.

4.3 *Network-Training Collaboration in Europe and China (NCEC)*

NCEC is a joint project between Chinese and European universities and institutions, aiming at producing an Internet-based interactive learning system, which is highly individualizable in order to satisfy diversified needs of user with reusable learning components namely learning objects paradigm. NCEC system consists of seven subsystems:

- LO authoring system
- User accessing system
- User management system
- Learning administration system
- LM composing system
- LO search engine system
- Repository accessing system

Online courseware authoring is a four step production process including textbook parsing, presentation style defining, XML validating and courseware temple designing. In NCEC project, all learning objects and their structures in the systems are formulated by XML. A LO is a self-contained component that can be organized together to compose Learning Material.

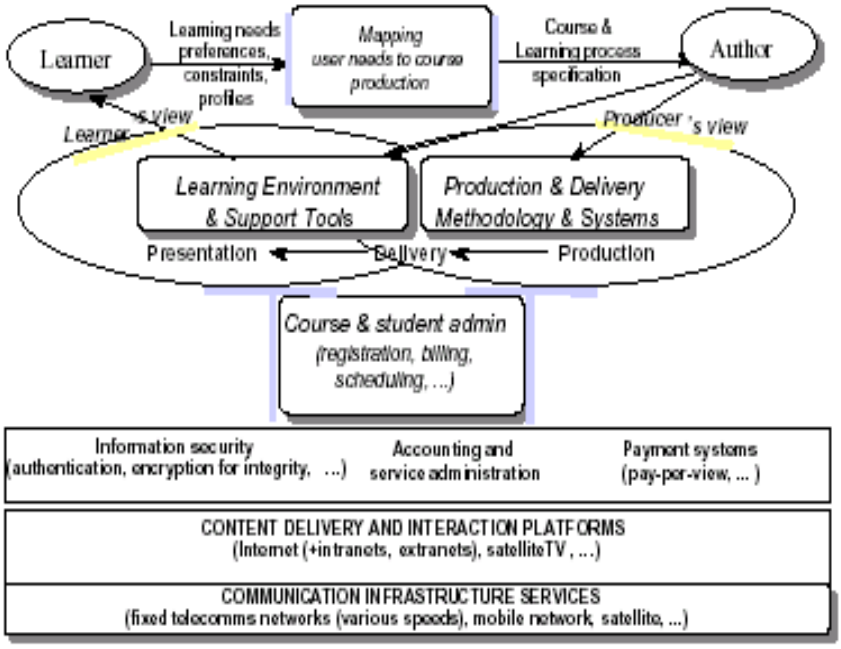

Fig 2: NCEC System Framework (source (Dehua et. al))

4.4 *Dynamic Courseware Generation on the WWW*

A tool for the authoring of adaptive CAL courses called the "Dynamic Courseware Generator" (DCG). This tool generates an individual courses based on the learner's goals and previous knowledge. The courses are dynamically generated according to the learner's level. The DCG runs on a WWW server. The DCG allows:

- Dynamically assembling CAL courses with different goals from a pool of teaching materials
- Adapt learning materials for different learners
- Changing the course dynamically according to the progress of the learner.

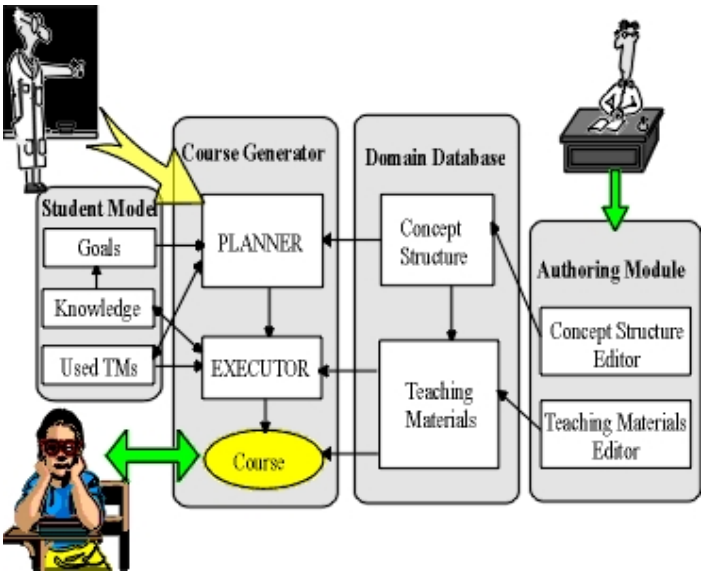

Fig 3 : Dynamic Courseware Generation Architecture (source (Vassileva & Deters, 1998))

5. Proposed Framework

The framework that is proposed for this paper is by improving the search method in selecting and composing the learning object. The actual algorithm is taken from the Teachware on Demand. We believe the Teachware on Demand algorithm's approach which is used in the composing and sequencing the material is the best way compared to the other existing framework., because this approach uses bottom up approach compared to the other framework that uses top down approach. Later, we will discuss why bottom up approach is preferable than the top down approach.

