The Development of Metadata Standards for Documents in the Teaching Domain — A Case Study in Taiwan

FU-CHING WANG[1], TIMOTHY K. SHIH[2], PAO-TA YU[3] AND MING-TSUNG LIU[3]
[1]Information Service Center
Industrial Technology Research Institute
[2]Department of Computer Science and Information Engineering
Tamkang University
[3]Department of Computer Science and Information Engineering
National Chung-Cheng University
168, University Road, Ming-Hsiung Chia-Yi, TAIWAN

Abstract: - Metadata (data about data) provides a common nomenclature to describe learning resources in a common way. The unification of metadata standard provides unambiguous way for describing the same attribute of learning resources. However, there exist indeed cultural differences among different nations and areas. Metadata standards should preserve enough flexibility to deal with cultural differences. In this paper, two metadata standards, Teaching Material Markup Language (TMML) and Metadata Lifecycle Model, adapted from SCORM and Dublin Core are introduced to overcome this problem.

Key-Words: Metadata, Standard, SCORM, Dublin Core, Cultural Differences

1 Introduction

Metadata (data about data) provides a common nomenclature to describe learning resources in a common way. Metadata can be collected in catalogs, as well as directly packaged with the learning resource it describes. Learning resources described with metadata can be systematically searched for and retrieved for use and reuse.

For example, Sharable Content Object Reference Model (SCORM) [1] defines a standard set of metadata element definitions that can be used to describe learning resources. Metadata are machine-understandable which means computers can search and retrieve learning object according to specific needs. Moreover, intelligent system such as agent-based system can utilize the information acquired from metadata to identify, recognize and analyze learning recourses [2][3].

The unification of metadata standard provides unambiguous and common way for describing the same attribute of learning resources. However, there exist indeed cultural differences among different nations and areas. Metadata standards should preserve enough flexibility to deal with cultural differences.

In this paper, two metadata standards, Teaching Material Markup Language (TMML) [4] and Metadata Architecture and Application Team Metadata (MAAT), adapted from SCORM and Dublin Core are introduced to overcome this problem.

2 Metadata Standards

In this section, two well-known metadata standards, SCORM and Dublin Core, and their evolution are discussed.

2.1 SCORM Metadata

Metadata for learning resources has been under development within a number of national and international organizations over the past few years. Advanced Distributed Learning (ADL) initiative references IEEE LTSC [5] Standard for Information Technology -- Education and Training Systems -- Learning Objects and Metadata (LOM) Working Group, the IMS [6] Global Learning Consortium, Inc. and the Alliance of Remote Instructional Authoring and Distribution Networks for Europe (ARIADNE) [7] as the bodies that are defining metadata specifically for learning resources. These groups, which have been working collaboratively, have developed a core set of specifications.

The SCORM references the IEEE LTSC LOM Standard. The LOM was developed as a result of a joint effort between the IMS Global Learning Consortium, Inc. and the ARIADNE to define a standard set of metadata element definitions that can be used to describe learning resources. SCORM has

adopted the same set of metadata elements described in the IEEE LTSC LOM Standard. SCORM will also reference a binding specification at such time that the binding specification becomes available. The binding specification will provide an XML representation for the IEEE LTSC LOM Standard.

SCORM applies the IMS metadata element definitions to three content model components: Asset, SCO and Content Aggregation. These three components define the metadata portion of the SCORM Content Aggregation Model.

This mapping of standardized definitions from IEEE to the SCORM Content Aggregation Model provides the missing link between general specifications and specific content models. The following sections define the application of IEEE definitions to the metadata portion of the SCORM Content Aggregation Model.

2.2 Dublin Core

Ongoing efforts of Dublin Core Metadata Initiative (DCMI) participants include the collaborative development and continual refinement of metadata conventions based on research and feedback between DCMI Working Groups.

The Dublin Core standard includes two levels: Simple and Qualified. Simple Dublin Core comprises fifteen elements; Qualified Dublin Core includes three additional elements (Audience, Provenance and RightsHolder), as well as a group of element refinements (also called qualifiers) that refine the semantics of the elements in ways that may be useful in resource discovery.

The Simple Dublin Core Metadata Element Set (DCMES) consists of 15 metadata elements as shown in Table 1.

Table 1: Simple Dublin Core Metadata Element Set

1	Title	2	Creator	3	Subject
4	Descriptio	5	Publisher	6	Contributo
7	Date	8	Type	9	Format
10	Identifier	11	Source	12	Language
13	Relation	14	Coverage	15	Rights

Each Dublin Core element is optional and may be repeated. The DCMI has established standard ways to refine elements and encourage the use of encoding and vocabulary schemes. There is no prescribed order in Dublin Core for presenting or using the elements.

3 Metadata Standards in Taiwan

Apply adaptively metadata standards in countries with different culture and educational background is a challenging issue for governments all over the world. Most solutions such as Cancore (Canada), UK LOM Core (United Kingdom) and CELTS-3 (China) are based on IEEE LOM standard with application profile.

In Taiwan, several research projects have been launched to focus on modifying metadata standards so that they can adapt to cultural differences without losing compatibility with original standards. Two of the research project outcomes are mentioned in this section. Some recommendations are also concluded for researchers attempt to adapt metadata standards.

