

Multimedia Portals based on Semantic Web Technology for GENES Communities - The Implementation of LOM-RDF Conversion Tools –

Susumu Sawai ¹, Toshiko Wakaki ², Masakazu Tachikawa ², and Susumu Kunifuji ³

¹ Gakujoken, 23,NangenCho,ShinjikuKu,Tokyo,
160-0012,Japan

NAC02440@nifty.com

<http://www.gakujoken.or.jp/index.html>

² Shibaura Institute of Technology, 307 Fukasaku,Saitama-City, Saitama,
330-8570 Japan

{twakaki, p02093}@sic.shibaura-it.ac.jp

³ School of Knowledge Science, JAIST, 307 Fukasaku,Saitama-City, Saitama,
330-8570 Japan

kuni@jaist.ac.jp

Abstract. We proposed new creation tool[1] for GENES communities, and implement semantic information of Semantic Web technologies, which were primarily designated for a computer to interpret information on the Web, in enabling the obtainment of a user's intended search results, and to achieve the retrieval of all LOM data available on the Web through its interoperability.

The aims of GENES project are to realize the multimedia portals based on Semantic Web technology where learners, especially children should be able to use digitalized cultural heritage for education, as a "Digital Cultural Genes".

1 Introduction

Today, as the utilization of learning content expands into various areas of our society, the importance of e-Learning is increasing even more significantly.

In order to backup "IT new reform strategy" in Japan, we proposed new creation tool [1] and implemented semantic information of Semantic Web technologies [2], which were primarily designated for a computer to interpret information on the Web, in enabling the obtainment of a user's intended search results, and to achieve the retrieval of all LOM data available on the Web through its interoperability.

The Ministry of Education, Science, Sports and Culture (MEXT) founded National Information Center for Educational Resources (NICER) in August 2002. NICER is the multimedia portals that has the roles to arrange and manage all the information about education and learning in Japan on the Internet.

GENES communities is the multimedia communities for researching cultural heritage and so on. GENES Project was launched in August 2001, in cooperation with Fujitsu and MEXT. GENES communities stands for Gakujoken network study-group. Gakujoken is the public organization managed by MEXT. The roles of Gakujoken is to collect and to offer the educational contents to learner.

2 Objective of GENES

The objective of GENES communities is to create the specific communities where digitalized cultural heritage contents are made by teachers in real school. The aims of GENES project are to realize the multimedia portals based on Semantic Web technology where learners, especially children should be able to use digitalized cultural heritage for education, as a "Digital Cultural Genes". Today the profile of the project to realize GENES communities is as follows:

- Cooperation with more than 1,200 school teachers.
- More than 10,000 learning object metadata (LOM) for educational contents.
- More than 38,000 educational contents (motion picture, photos and texts) are combined in this digital-archives (archives). These educational contents contain the following specifications: a school grade, a subject and a unit, a coverage person, a coverage date, a coverage place, copyright, etc.

3 Applying tacit knowledge to Web page Search Systems

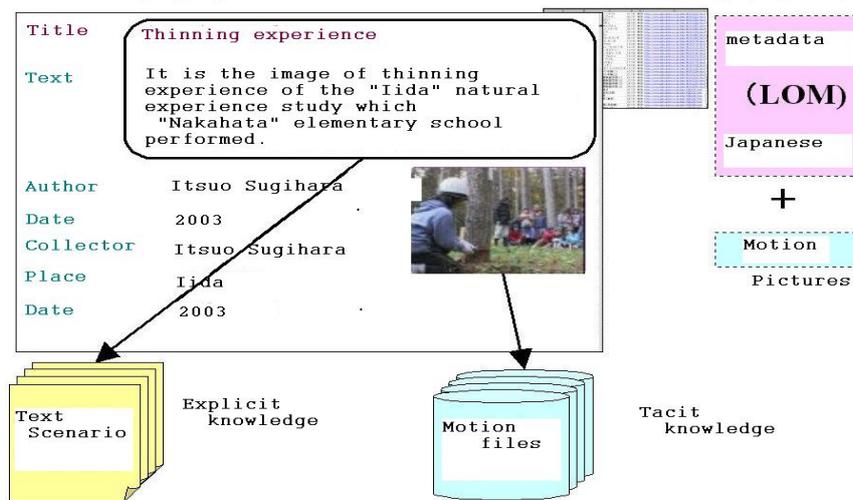


Fig. 1. New Knowledge Representation

Advantages of implementing the new knowledge representation based on the combination of motion pictures, natural language and metadata are that it enables text-search within the realm of tacit knowledge [3] derived from still or motion pictures.

Web pages targeted in this research consist of text scenarios of explicit knowledge and still or motion digital image files of tacit knowledge. The explanation of the motion file is in the text scenario database.

