論文
3-2-4

Development of a Metadata Management System for Web-Based Coursewares
(웹 기반 코스웨어를 위한 메타데이타 관리 시스템의 개발)

Gyo-Sik Moon*
(문 교 식)

ABSTRACT

The amount of information we can obtain from the WWW is enormous. However, the users are often confused with search results due to irrelevant information, misinformation, insufficient information, etc. This paper proposes a metadata management system trying to facilitate obtaining right information on the Web. Elements of the metadata for Web coursewares are selected and their relationships are identified in order to organize them in a relational database. Procedures to obtain the metadata are implemented in the Microsoft ASP(Active Server Pages). The system provides user-friendly interfaces to give essential information on the Web coursewares and to handle users' enquiries correctly and efficiently.

요 약

WWW으로부터 얻을 수 있는 정보의 양은 엄청나다. 그러나 사용자는 종종 정보검색의 결과에 혼란스러움을 느끼는데 그 이유는 관계없는 정보, 잘못된 정보, 충분치 못한 정보 등에 있다. 이 논문은 올바른 정보의 검색에 편의를 주기위한 메타데이타 관리 시스템을 제안한다. 웹 코스웨어를 위한 메타데이타의 요소들을 가려내고 그 관계를 규명하여 관계 데이터베이스를 구성한다. 메타데이타를 취득하는 프로시저는 마이크로소프트의 ASP(Active Server Pages)로 구현하였다. 웹 코스웨어에 대한 긴요한 정보를 제공하고 사용자의 질의를 정확하고 효율적으로 처리하기위하여 사용자 편이성이 보장된 인터페이스를 시스템이 제공한다.

1. Introduction

Search engines mostly use keyword-matching techniques to retrieve information on the Web by reading the heading/body of Web pages. Different search engines produce different results because they employ different search algorithms and databases. Web users often experience difficulties using keyword-matching type of search engines, which can be summed up as follows;

1) Unwanted search results due to lack of ability to filter out irrelevant Web pages.

2) Excluding important documents due to the failure to assess the significance of the documents.

* 정회원 : 대구교육대학교 전산교육과 조교수

논문접수 : 2002. 1. 19.
심사완료 : 2002. 2. 9.

3) Irrelevant information due to the limited means of accepting user input.

It should be noted that users have no way to express specific enquiries about WBI

(Web-based Instruction) courses such as courseware type, author, multimedia, and so on.

General-purpose search engines have the above-mentioned shortcomings which are not going to be overcome in the near future. The paper shows that a reliable and efficient search mechanism can be devised by utilizing metadata on Web courses rather than using keyword-matching. Organizing and managing metadata on Web courses allows users to search for their educational needs. The system requires an initial build-up of metadata database on selected Web courses, and continual updates to the database are necessary to keep the database meaningful.

The elements of the metadata proposed in this paper can be considered as an extension of the Dublin Core format defined in 1995[12]. Nevertheless, the format is not sufficient in representing Web-based instructional materials because it was originally designed for general type of electronic documents, not specifically for Web coursewares[8].

2. Domains of Metadata for WBI Courses

2.1 Identifying Four Domains

The categorization of metadata for WBI courses clarifies the nature of metadata and gives the clue to designing a metadata database. The categorization presented in the paper emphasizes the importance of the Web environment in which teaching and learning activities are mainly taken place.

Metadata can be categorized into four domains in terms of the characteristics of information on WBI courses. The four domains are document information domain, Web technology information domain, presentation information domain, and instruction information domain. Each domain comprises a set of elements which characterizes the specific information domain.

1) Document Information Domain : This domain of information describes the overall picture of a courseware. Many educators concern about the authenticity, validity, and reliability of Web materials. To reduce the negative aspects of Web courses the designer/developer must specify essential information on the courses. For example, the information on authors suggests the quality of a course. Elements of the document information domain are courseware information, author information, location information, revision information, types of courseware, user review, hardware and software requirements, and user manuals [1],[4],[5].
2) Web Technology Information Domain : This domain describes how technological features are incorporated into Web courses. Web-specific techniques are necessary to develop courses for Web-based learning environment. As the Web techonology advances new instructional techniques must be devised and properly incorporated into Web courses. Elements of the domain include: accessibility to the course, ease of use, link-related characteristics, course management techniques, operational capabilities, interactivity, application program utilization, and external device utilization [4],[7],[11].
3) Presentation Information Domain : The domain focuses on the appropriateness of presentation of courseware contents. The main concern in the domain is how an instructional goal can be achieved via the Web presentation as opposed to traditional media such as books and CD-ROMs. The elements of the domain consist of human interfaces, textual presentation, multimedia presentation, types of presentation structure [2],[5].