3.1 Teaching Material Markup Language (TMML)

TMML is developed by a project named "A Study of Recommending the Standard Format for e-Learning Systems, Platform, and Content" launched in 2002. This project, sponsored by National Science Council (NSC) in Taiwan, had the mission to localize and customize the metadata standard for the educational society in Taiwan.

TMML divides metadata elements into two levels, generic metadata level and specific domain metadata level. The architecture of TMML is illustrated in Fig. 1. The first level, generic metadata level, defines metadata elements commonly used in general domain. While in the second level, specific domain metadata level, metadata elements applied to describe learning resources in specific educational domain

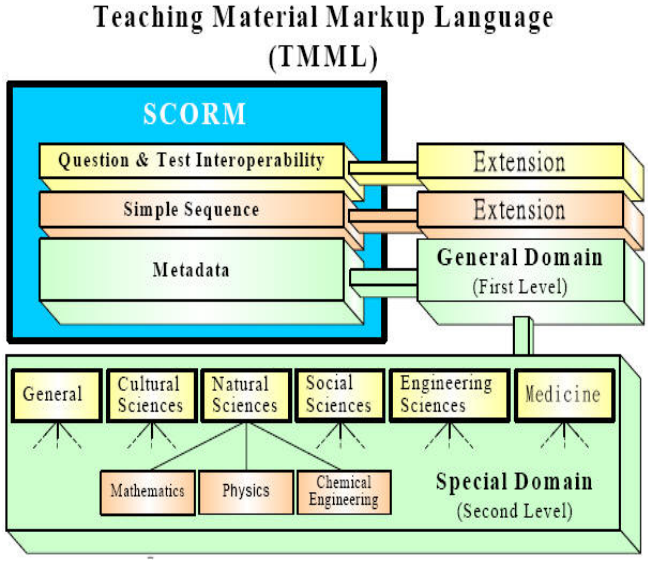

Fig. 1: The architecture of TMML

All the metadata elements are divided into fourteen categories, nine categories inherited from SCORM Metadata and five expanded categories adapted from IMS Question & Test Interoperability

Specification and IMS Simple Sequencing Specification. Further information about TMML and the download service for specification related documents can be found in the following website: http://e-learning.nctu.edu.tw.

3.2 Metadata Architecture and Application Team Metadata (MAAT)

The "National Digital Archives Program" (NDAP) project in Taiwan was launched on January 1st, 2002 and is sponsored by the National Science Council (NSC). The aim is to promote and coordinate content digitization and preservation at leading museums, archives, universities, research institutes, and other content holders in Taiwan.

Before the NDAP was launched, years were spent studying the impact of Information and Communication Technologies (ICT) on socio-economics. These studies lead us to believe that mankind is facing critical cultural and social change, including changes in the means of communication from printed matter to electronics, and other changes to the way of life in society. ICT has advanced to a level that affords us the opportunity to digitize our cultural treasures and heritage, so they can be preserved and utilized in the digital era. Otherwise, some of them might gradually disappear and possibly become extinct.

MAAT, one of the technical teams in NDAP, develop Metadata Lifecycle Model to standardize the process of defining metadata for specific domain such as content digitization and preservation. As shown in Fig. 2, there are four phases in the Metadata Lifecycle Model.

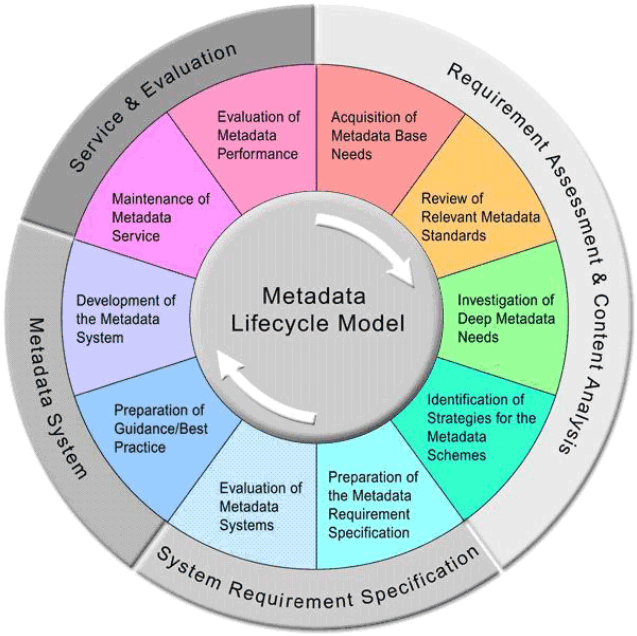

Fig. 2: Metadata Lifecycle Model

In the Requirement Assessment & Content Analysis Phase, The first step of the metadata lifecycle is to interview the content experts or providers about their metadata requirements for each collection project, and to analyze the attributes of collection projects. Then relevant metadata standards are reviewed and deep metadata needs are investigated. The last step is Identification of strategies for the metadata schemes and achieving interoperability with Well-known Metadata Standards.

