

Knowledge Processing and Service System for Digital Education based on Semantic Web

Xia Li¹, YingLong Wang²

1 Humanity & Social Science College, Wuhan University of Science and
Engineering

Liniya_@163.com

2 Department of Information Technology, HuaZhong Normal University
Wangylccnu@mail.ccnu.edu.cn

Abstract. Recent efforts in the area of learning technology have resulted in a considerable improvement in the interoperability of digital education resources. On the basis of sharing and integration of digital educational resources, knowledge Processing and service considers the Semantic knowledge extraction of metadata of educational resources, Construction ontology for educational resources, and combining the users personality characteristics, providing semantic-based personalized retrieval and delivery services. In this paper, we work out the architecture of knowledge processing and service system for digital education, then discuss the semantic representation of digital education resources, semantic mining and discovery of digital education resources and personalized knowledge delivery service.

1 Introduction

Recent efforts in the area of learning technology have resulted in a considerable improvement in the interoperability of digital education resources across different Learning Management Systems (LMS) and Learning Object Repositories (LOR) [1]. On the basis of sharing and integration of digital educational resources, knowledge Processing and service considers the Semantic knowledge extraction of metadata of educational resources, Construction ontology for educational resources, and combining the users personality characteristics, providing semantic-based personalized retrieval and delivery services

Knowledge Services is the innovation of traditional service model through computer, network and multimedia technology-based information technology, which is a brand new information service models. Knowledge Service makes use of service

Please use the following format when citing this chapter:

Li, X., Wang, Y. L., 2007, in IFIP International Federation for Information Processing, Volume 251, Integration and Innovation Orient to E-Society Volume1, Wang, W. (Eds), (Boston: Springer), pp. 499-506.

description, generation, acquisition, organization, positioning, and integration to meet the demands of users to use the information. It is closely related to two areas : knowledge-related issues and service-related issues. Dealing with issues related to knowledge of the objectives is to provide knowledge services necessary for the proper and accurate knowledge. Due to the large volume of knowledge is the source rich and complex features, it is related to knowledge modeling and knowledge base of the building. Service related to the goal is to understand user demand, which in turn service through knowledge of the operation, on the Internet in an appropriate manner to meet these demands.

Semantic Web technology is able to provide the required computational semantics for the automation of tasks related to Digital Education resources as selection or delivery [1]. Ontology can play an important role in sharable semantics and concept for digital education interoperation and it can also make a knowledge base for ontology driven digital education resources retrieve and delivery

In this paper, we work out the architecture of knowledge processing and service system for digital education, then discuss the key technology involved. The paper is organized as follows: Section 2 introduces the Semantic Web and ontology, which is the key technology used for the system; Section 3 discuss the architecture of knowledge processing and service system for digital education; section 4 discusses the implementation of the key technologies in the system and provides an overview of various components of the system such as semantic representation of digital education resources, semantic mining and discovery of digital education resources and personalized knowledge delivery service; Section 5 is a brief review of related work; while conclusions are presented in section6.

2 Semantic Web and Ontology

There are a number of important issues related to the Semantic Web. Roughly speaking, they belong to four categories: languages for the Semantic Web, ontologies, semantic markup of pages on the Semantic Web, and services that the Semantic Web is supposed to provide [2].

Currently, Web content is formatted for human readers rather than programs. HTML is the predominant language in which Web pages are written. For people the html information is presented in a satisfactory way, but machines will have their problems. The semantic web representation is far more easily processable by machines. The term metadata refers to such information: data about data. Metadata capture part of the meaning of data, thus the term semantic in Semantic Web.

In computer science, the ontology could be used to build vocabulary of specific domain application that is used for a formal explicit definition of some concepts that are shared among researches. For environment of knowledge sharing, the ontology cloud be meant to “*a specification of a conceptualization*” [3]. The conceptualization is the objects, concepts and other entities that are assumed to exist in some areas of interest and the relationships that hold among them [4].

Ideally, creation of digital educational Web contents with ontological annotation should be supported by ontology-driven authoring tools and class hierarchies based on a number of underlying ontologies[2]. Digital Education resources can then be presented, edited, modified, and mixed consistently. The reality, however, is still far away from being ideal and there are a lot of further steps and efforts to make in order to move forward.

3 The architecture of knowledge processing and service system for digital education

The architecture of knowledge processing and service system for digital education is as the fig.1. The system includes two components, which are knowledge processing section and knowledge service section separately.