4) Instruction Information Domain : The domain tells what instructional goal is and what strategy is used for achieving the goal. This type of information would benefit instructors for selecting suitable courses for various instructional purposes. Elements of the domain are the following; instructional goal, instructional strategy, and instructional model [3],[4],[10].

The categorization presented here includes most of the elements relevant to evaluate WBI courses allowing users to choose courses that suit their needs.

2.2 Extracting Metadata out of the Four Domains

We categorize information on WBI courses into four domains each of which consists of a set of metadata as the following;

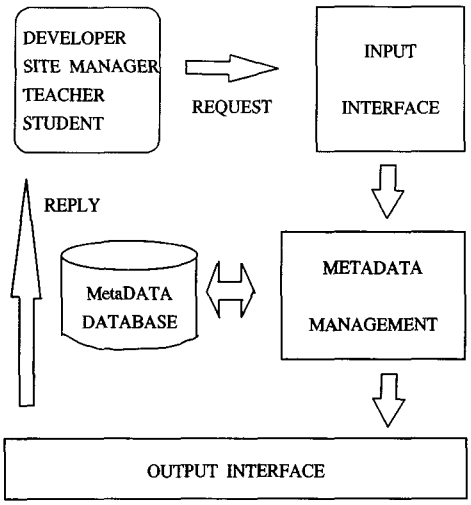
1) Document Information Domain : Courseware information (courseware title, subject, topic, grades, courseware type), author information (number of authors, author name, occupation, research area, E-mail address, phone number), help information (reference, how-to guides, manual, troubleshooting guide), location information (URL, E-mail address of the site manager, introductory statement to the course producer), revision information (revision dates, version number, revision details, revision plan), evaluation information (accuracy, authenticity, usability, effectiveness, suitability).
2) Web Technology Information Domain : Accessibility to the site, transfer time, independence of Web browers/programs, ease of use, link-related characteristics, operational capabilities, interactivity, application program utilization, external device utilization.
3) Presentation Information Domain : Human interfaces (page layout, organization, economy, communication, consistency, aesthetic integrity), textual presentation (legibility, interactiveness), multimedia presentation (types of media, content description), presentation structure (hierarchical, sequential, hypermedia).
4) Instruction Information Domain : Instructional goal, instructional strategy, instructional model.

3. Evaluation Procedure

3.1 Overview of the Interface

Metadata of WBI courses can primarily be collected during the evaluation procedure by different types of users - developers, site managers, teachers, and students. The developers initially provide intrinsic information on the courses from their point of view. The information provided by the developers encompass the entire range of the four domains, which is stored on a metadata database before the course is delivered to end-users (i.e., teachers and students).

The end-users may ask the system to find Web courses for their needs through the system's interface. And then the metadata management system responds to the search request by looking at the metadata stored by the course provider. Now, the end-users are asked to comment about the course. The user's evaluation on the course is stored on the metadata database which later can be used by other end-users. Also, the developers might access to the evaluation as a feedback so that they can improve or develop a better Web course in the future. [Fig. 1] shows the overview of the interfaces.



[Fig. 1] Metadata management system diagram

The metadata management system provides an evaluation form to be filled by evaluators so that the assessment can be done formally and efficiently. The evaluation form provides a formalized way to assess coursewares on the Web. Questions in the form may be administered differently according to different types of users. Two types of evaluation forms are designed for two types of end-users - teachers and students - because they have different perspectives in judging the quality of WBI courses.

3.2 Interface Design for Input

There are four types of users accessible to the system - developer, site manager, teacher, and student. Each user has different purposes and perspectives from which different interfaces can be devised. Users can input their responses through different interfaces according to the user type. Especially, utilizing feedback from end-users is conducive to better learning materials in the future. Interfaces for developer and site manager are not accessible to end-users so that system-related information is secured.

The developer and site manager can give useful information on WBI courses for end-users to access what they want. The developer can provide the essential information on the material such as courseware and author information. Location information and revision information are provided by the site manager whose duty is to maintain the system reliable.

Teachers are mainly interested in the effectiveness of Web coursewares. Teachers should get enough information to decide whether or not coursewares can be used in any form of instruction for their instructional needs. A preliminary review must be done before actual utilization. It would be quite beneficial to other users if teachers keep record of their experiences with coursewares.