In the System Requirement Specification Phase, the preparation of the metadata requirement specification and evaluation of metadata systems are performed. This stage involves the evaluation of potential metadata systems. The collection project members can select an existing system developed by homogeneous or similar collection projects.

In the Metadata System Phase, the preparation of best practice guidance and development of the metadata system are performed. System developers develop metadata tools and systems based on the metadata requirement specification.

In the Service & Evaluation Phase, maintenance of metadata service and evaluation of metadata performance are performed. The last stage of the MLM seeks to review results of the whole metadata process and performance. The evaluation is conducted according to the assessment of metadata record quality, the effectiveness of adopting a metadata scheme for retrieval, the use of metadata creation tools within the collection project, and the application of the Metadata Lifecycle Model in each stage.

4 Conclusion

After analyzing the development of Metadata Standard in Taiwan, we deem that two important properties of Metadata Standard should be maintained during the development phase of metadata specification — layered structure and well-defined lifecycle.

Layered structure divides the elements into layers which preserves the most flexibility. Applications can choose proper level of detail in the metadata standard after evaluating the basic and deep needs. Fig. 3 illustrated the layer structure we recommended. The bottom layer is Generic Metadata Layer whose elements can be used to describe general resources. Metadata elements in this layer are designed for general purpose and compatible with Internet and Library Resource Metadata [8]. The middle layer is Learning Resource Metadata Layer

designed for describing learning resources. The top layer is Specific Domain Metadata for learning resources in specific domain including elementary education, high school education, college education, and lifelong education.

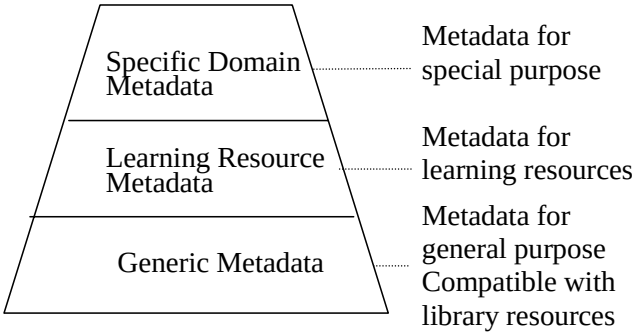

Fig. 3: The Recommended Layered Structure for Metadata Standard

Metadata is an emerging approach to organizing structured digital collections, in order to support precise retrieval, long-term preservation, and interoperability [9] on an extraordinary Internet scale. Although there are many metadata practices in digital libraries, few literatures have addressed the methodology of a best practice for developing metadata. In the light of metadata provision, digital library projects often face a series of issues, including: how to get started, how to acquire metadata needs, how to choose a suitable metadata standard and adopt it, how to develop metadata specification, how to evaluate a metadata system, and so forth. A set of effective methods to develop metadata is thus very important.

Analysis	● Investigation of metadata need ● Review of relevant standard ● Best practices
Design	● Metadata element acquiring and expanded ● Assessment of the investigation result ● Consulting domain experts ● Vocabulary
Test	● Prototype implementation ● Test-bed ● Feed back and suggestion
Implementation	● Released version ● System and tool develop
Evaluation	● Integration evaluation ● Report ● Technical support service

Fig. 4: The Recommended Lifecycle for Metadata Standard

As shown in Fig. 4, five stages of the lifecycle are analysis, design, test, implementation and evaluation. The metadata lifecycle list here is an evidence-based approach that normalizes metadata requirements, has been proved to achieve efficiency, quality assurance, and consistency by providing a systematic way to develop metadata systems.

References:

[1] ADL, “Sharable Content Object Reference Model Version 1.2”, http://www.adlnet.org, 2001.
[2] Dieter Fensel, “Ontology-Based Knowledge Management”, 2002.
[3] Stefan Decker and Sergey Melnik and Frank Van Harmelen and Dieter Fensel and Michel Klein and Jeen Broekstra and Michael Erdmann and Ian Horrocks, “The Semantic Web: The roles of XML and RDF”, 2000 October.
[4] Teaching Materials Markup Language (TMML) http://e-learning.nctu.edu.tw/TMML/index.htm
[5] IEEE Information Technology - Learning Technology - Learning Objects Metadata LOM: Working Draft v6.1 (2001-05-03) http://ltsc.ieee.org/.
[6] IMS Learning Resource Metadata Specification Version 1.2. http://www.imsglobal.org/.
[7] Alliance of Remote Instructional Authoring & Distribution Networks for Europe (ARIADNE) http://www.ariadne-eu.org/
[8] The Library of Congress, “Library of Congress Classification,” http://lcweb.loc.gov/catdir/cpso/lcco/lcco.html
[9] Margaret St. Pierre, “Issues in Crosswalking Content Metadata Standards”, http://www.niso.org/, 1998.
[10] World Wide Web Consortium (W3C), Resource Description Framework (RDF) Model and Syntax Specification, http://www.w3.org/
[11] Michael Brennan, “The Learning Learning Content Management System: A New eLearning Market Segment Emerges”, http://www.idc.com
[12] Aviation Industry CBT (Computer-Based Training) Committee (AICC) (http://www.aicc.org)