The section of knowledge processing consists of the following modules:

- Digital Resources Parsing Module: Parse resources collected by the various formats of digital education resources for a unified XML format;
- Knowledge Extraction Module: Extract the information about the title, disciplinary, author, keywords, such as summary metadata information from the XML documents;
- Ontology search module: Construct the ontology base for the digital education resources and provide the ontology driven retrieval for the user to find the required resource with correct location and information;

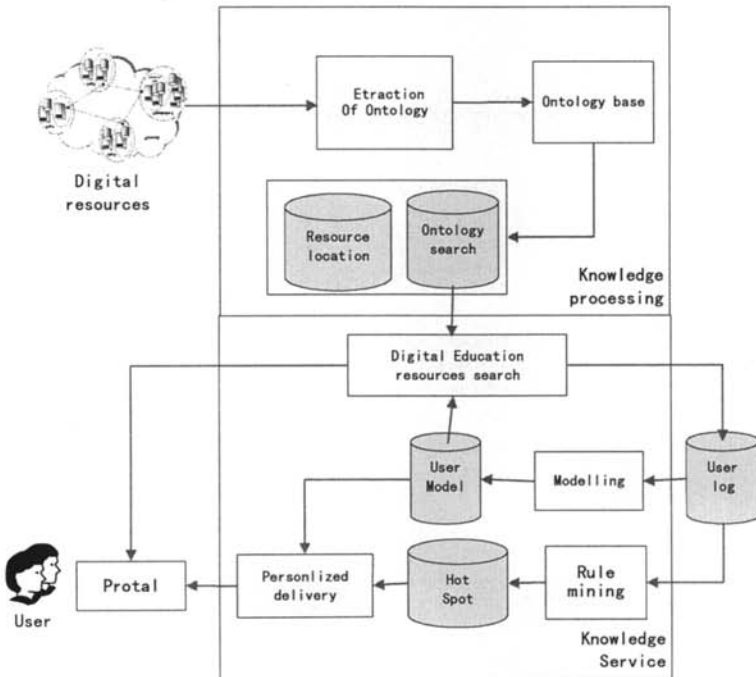


Fig. 1. The architecture of knowledge processing and service system for digital education

- The section of knowledge service consists of the following modules:
- User modeling module: Generate and adjust dynamically the user model in term of the search requests ;
 - Digital Education Resources search module: Parsing and Analyzing semantically the inquiry of user, such as keywords portfolio, metadata, knowledge ontology in the form of inquiries, According to the ontology Association to location the education resources meeting the required condition, and filtering and sorting the results in term of the user model;
 - Personalized knowledge push and delivery: Achieve the initiative, Efficient knowledge Push and delivery Service According to the user model and knowledge access character of the digital education resources though various network platform;

4 The implementation of the key technologies in the System

There are mainly three key technologies in the knowledge processing and service system for digital education, which is semantic representation of digital education resources, semantic mining and discovery of digital education resources and personalized knowledge push and delivery service.

4.1 Semantic Representation of Digital Education Resources

Extraction of ontology from the digital education resources plays an important role in the education resources semantic representation. Current mainstream search engines, such as Google, Baidu, etc can not extract ontology and metadata information from the web resources, and unable to support ontology driven retrieval resources. In the next generation search engines, metadata and ontology extraction module has become an important component of them. It can enhance the retrieval of flexibility, retrieval accuracy and the rate of return. According to the characteristics of digital educational resources and user search habits, We reference IEEE learning object metadata standard IEEE LOM (learning object metadata standard)[5], and other standards, make use of natural language processing and machine learning methods to automatically collected subjects, the title, author, keywords, Following is a brief summary of 10species of elements in ontology.

Table 1. Extraction of ontology elements for educational resources

Number	Element	
1	Identifier	Identifier of educational resource, which is unique in the world
2	Title	The name of educational resource
3	Author	Creator of educational resource

4	Abstract	Text description of the educational resource content
5	Keywords	Important descriptive word of content
6	Subject	Subject of the educational resource
7	Language	Language of the educational resource
8	Context	Main scenario using educational resource
9	Format	Technical data type of educational resource, which is used to decide the running time software required by educational resource
10	Resource Type	Concrete type of educational resources

4.2 Semantic Mining and Discovery of Digital Education Resources

Currently, the commonly used full-text index method is mainly through the text word segmentation, extracting the text contained in the feature words, to feature words for indexing projects, the establishment of words and text content of the link between indexes, the inverted index Based on the characteristics to achieve the full text word search function. However, it was found that this method the following drawbacks:

- Index files too large. Feature-based Translation indexing methods due to lack of understanding and context of broad, will introduce text in the number of message text indexing, document indexing to occupy a space too large.
- Retrieval inefficient. As the use of the word inverted index technology, With the retrieval of text and the size to the number of words and the size of the index surge, which not only affect the speed, but also the quality of inquiries and efficiency.
- Low rate of precision, high rate of repetition. Overall, such as indexing mechanism is simply to achieve a feature words and text on the link between primers, They lack the resources to the original description of semantic information (such as metadata, ontology), failed to establish friendly understanding between the three elements as "primary resources" , "ontology based indexing mechanism" and "user". it can not meet the semantic matching of user query and text content.

Based on the characteristics of the content indexing words of the aforementioned problems, we construct a "meta-data", "ontology", content and resources Association-indexing mechanism, then achieve "metadata" "ontology" and the link between digital educational resources inquiries. The model includes semantic information resources ("ontology"), to help raise resources retrieval accuracy. The same time can reduce the storage space for the occupier. The issue of resources search model involves the number of attributes resources, the need to establish an ontology driven index mechanism, hence realize the semantic mining and discovery to meet the user inquiry.