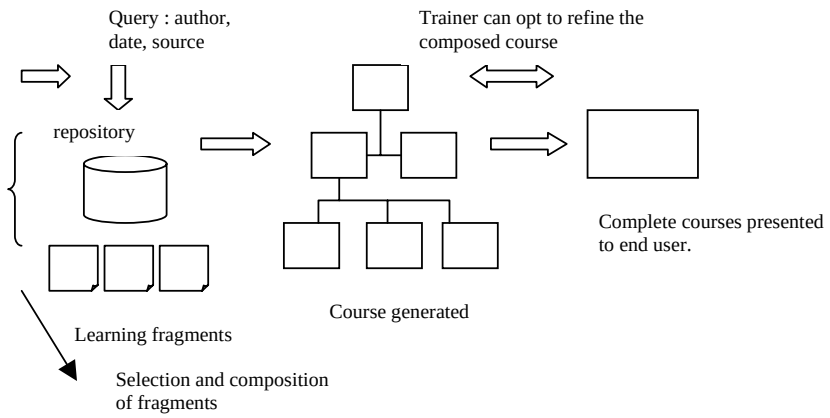

Fig 4: Proposed Framework

In our approach, a learning object is a self contained, modular piece of course material. Self contained here shows that the learning object is independent and is a module of course material. The currently implemented select and sequence algorithm is based on a very simple local search algorithm in the Teachware on Demand (Caumanns, 1998):

1. Define a set of keywords and topics that should be taught by the application.
2. Find a minimum set of learning objects that contains all of these keywords and topics as index entries within the objects' joined taught knowledge.
3. Define the set of learning object found in step 2 as the set of objects used. Define their joined taught knowledge as the knowledge provided, their joined required knowledge as the knowledge needed.
4. While the knowledge needed is not empty:
    a. Find a learning object within the leading knowledge graph that best suits
2. with the current set of objects used.
    a. Add this learning object's taught index entries to the set of taught knowledge
3. and take away the same index entries from the knowledge needed.
    a. Join all the learning object's required index entries (without the ones being members of the knowledge taught) with the knowledge needed.
4. Find learning object within the other layers that suit into the existing graph.
5. Try to improve the learning graph by adding or exchanging single learning objects. Add redundant learning object from all layers according to the desired level of detail.
6. Spilt the learning graph into a set of chapters and sections by using structural information from all learning object' joined index. Add redundant learning objects from all layers to make chapters and sections "look round".
7. Split each chapter into a set of frames and add hyperlinks and other navigational items.

To find the best matching learning fragment, we propose to rank the metadata. The evaluation function based on heuristics, such as to prefer learning object with a high outer grade (many keywords taught), and to avoid index entries with a low inner grade (few explanations available). Learning object that teach keywords with a high outer grade (very often used keywords) seem always to be a good choice, while learning object teaching seldom used keywords very often produce 'unwanted contents.'

We propose to code the algorithm in XML using the LOM standard. XML provides several benefits over the HTML (Hyper Text Markup Language) or other data formats in an e-learning framework. The most attractive of XML is its simplicity. XML can enable learning content to be served, received and processed on the Web as easily as HTML and it also works easily across organizational boundaries. We choose to use the LOM schema because of this few factors:

- The LOM standard is the first standard for learning content to be released by an accredited standards organization.

