DETERMINING IMPORTANT METADATA FOR ACCESSIBILITY AND REUSABILITY OF LEARNING OBJECT

Siti Fadzilah Mat Noor, ^1^Norazah Yusof, ^2^Siti Zaiton Mohd Hashim
Software Engineering Department
Faculty of Computer Science & Information System
Universiti Teknologi Malaysia.
^1^fadzilah@ftsm.ukm.my
^2^norazah@utm.my
^3^sitizaiton@utm.my

ABSTRACT

The research in WBE (Web-based Education) systems is centred in reusability, accessibility, durability and interoperability of didactic materials and environments of virtual education. Most of the educational materials located in the web environment or learning portals are in the form of digital objects. These digital objects are also known as learning objects as they are used specifically for learning purposes in the educational environment. Most of the researcher in WBE agreed that by reusing learning object will save time and cost, instead of repeatedly authoring it and also enhances the quality of digital learning experiences. In order to easily access and reuse learning object, metadata is used particularly to describe the learning object. Description of the learning object may enables user to access and finds the right learning object for the right job. For the purpose of reusing the learning object, an experiment has been done to extract metadata element from the learning object. The result shows that the most frequently used metadata elements in the learning object are title, keyword, type, format and identifier. These important metadata elements in the learning object will be used for accessing and reusing the learning object. In order to extract the metadata, extraction method is used on content analysis of the learning object.

Keywords : Learning Object, Metadata, Reusability, Accessibility.

INTRODUCTION

The term Learning Object (LO) is one of the main research topics in the e-learning community in the recent years. LO in the educational field is considered as any kind of material that can be reused in teaching, such as a lesson plan, video, or section of a program code (Spalter & Dam, 2003). LO can also be in any digital or non-digital objects and used or reused for the technical support of learning (Friesen, 2001), which is described by a metadata record. Metadata records provide information about the object and its prospective educational usages. Learning object metadata is thus the key to reuse (Elena et al 2006). In this respect, most researchers pay attention to the issue of LOs’ reusability (Duval & Hodgins 2003, Motelet et al 2006). The most obvious motivation is the economic interest of reusing learning material instead of repeatedly authoring it and also enhances the quality of digital learning experiences (Mohan & Brooks 2003, Koutlis et al 1999, Duval & Hodgins 2000). Other motivations can be found in the pedagogical area since learner-centric teaching theories invite instructors to use a wide variety of didactic material (Motelet et al 2006).

BACKGROUND OF RESEARCH

The reality now doesn’t show such repository of LO allow of reuse learning object but more into shared LOs in coarse-grained form (Friesen 2004, Najjar 2003, Najjar 2004, Kabel 2003). In an empirical study of MERLOT (Multimedia Educational Resource for Learning and Online Teaching) repository (Heyer , 2005), Heyer found that the majority of learning resources integrate several information objects and educational components in a fixed, immutable way, which implies that the degree of reusability is extremely low. In order to make it reuse, the learning object have to be in a fine-grain form because raw media elements are often much easier to reuse then aggregate assemblies. In other words, as the LOs size decreases (lower granularity), its potential for reuse increases (Verbert et al 2005 & Duval & Hodgins 2003). Currently the problem with having fine-grained form LO in the existing learning object repository hardly exists; therefore this research is looking into that. Making LO as small as possible allows them to be easily reused without change, or with minimal change, and to be
combined in a variety of ways in other applications. This helps increase speed and efficiency of instructional design and provides ultimate flexibility. In order to access and reuse LO, metadata is used particularly to describe the LO. One of the key issues in reusing LO is each reusable LO (RLO) must have a description that enables user to access and find the right RLO for the right purpose (Valderrama et al 2004). According to Smith (2004), the use of appropriate metadata in LO ensures that LO can work with other LO, user can use the LO in different ways, and the ownership and attribution rights remain attached to LO’s, no matter where it goes or how it is used. This paper will focus on metadata extraction from collected LO to determine important metadata elements that are mostly used in LO.

**Metadata Standard for Educational Resources**
The purpose of metadata is to facilitate search, evaluation, acquisition and use of resources (IEEE, 2001). Furthermore, in order to reuse digital content (LO), interested user must be able to identify and locate relevant documents. This requires descriptive data which generally referred to as metadata. Each LO must have a metadata, the definition about the LO that enables content developers to search and to reach related LO (Wiley, 2000).

Metadata for educational resources aims at effectively facilitating the retrieval and reuse of LO. There are two well known metadata standard used for educational resources which are IEEE Learning Object Metadata (LOM) and Dublin Core (DC). IEEE LOM was released as IEEE 1484.12.1 in June 2002 (IEEE, 2005). DC was approved by the American National Standards Institute in September 2001 as ANSI/NISO Z39.85 and ratified by the International Standards Organization in January 2003 as ISO 15836 (Dublin Core, 2003).

The LOM specification contains almost 77 elements. These elements cover various aspects of the learning material including data about general content, life cycle, meta-metadata, technical characteristic, educational usage, rights, relations, annotation, and classification. The Dublin Core Metadata Element Set includes: title, creator, subject, description, publisher, contributor, date, type, format, identifier, source, language, relation, coverage and rights (Dublin Core, 2003). Based on study by Heath et al (2005), although LOM elements are useful in describing learning resources but the most useful ones are also in (and mappable to) DC. Table 1 shows LOM and corresponding DC elements. The next section will further discuss extraction method based on content analysis by using DC.dot, a web based application system to extract important elements in LO based on Dublin Core metadata.