4.3 Personalized Knowledge Push and Delivery Service

Due to the differences between different users in age, sex, education, occupation, hobbies etc, the field that users are interested in is also different. Thus the results of

the same search requests have different appraisal. However, the existing search engine that the service provided is the "search for" and not "user-oriented." that is, for different users to the same search requests. Search engines provide the same search result. It is hard to meet the learner's individual needs.

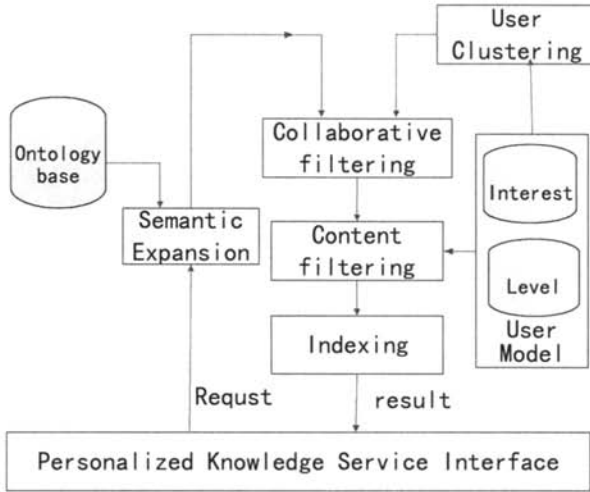


Fig. 2. Personalized Knowledge Service Framework

To address the above issues, we make use of the personalized service technology, which provide adaptive services through the collection and analysis of information of the user's interest, as the fig2 presents. At present, the realization of personalized service is mainly through the rule-based and information-based filtering technical means. The former utilizes predefined rules to filter information; the advantage is simple, direct and shortcoming is that it is difficult to ensure the quality of the rules, and the rules are of poor scalability. The latter can be further divided into content-based filtering and collaborative filtering. Content filtering is the similarity of the use of resources and users interested in, the advantage is simple, effective, the drawback is not for users interested in and users only have found similar resources with interest. Collaborative filtering is the use of user's similarity between the information to recommend information, it can discover new elements of interest for users, and the drawback is the need for user participation.

5 Relative Work

The benefits of ontologies have already been recognized in the learning technology community [6] [7]. First, the Semantic Web technologies are applied to some important learning technology standards. For example, there is an RDF Binding [8]

for the IEEE Learning Object Metadata (LOM) that is a part of the regular IEEE LTSC standardization efforts. Next, in order to better define different aspects of applications in e-learning, researchers have given several classifications of ontologies [9].

In semantic web enabled e-learning domain, several approaches are currently investigated, ranging from federated or distributed learning repositories (ARIADNE [10] or EDUTELLA [11]) or learning management systems, which focus on course delivery and administrative aspects, and adaptive web-based educational systems which offer personalized access and presentation facilities to learning resources for specific application domains.

6 Conclusion

In this paper, we first have an overview of the architecture of knowledge processing and service system for digital education. Then discuss various components of the system and the key technology involved, which is semantic representation of digital education resources, semantic mining and discovery of digital education resources and personalized knowledge delivery service separately.

We plan further implementation of the knowledge processing and service system in order to construct a digital education oriented knowledge service system and utilized the semantic web technology to realized ontology driven knowledge retrieval and delivery. The final goal is to develop a user oriented system to meet the personalized user need, which can be used in education applications for querying digital education resources, and thus improving the effectiveness of education systems in general.

References

1. LUISA (Learning Content Management System Using Innovative Semantic Web Services Architecture), <http://www.luisa-project.eu/>.
2. V. Devedzic, "Education and The Semantic Web", *International Journal of Artificial Intelligence in Education (IJAIED)*, Vol.14, 2004, pp. 39-65.
3. T.R.Gruber, A translation approach to portable ontologies, Knowledge Acquisition, 1993. [Online].
4. S.Niwattanakul, M.Eboueya, D.Lillis, "Describing and Researching of Learning Resources with Ontology Model," *jva*, pp. 214-222, IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing (JVA'06), 2006.
5. Duval, E., Ed. (2002). 1484.12.1 IEEE Standard for Learning Object Metadata. June 2002
6. Sampson, D.G., Lytras, M.D., Wagner, G., & Diaz, P. (2004). "Guest Editorial:Ontologies and the Semantic Web for E-learning", *Educational Technology & Society*, 7 (4), 26-28.
7. Lytras, M., Tsilira, A., Themistocleous, M. (2003). Towards the semantic e-learning:An ontological oriented discussion of the new research agenda in e-learning. In Proceedings of the 9th Americas Conference on Information Systems, Tampa, USA, pp. 2985-2997.
8. Nilsson, M., Palmer, M., Brase, J. (2003). The LOM RDF binding - principles and implementation, In Proceedings of the 3rd Annual ARIADNE Conference, Leuven, Belgium

9. Gasevic, Dragan; Hatala, Marek, "Ontology Mappings to Improve Learning Resource Search", British Journal of *Educational Technology*, v37 n3 p375-389 May 2006.
10. Ariadne: Alliance of remote instructional authoring and distributions networks for europe, 2001.<http://ariadne.unil.ch/>.
11. Edutella, 2001. <http://edutella.jxta.org/>.