- This standards ensure that LOM will have the high degree of semantic interoperability and learning object can be reused and exchanged in different Learning Management System (LMS)

We believe the algorithm used by Teachware on Demand is most suitable because of this few factors (Caumanns, 1998):

- Using a bottom-up approach structure, which with this approach it can reduce the amount of manual work in the redesigns and modifications.
- Bottom up approach allows for a finer grained adaptation, as it is completely based on the smallest building blocks – learning objects.
- The approach is more flexible, because any top-down generation depends on both an existing learning object and an predefined conceptual network, which can be seen as the applications overall structure. The static learning object has just been replaced by a fixed mapping of a knowledge domain to a hierarchy of concepts. On the contrary, the bottom up approach only depends on an existing media object base and "calculates" all the other layers, so that even the structure of the application can be adapted to the current user.
- The conceptual network used by the top down approach reflects the provider's model of knowledge domain, and not the users'
- The bottom up approach is better suited for converting existing printed texts into multimedia teachware, as it just depends on learning fragment, that can be easily be generated from text paragraphs, diagrams, and images.

6. Summary

The presented proposed framework is a first approach that we believe will help in improving selecting and composing the fragment by using the algorithm used in TOD. Future work would include on how to get the best match of learning fragment which we suggest to rank the metadata with certain value to make the retrieval process more precise and to generate full automation of metadata.

References

A. Ip, A. Young, I. Morrison. *Learning Objects-Whose are They?* Proceedings of the 15th Annual Conference of The National Advisory Committee on Computing Qualifications. ISBN 0-473-08747-2 p 315-320 : Melbourne. (2002)

A. Ip, I. Morrison, M. Currie. *What is a Learning Object Technically?* WebNet 2001 : Orlando, Florida. (2001)

A. Suppaptnarm, S. Ahmed. *E-Learning From Knowledge and Experience Capture in Design. Engineering Design Centre*, Cambridge University : UK.

J. Caumanns. *A Bottom-Up Approach to Multimedia Teachware*, Proceedings of ITS '98, LNCS 1452, p 116-125. (1998)

D. Grant Campbell. *The Use of the Dublin Core in Web Annotation Programs*. Proceeding International Conference on Dublin Core and Metadata for e-Communities, Firenze University Press, p 105-110. (2002)

E. R. Jones. *Implications of SCORM and Emerging E-Learning Standards On Engineering Education.* Proceedings of the 2002 ASEE Gulf-Southwest Annual Conference, The University of Louisana : Lafayette. (2002)

J. Vassileva, D. Ralph. *Dynamic Course Generation on the WWW*. British Journal of Education Technologies, p 5-14. (1998)

K. Nakabayashi. *Trends of e-Learning Standardization*. Workshop on Development of New Training Technology Application. : OVTA Center, Chiba, Japan. (2001)

M. Magee, N. Alliance, D. Norman, J. Wood, R. Purdy, G. Irwin. *Building Digital Books with Dublin Core and IMS Content Packaging.* Proceeding International Conference on Dublin Core and Metadata for e-Communities, Firenze University Press, p 91- 96. (2002)

P. R. Lakshmi Eshwari, S. Thomas, D. Miara, N. Sarat Chandra Babu. *XML and Component Technology in e-Learning Tool Development.* Centre for Development of Advanced Computing : Hyderabad.

S. Hollfeder, R. Stecher, P. Frankhauser. *Tailoring Training Courses Using XML-based Metadata. FhG-IPSI, Darmstadt* : Germany.

T.Apostolopoulus, Kefala A. *An XML-based E-Learning Service Management Framework*. Proceedings of the IEEE International Conference on Advanced Learning Technologies. (2004)

T. Newman. *What Is Scorm?*. Communication Department, Information Transfer : Burleigh House, Cambridge. (2002)

T. L. Wentling, C. Waight, J. Gallaher, J. L. Fleur, C. Wang, A. Kanfer. *E-Learning-A Review of Literature*. Knowledge and Learning Systems Group. University of Illinois : Urbana-Champaign. (2000)

V. Theoktisto, A. Bianchini, E. Ruckhaus. *AVANTE: An architecture of CORBA components and XML Metadata for Web Based Instruction*. Department Computacion, Universidad Simon Bolivar : Venezuela.

Y. Dehua, L. Hongying, L.Zhangxi. *Application of LOM and XML in Web-Based Distance Learning.* Tongji University, Shanghai 200092, China.

Z. Shen, Y. Shi and G. Xu. *A Learning Resource Metadata Management System Based on LOM Specification.* Computer Science Department, Tsinghua University.

URL :

http://dublincore.org.

http://ltsc.ieee.org/doc/wg12/LOM3.6.html

http://www.xmlfiles.com

http://www.w3schools.com/xml

http://xml.coverpages.org/conf.html

http://www.learningcircuits.org/2002/dec2002/shank.htm