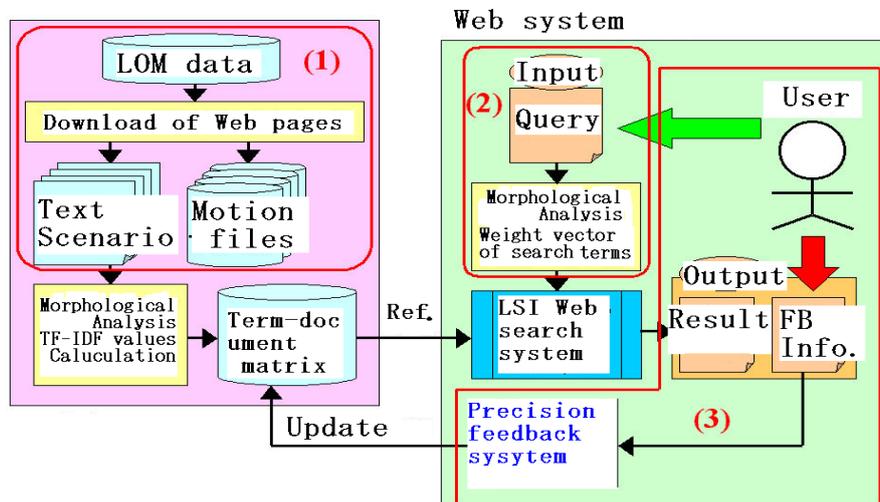


Fig. 2. Web page Search Systems

The benefits of this are that sharing methods utilizing web-page search systems will become available, for example the application of tacit knowledge through motion pictures to e-learning.

Web pages similar to the imputed query will be searched within the database created by the text scenario of web pages, using Latent Semantic Indexing (LSI) [4], [5] as a search method. Learning Object Metadata (LOM) which enables search for digital archives will serve as an interface for the actual search. This will enable the application of NICER to the search system.

In addition, this will help applying tacit knowledge from motion files to e-learning as well as conducting feedback searches.

Search criteria should be the sentences close to natural language. This will enhance universality and contribute to raising the system's utility value.

4 RDF Expressions of LOM

LOM (Learning Object Metadata) is an IEEE standard (approved in 2002) consisted of 96 elements regarding the definition of structures and word assignment methods concerning learning objects.

When describing an LOM in RDF expression, the URL of the learning material archive shall be set as the nominative, the attributes of each LOM as the property, the content corresponding each attribute as the object. The property may be diverted from an existing metadata or uniquely defined using RDFS.

4.1 Utilizing Existing Metadata

We have primarily used the Dublin Core [5] in cases where we utilized an existing metadata. The Dublin Core is a metadata set out for the description of bibliographic information, consisted of 15 basic elements and numerous extended elements.

4.2 Unique Definition Using RDFS

In cases where existing words could not be used, the properties and classes will be defined by new entry. Figure 3 is an example of the expression of the thumbnail item, and the property is defined in RDFS. As a thumbnail is an image, the property is defined by using `dct:Image` (a Dublin Core-typed element, a class indicating non-textual, visual symbolic expressions) in the property value range, manifesting that the object is an image.

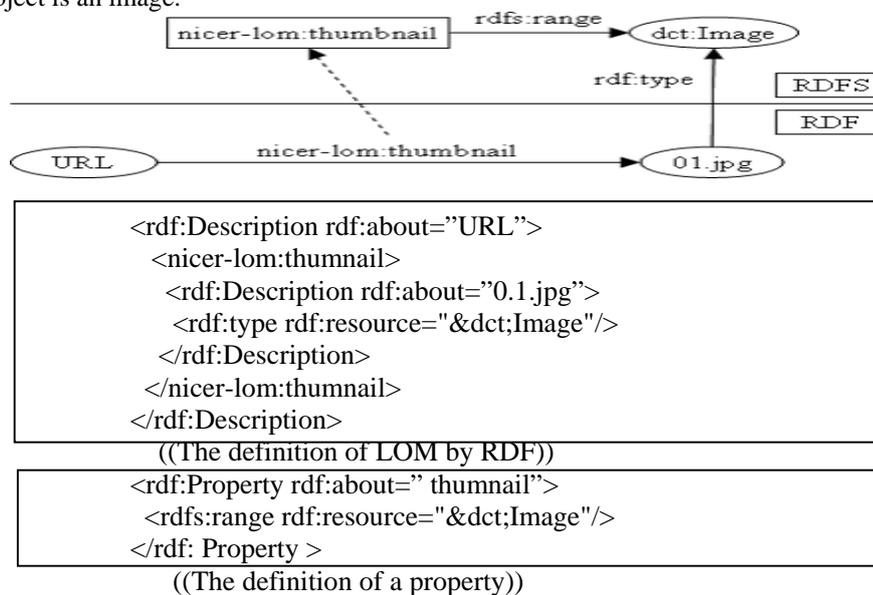


Fig. 3. An example of new definition using RDFS

4.3 Course of Study

The course of study is a curriculum standard announced by MEXT, which sets down the content and the details of the studies to be taught through elementary school to high school. The NICER LOM expresses the course of study in a tree structure, indicating the School Name, Subject, Item and so on. In RDF/RDFS expressions of LOM, the relationships between school types and subjects in adjacent sectors are defining their class with the organizing subjects segmented in upper level schools such as high schools.

