

The Semantic Retrieval System for Learning Resources Based on Subject Knowledge Ontology

Bei Wu

College of mechanical and Electronic Information, Wuhan University of Engineering Science, Hubei, 430200, Wuhan, P. R. China

Abstract—There are many kinds of learning resources in the online education platform, but there are still many problems, such as learning resources are difficult to retrieve, search results are redundant, and can only be retrieved by keywords, but lack of semantic support. This paper integrated semantic web into the retrieval of learning resources, and designed the subject knowledge ontology as the knowledge description of learning resources semantic retrieval system, and implemented the semantic retrieval of learning resources through ontology concept similarity analysis, ontology reasoning and ontology in computing and other sectors, to reduce the redundancy of information retrieval results and realize the display of high related resources.

Keywords—knowledge retrieval; knowledge navigation; learning resources

I. INTRODUCTION

Digital learning resources have massive, distributed, heterogeneous, multi-level features, and there are different standards, high cost, personalized, low level intelligent services, resulting in a large number of digital learning resources can not be shared. One hand, digital educational resources can be shared and poor, on the other hand, digital educational resources is redundant.

Serious waste retrieval of educational resources caused great distress to the learner. How to integrate these various types of the public the mass education resources for learners to provide a unified retrieval services, and the field of digital learning problems to be solved.

Existing solutions such as search engine technology, and other applications in the field of network resource retrieval, the far-reaching social implications. Information retrieval technology is currently mainly solve the massive resources of the indexing and retrieval can not solve digitized learning the knowledge organization and meta-data semantic description such as, therefore general-purpose search engine technology in the field of digital education resources integrated retrieval also can not completely meet the digitized learning in a areas learners personalized needs.

There are two reasons: the description of learning resources is different and lack of a unified educational resources metadata standard description system, no fixed metadata mapping mechanism; Second, the distribution of educational resources is too wide, the various disciplines are not formed unified body of discipline, lack of effective methods and means of the educational resources of knowledge organization. Search engine using traditional principle, there is no strict norms of scientific terminology entry, and use key words to search for related information,

and show the advantages of rapid positioning precision appeared to be inadequate.

In response to these problems, the paper proposed the digital learning resources knowledge organization model which is applied to retrieve digital learning resources. Based knowledge organization model of digital educational resources, it can effectively solve the semantic gap problem of digital educational resources metadata, semantic mapping mechanism is formed between the metadata descriptions, and to lay the basis of resource integration for integrated retrieval of digital educational resources. Therefore, this paper will apply knowledge organization model of educational resources to the organization, navigation and retrieval of digital learning resources, to build knowledge-based organizational model of digital learning resources navigation prototype system. The paper uses knowledge organization model of learning resources to knowledge organization for Educational Technology website navigation on the classification of the network of educational resources.

II. THE APPROACH OF THE CONSTRUCTION FOR KNOWLEDGE RETRIEVED AND NAVIGATION SYSTEM

An easy way to comply with the journal paper formatting requirements is to use this document as a template and simply type your text into it.

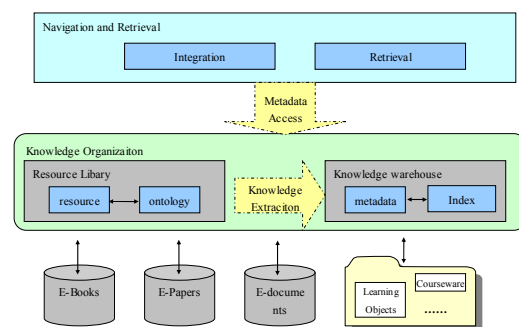


FIGURE I. THE APPROACH OF THE NAVIGATION AND INTEGRATED RETRIEVAL FOR LEARNING RESOURCES

Ontology-based knowledge organization model is applied to the integrated retrieval. The approach of the navigation and integrated retrieval for learning resources is shown in Figure I.

The navigation and integrated retrieval for learning resources includes the following modules:

The resource layer: the main data file of digital learning resource library includes the following resources: teaching

cases, teaching materials, and digital learning courseware; e-paper and e-books.

The layer of knowledge organization: knowledge organization model of digital learning resources with the proposed organization and management of a centralized digital education resources, the formation of a resource library, disciplinary ontology library, metadata extraction and automatic classification on the basis of formation by learning resources metadata warehouse of knowledge and education resource classification index.