<table>
<thead>
<tr>
<th>LOM No.</th>
<th>LOM Element Name</th>
<th>DC mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>General</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>Title</td>
<td>dc.title</td>
</tr>
<tr>
<td>1.3</td>
<td>Language</td>
<td>dc.language</td>
</tr>
<tr>
<td>1.4</td>
<td>Description</td>
<td>dc.description</td>
</tr>
<tr>
<td>1.5</td>
<td>Keyword</td>
<td>dc.subject</td>
</tr>
<tr>
<td>2</td>
<td>Lifecycle</td>
<td></td>
</tr>
<tr>
<td>2.3.2</td>
<td>Entity</td>
<td>dc.creator,dc.publisher</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Date</td>
<td>dc.date</td>
</tr>
<tr>
<td>3</td>
<td>Metametadata</td>
<td></td>
</tr>
<tr>
<td>3.1.2</td>
<td>Entry</td>
<td>dc.identifier</td>
</tr>
<tr>
<td>4</td>
<td>Technical</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>Format</td>
<td>dc.format</td>
</tr>
<tr>
<td>4.3</td>
<td>Location</td>
<td>dc.identifier</td>
</tr>
<tr>
<td>5</td>
<td>Educational</td>
<td></td>
</tr>
<tr>
<td>5.2</td>
<td>learningresourcetype</td>
<td>dc.type</td>
</tr>
<tr>
<td>6</td>
<td>Rights</td>
<td></td>
</tr>
<tr>
<td>6.3</td>
<td>description</td>
<td>dc.rights</td>
</tr>
<tr>
<td>7</td>
<td>Relation</td>
<td></td>
</tr>
<tr>
<td>7.2.1.2</td>
<td>Entry</td>
<td>dc.source,dc.relation</td>
</tr>
<tr>
<td>9</td>
<td>Classification</td>
<td></td>
</tr>
<tr>
<td>9.2.2.2</td>
<td>Entry</td>
<td>dc.subject</td>
</tr>
</tbody>
</table>
Metadata Extraction Methods
There are two metadata extraction methods which are generation techniques based on content analysis and generation based on context analysis (Motelet et al, 2006). For this paper we are focusing on metadata extraction technique based on content analysis by using DC.dot, a web based application generating DublinCore metadata for a web resource (DCMI, 2005). DC.dot is chosen because as mentioned earlier, the most useful metadata elements are also in DC. Generic content analysis is useful to generate keywords or description. More specific characteristics like title, author, size, creation date, etc. are generally retrieved from already existing metadata embedded in the educational resource itself. This method is used by DC.dot. Although DC.dot will generate metadata based on Dublin Core, it can however be changed into IEEE LOM Standard format.

EXPERIMENT
An experiment was conducted by using DC.dot, a generator developed by UKOLN (UK Office for Library and Information Networking) based at the University of Bath. Metadata creation with DC.dot is initiated by submitting a URL of LO. Figure 1 shows an example of the metadata extraction using DC.dot.

![Metadata Extraction Using DC.dot](image)

Figure 1: Metadata Extraction Using DC.dot

Data Set & Workflow
In this experiment, the data set consists of 200 web based LO collected from existing learning object repository such as MERLOT (Multimedia Educational Resource for Learning and Online Teaching) and SMETE (Science, Mathematics, Engineering, and Technology Education) that comprise subject in computer science and biology. The steps for extraction metadata of a web based LO are as follows:

i. By submitting URL of the LO, the system will extract metadata elements from the web based LO. The metadata attribute checking process is carried out.
ii. The system copies resource ‘identifier’ metadata from the Web browser’s address prompt, and harvests ‘title’, ‘keywords’, ‘description’,” and ‘type’ metadata from resource META tags. System also automatically generates ‘type’, ‘format’ and ‘date’ metadata. Date is automatically track by system from the source code programming.
iii. The system will display all the metadata elements that have been extracted.
iv. Finally all the metadata elements will be kept in the database.

Metadata Elements Analysis
Metadata elements analysis is conducted to determine the most important metadata elements in LO. From the data set consists of 200 web-based LO, an analysis is carried out and the results are presented by illustrating the frequencies and percentage of metadata elements usage in LO.
Figure 2 shows that out of 15 elements in Dublin Core Metadata there are only 11 elements which are commonly used in LO namely title, author, keywords, description, publisher, date, type, format, identifier, language and right-management. The missing metadata elements such as contributor, source, relation and coverage suggest that these elements are not quite important in describing the LO.

Figure 2: Frequencies of metadata elements in data set

Figure 3 ranks the metadata elements used in LO. Based on Figure 3, the most frequently used metadata element in LO is the format, whereby 97% of LO use this element. In general, there are 5 metadata elements which have a percentage of usage of more than 90% and they
CONCLUSION

The experiment shows that the most frequently used metadata elements are Title, Keyword, Type, Format and Identifier. Although the data sets are limited to 200 web based LO, this finding may assist us to select important metadata elements and this metadata element will ensure that ownership and attribution rights remain attached to LO’s. Even though existing LO may be used for a different purposes but the original metadata will be retained to ensure the origin of LO remain attached.

Unfortunately, metadata elements that have been identified were not enough to describe the LO. Apart from format, identifier (URL of the LO), title, type and keyword another two metadata elements namely author and description were quite important to remain identity of the LO but the lack of available metadata to describe the LO makes it very difficult to share and reuse these objects (Ochoa et al. 2005). Once the LO has been granules into text, image and other form; process of tagging granule-LO with the identified metadata elements that have been extracted before will be carried out. Future research will look into utilizing intelligent-techniques to implement the tagging process in order to access and reuse the existing LOs.

REFERENCES

DC.dot http://www.ukoln.ac.uk/metadata/dcdot/
MERLOT (Multimedia Educational Resource for Learning and Online Teaching) http://www.merlot.org


SMETE (Science, Mathematics, Engineering, and Technology Education)
http://www.smete.org