5 LOM-RDF Conversion Tool

There are two prominent types of methods in converting LOM to RDF expressions. One is to apply the content to a RDF template in which items allowing free input such as titles, summaries are predefined. The other is to convert selected items such as course of studies and language into RDF expressions corresponding to the item content. Figure 4 illustrates the system block diagram of LOM-RDF conversion tool .

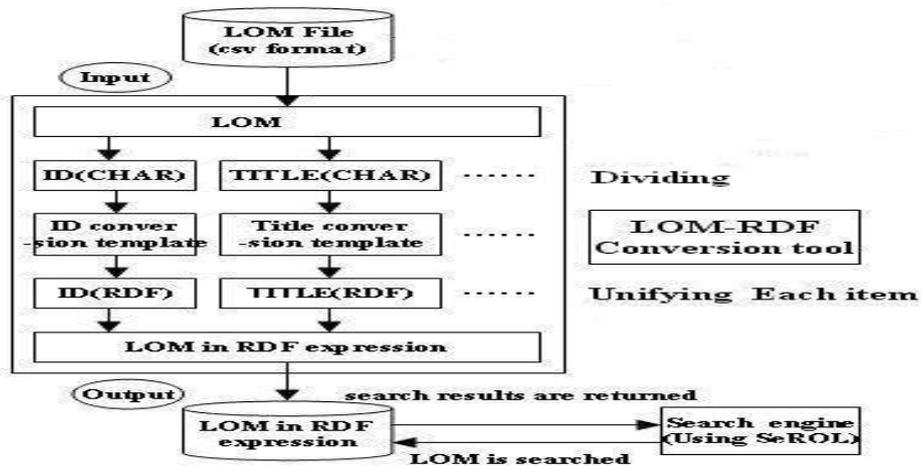


Fig. 4. System block diagram of LOM-RDF conversion tool

6 RDF Searches and Assessments

We used a tool called Sesame [6] in order to assess the usability of the search against learning material archives using LOM-RDF expression. Sesame handles RDF as a database, conducting search using an SQL-like language named “SeRQL”.

The adequacy of searches were assessed using 400 LOM records of learning material archives belonging to the subject “social studies” throughout elementary school to high school. Tables1 shows the comparison of the results of searches conducted against elementary school-level civics’ archives using different criteria. The conditions were set as follows:

Condition D: Course of study [elementary school/ social studies/ 6th grade]

Condition E: Course of study [subjects in the area of history] and Target age [6th grade]

Condition F: Course of study [subjects in the area of civics] and Target age [6th grade]

Super-ordinate concepts were used for the criteria concerning “course of study” in conditions E and F.

More results were returned in the search using condition D. However, since history-related and civics-related subjects are both taught in 6th grade, the results of condition D include numerous archives related to the history subject (such as those on

the “Gifu Castle”), and has not been able to return the user’s desired results. At the same time, condition F indicates that elementary school-level civics-related learning material archives can be retrieved by setting the “Course of study” criteria as “civics” in abstract and narrowing down the difficulty level by setting the “Target age” criteria.

Table 1. An example of data obtained as search results (civics)

The title of a teaching-materials archive	Search conditions	
	D	E
Houkoku shrine and a garden	○	
Nagoya Castle & tower (Aichi Prefecture)	○	
virtual tour of Supreme Court	○	
Gifu castle	○	
virtual tour of Supreme Court	○	○
virtual inspection tour of Bank of Japan		○
directions of shopping center reproduction		○
numbers of relevance of search conditions	170	27

7 Postscript

RDF searches have proved to provide appropriate search results to a certain extent. However, due to the yet limited number of LOM and limited search target subjects, it is necessary that we assess the results against a larger parameter of LOM and subjects. Additionally, the assessment in this paper does not consider the detailed content of the learning material archives, therefore such will also be taken into consideration on conducting further assessment.

In order to backup "IT new reform strategy" in Japan, we proposed new creation tool for GENES communities, and implement semantic information of Semantic Web technologies. And we emphasize that learner, especially children should be able to use every digital Cultural Heritage for education, as a "Digital Cultural Genes".

References

1. Susumu Sawai, Tosiko Wakaki, Mitsuru Ikeda, Susumu Kunifuji: Ubiquitous Networking for GENES Society: e-Learning Tools and Digital Archives for Education with Significant Use of Cultural Heritage Contents, proc. of the 1st Int. Workshop on Interoperability of Web-based Educational Systems, WWW2005(2005)71-74
2. G. Antoniou, F.v.Harmelen: A Semantic Web Primer, The MIT Press(2004)61-108
3. Michael Polanyi: THE TACIT DIMENSION, Routledge & Kegan Paul Ltd., London (1966)
4. K. Kita, K. Tsuda, M. Sisibori: Information Retrieval Algorism, Kyuritsu Press(2002)51-89
5. Dublin Core Metadata Initiative: <http://dublincore.org/>
6. Sesame: <http://openrdf.org/>