The layer of Integrated retrieval: mainly includes integrated metadata access and integration and sharing of digital educational resources, when the user access to the knowledge warehouse integrated search portal through digital educational resources through metadata access interface knowledge warehouse metadata unified access.

III. THE KNOWLEDGE RETRIEVED AND NAVIGATION SYSTEM FOR WEB BASED LEARNING RESOURCES

The function modules of knowledge retrieved and navigation System for web based learning Resources is shown in Figure II. The system embodied knowledge organization model associated learning resources both inside and outside the organization, to achieve the educational resources the pretreatment and knowledge organization and navigation, and is reflected in the formation of a database of educational resources, including indexing library, professional word library and abstract database.

The function of the system is divided into four modules: resource module, pretreatment module, database module and knowledge organization and navigation module, which is as follows:

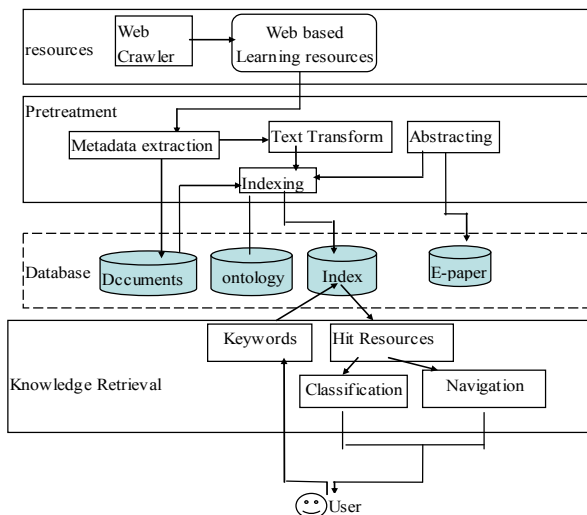


FIGURE II. KNOWLEDGE RETRIEVED AND NAVIGATION SYSTEM FOR WEB BASED LEARNING RESOURCES

A. The Module of Knowledge Retrieved and Navigation System

The resources module: This module consists mainly collected from the Internet to the various types of educational resources on the same topic, these different resource formats, such as. Html Doc, ppt, pdf, etc. It Collect the resources using a Web crawler. In this system, it mainly use Heritrix web crawler, the reptiles by secondary

development on the basis of its source code, open source web crawler using java language development, to meet the needs of the project. Filter which independent web pages little relationship with the subject, such as advertising, too many hyperlinks page.

The pretreatment module: pretreatment layer search to educational resources for processing, its main features include: metadata extraction and mapping, document conversion, automatic abstracting and indexing. Metadata extraction and mapping for conversion extraction of metadata and metadata standards of educational resources, which identify classes extracted identifier in HTML description metadata extraction, the semantic content of the class meta-data extraction Chapter keyword extraction method. The main function is to convert the document to varying resource file format converted to txt format, in order to facilitate the extraction of resources, the resources stored in the document library after document conversion. The automatic summarization module is mainly used resources generated abstracts which can help users quickly understand the main content of the resource, and thus determine the need for further view the entire contents of the resource. It will improve the speed of retrieval of resources.

The database module: mainly during system operation needs to access a variety of data resources and knowledge resources entities document library, the document describes ontology database indexing library, abstract database. And provide a mechanism for access to knowledge resources entity, body and other data resources, For example, the entity of resources, unified resource access interface; JDBC (Java Database Connectivity) will provide access interface for the database; OWL file Education Resource Description constituted body, will provide access to the Jena API and other standards-based interface.

Knowledge organization and navigation module: It is mainly used to retrieve the results of the second organization, and navigation resources generated automatically. First through the conversion of metadata heterogeneous metadata are mapped to the same standard, to achieve heterogeneous metadata reorganization. It is a resource secondary tissue by classifying the retrieval result, which will provide to the users to more clearly understand the results.

B. The Assessment to Web Based Learning Resources

Access to web based learning resource is through the web crawler, which is a program or script a according to certain rules, automatically crawl the World Wide Web information. Some infrequently used name there are ants, automatic indexing, a simulation program or worms. We primarily use an open source crawler software heritrix. On the basis of the heritrix secondary development interface localization, it is to adapt to the needs of the application.

Heritrix the runs as follows:

Through the command heritrix - admin = admin: admin start Web crawler.

C. Metadata Mapping

Metadata Mapping is of mutual conversion the information loss minimum between different metadata standards, and its essence is a meta-data standard elements and qualifiers converted to another metadata standard