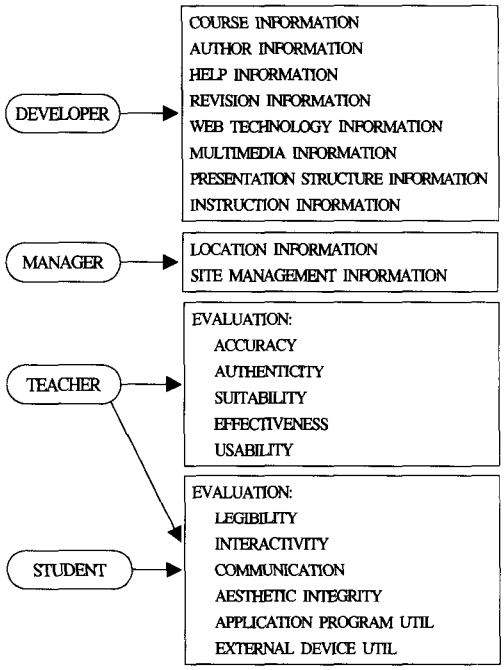
Students can search information on the Web and they can utilize the information provided by the system. Also, students can express their experiences, feelings, and so on. The responses generated by students will encourage or discourage other students to try.

The following is the list of information to be provided by developers - information on courseware, author, help, revision, browser, external device, application program, multimedia, presentation structure, and instruction. Information on location and site management is provided by the site manager.

Teachers' responses can comprise the following; ability to achieve the instructional goal, appropriateness of the instructional strategy used, accessibility to the site, ease of use, link-related characteristics (consistency, aesthetic integrity, economy, problems with navigation), interactivity, multimedia utilization, problems with external devices, application program utilization, legibility, and so on.

Students responses are often valuable opinion which courseware developers should pay heed to. There can be some discrepancies between what the developers expect students to learn from the product and what they actually learn from it because they sometimes do not look and feel the same way as developers do. Data can be gathered for various kinds of responses such as legibility, ease of use, interactivity, multimedia utilization, problems with network, and so on.

[Fig. 2] depicts relationship between types of users and metadata in which an arrow is drawn if a user is permitted to store information.



[Fig. 2] Types of users with write privilege on metadata

3.3 Interface Design for Output

The output interface is designed to provide essential information on Web coursewares to various types of users. Information can be searched by one or more of the following search variables; topic, subject, courseware title, grade, author, and description of multimedia. Search results will be the list of courseware titles that are satisfied by the search conditions provided by the user. The above variables can be used in a search enquiry individually or collectively. Logical connectors are used to form a query when there are two or more search variables are involved. Other variables can be included in a query for a better search result.

For efficiency consideration, we need to elaborate on the use of two or more variables for a search. Using topic as a search variable obviates the use of subject because the former is more precise than the latter. Keywords extracted from words appear in the topic are passed to the search engine and the result is given out as the list of all coursewares matching at least one keyword. The resulting list is sorted by the number of matches in a descending order. Courseware title can be searched similarly. We note that subject or topic is redundant when courseware title is used because title information is more precise. Using subject alone could result in very long list of coursewares. Grade can be used with any other variables.

Multimedia content description can be used to search a multimedia content matching the description inputted by the user and the result will be the list of all media contents satisfying the enquiry.

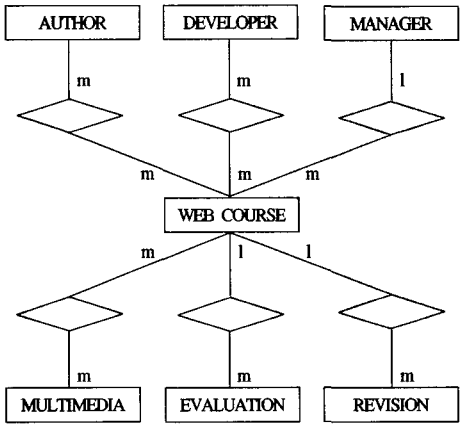
Other data can be used as search variables for a variety of reasons. Although every information stored on a database can be involved in a search, not all data are practical. If someone tries to find out Web courses of game type for a mathematical topic, then courseware type must be used as a search variable. Some users might be interested in courses written by a particular author, or developed by a particular institute.

4. Database Design and Implementation

We organize metadata extracted from the four domains of information in order to build a relational database, which is depicted briefly in [Fig. 3] as an E-R (Entity-Relationship) diagram consisting of seven entities (whose attributes are omitted for brevity).

The design was implemented in the Microsoft ASP (Active Server Pages) and the database was built on an NT server. Metadata were obtained from manual search and evaluation by a group of students and teachers trained for the purpose.

Automatic evaluation does not seem to be practical. However, standardized documents would facilitate gathering and organizing metadata. Currently, XML-based documents are emerging and gathering much attention[6],[9].



[Fig. 3] E-R diagram for Web course management database

Relations and their attributes are shown as follows;

<Table 1> Web Course Relation

Attributes	Data Type	Write Privilege
course ID	serial no.	manager
operater ID	serial no.	manager
version	number	developer
title	text	developer
subject	text	developer
topic	text	developer
grade	number	developer
type	text	developer
URL	text	developer
goal	text	developer
strategy	text	developer
model	text	developer
layout	text	developer
structure	text	developer
reference	text	developer
how-to-guide	text	developer
manual	text	

<Table 2> Author Relation

Attributes	Data Type	Write Privilege
author ID	serial no.	manager
name	text	manager
occupation	text	manager
affliation	text	manager
E-mail	text	manager
phone	text	manager

<Table 3> Developer Relation

Attributes	Data Type	Write Privilege
developer ID	serial no.	manager
name	text	manager
afflication	text	manager
E-mail	text	manager
phone	text	manager

<Table 4>Site Manager Relation

Attributes	Data Type	Write Privilege
manager ID	serial no.	manager
URL	text	manager
name	text	manager
E-mail	text	manager
site-information	text	manager

<Table 5> Revison Relation

Attributes	Data Type	Write Privilege
revision ID	serial no.	manager
course ID	serial no.	manager
date	text	developer
detail	text	developer

<Table 6> Evaluation Relation

Attributes	Data Type	Write Privilege
eval ID	serial no.	manager
course ID	serial no.	manager
version	number	manager
date	date	teacher, student
evaluator type	text	teacher, student
evaluator name	text	teacher, student
evaluator grade	number	student
accuracy	text	teacher
authenticy	text	teacher
suitability	text	teacher
effectiveness	text	teacher
usability	text	teacher
legibility	text	teacher, student
interactivity	text	teacher, student
communication	text	teacher, student
aesthetic integrity	text	teacher, student
application program	text	teacher, student
external device	text	teacher, student

<Table 7> Multimedia Relation

Attributes	Data Type	Write Privilege
multimedia ID	serial no.	manager
type	text	manager
size	text	manager
location	text	manager
author name	text	manager
description	text	manager

5. Conclusion

Searches based on metadata database outperform general-purpose search engines based on keyword matching in terms of correctness, authenticity, and efficiency. Users feel comfortable using well-organized, user-friendly interface to access educational materials for their special needs. In addition, users can participate in the improvement of the database by commenting on Web materials they experienced.

Metadata system should be maintained regularly by experts, which involves addition of new materials, deletion of obsolete ones, and modifications to existing ones.

※ References

[1] Alessi,S.M., Trollip,S.R.(1998), "Computer-Based Instruction Methods and Development", Prentice-Hall.

[2] Boyle,T.(1997), "Design for Multimedia Learning ", Prentice-Hall.

[3] Dills,C.R., Romiszowski,A.J.(editors).(1997) "Instructional Development Paradigms", Educational Technology Publications.

[4] Driscoll,M.(1998), "Web-Based Training", Jossey-Bass/Pfeiffer.

[5] Khan,B.H.(editor).(1997), "Web-Based Instruction ", Educational Technology Publications.

[6] Kim,N.W., Jeong,C.K.(2001), "Retrieval algorithm for Web Document using XML DOM", Journal of the Korea Computer Industry Education Society, vol.2, No.6, pp.775-782.

[7] Klassen,J., Vogel,D.R., Moody,E.(2000), "Design for Interactivity", ICCE 2000.

[8] Ko,M.J., Chae,K.J.(2000), "Document Management System using XML/EDI in Education Web", The Journal of Korean Association of Computer Education, Vol.3, No.1, pp.151-160.

[9] Lee,J.M., Min,K.S., Park,S.W., Kim,H.J.(2001), " XML Data support in RDBMS through XML-View", Journal of KISS: Computing Practices, Vol. 7, No.3, pp.202-210.

[10] Merrill,M.D.(1994), "Instructional Design Theory", Educational Technology Publications.

[11] Serim,F.(1996), Koch,M., "NetLearning: Why Teachers Use the Internet", O'Reilly.

[12] DC(Dublin Core), http://dublincore.org/documents/dces/

문 교 식

1982년 경북대학교 공과대학
컴퓨터공학과 (공학사)
1982 - 1986년 KIST
시스템공학연구소, 연구원
1989년 University of Oklahoma
대학원 전산학과 (이학석사)
1995년 University of North Texas
대학원 전산학과 (이학박사)
1996. 3. - 1997. 2
(부산) 동명정보대학교
컴퓨터공학과 조교수
1997. 3. - 현재
대구교육대학교 전산교육과
전임강사, 조교수.
관심분야 : 전산교육,
알고리즘, 인공지능