Scientific
Research

Sharing and Implementation of Heterogeneous Database for Education Resource Based on XML

Shixi TANG

YanCheng Teachers University College of Science & Technology, Yancheng, China.
Email: tsxlyh@163.com

ABSTRACT

The problem of sharing heterogeneous database for accessing different educational resources has to be considered. The study is carried out to realize the heterogeneous database sharing for educational resources using multi-media educational resources as the researching object. XML is applied as middleware for the practical requirements of education. The study has important practical significance for the intellectualization of educational and teaching resource platform.

Keywords: *Heterogeneous Database, XML, Education Resource*

1. Introduction

Nowadays, material resource, energy resource and information resource are three pillar resources during the development of technology and economics. Educational information resource, as an important part of information resource, plays an important role in improving the teaching quality and mining the potential of education. Therefore, many countries have set up national educational resources centers, such as the National Educational Resources Information Center of United States, the National Network for Basic Education Resource held by Basic Education Curriculum Development Center of Ministry of Education and Central Audio-Visual Education Center of China. The construction of educational resources base with corporations such as Resource of China School, K12, Clever, ZhongJiaoYuxing, Cisco Tong Fang, Tuteng, Tinghua Tangfang, becomes more and more mature. Education resource has already become an important part of network resource. It plays important role in solving the problems such as information selection, information identification, information digestion, and information individuation when users get amount of information through education resource base.

It turns out to be the most potential resource for global information transferring and sharing with the rapid development of WWW. The requirements and development of new fields, such as e-commerce, e-books and distance education have made Web data more complex and diverse. Therefore, it is difficult to store and manage all the different Web data by traditional database technology. XML is becoming the data description and exchange

standard on Internet. Meanwhile, XML and a series of related standards have been widely accepted and used, including the generation, storage, analysis of XML documents, which lay the foundation for XML as database, as well as offer the possibility to realize the sharing of data in heterogeneous databases. The automatic transforming XML into different relational databases effectively has different ways. Oracle XML SQL Utility models XML documents element as a group of nesting tables, through the element that oracle object data type modeled; IBM DB2 XML Extender saves the XML documents as BLOB type object, decomposes them to a group of tables, and defines the XML collection by the XML1.0 grammar. Microsoft solves the problem through expanding SQL-92 and involving the OPENXML line collection; Sybase Adaptive Server takes ResultSetXml Java class as the foundation of processing XML documents in two directions [1−3].

However, all manufacturers have a general character that the XML durability establishes in a special foundation, and there is no general facility to save all the XML documents. If the XML documents use a new grammar, it needs a special mapping. This is very disadvantageous for the visiting and sharing the different heterogeneous education resource. Firstly, education resource's description is very complex, each education resource's comprehensive description reaches more than 160 terms based on international standard, which has the internationalization request, and needs a standardized description frame. Secondly, the existing education resource's description is not normal, on the one hand we need to preserve its original description; on the other hand, we need a stan-

dardized description mapping. Thirdly, when users use an education resource, its description must personalize to meet the users' especial requirements by cutting out and transforming its description. Fourthly, education resource's description attribute value is dissimilar from different understanding aspects based on education resource belonging to different disciplines. Fifthly, the education resource has very wide manifestation, including text, image, sound, video, animation and so on. Its respective description must reflect both the general character and the difference. Only using existing tools given by manufacturers can not solve these problems. This article takes the XML documents as the middle data exchanging model to study the sharing problem of education resource heterogeneous database by taking the multimedia education resource as the study object, embarked from the actual requirements of education teaching.

2. Sharing Technology Selecting for Heterogeneous Database of Education Resource

Heterogeneous database systems [4] are collections of many related database systems which could achieve sharing and transparent accessing the data. Each database system with its own DBMS has already existed before adding in the heterogeneous database system. All components of the heterogeneous database have their own autonomy; each database system still remains its own application characteristics, the integrity control and security control when sharing the data. After carrying out data conversion, on the one hand, all the information to be shared is converted from source database to the destination database; on the other hand, such a conversion can not contain related redundant information. We use the interoperability technology and data integrity technology of heterogeneous database to achieve the tasks of data sharing in heterogeneous database. Heterogeneous databases interoperability is a prerequisite for data integration. The core is data accessing, so as to provide underlying technology for achieving data integration [5–6].

The methods used to achieve the exchanging of data among databases are various. Development tools with data transfer tool, such as data pipeline in the PowerBuilder, can be used, but it relies on the database structure, and its flexibility is poor. E-mail system can not meet the data exchange requirement between heterogeneous database systems, but you must connect two databases before data transmission, which in turn lower its flexibility. We solve the referred problems by using XML as a middleware of heterogeneous databases system. First of all, as a middleware, XML makes heterogeneous database system independent, and incompatibility of heterogeneous database systems is solved. If only front desk application can support XML, it can transfer

the exchange of information among heterogeneous database systems to mapping relationship between XML. Secondly, XML documents are easy to read and modify. XML documents could be opened and modified by an ordinary WordPad, and the structure is relatively simple which is easy to read, modify and convert. All these increase the flexibility of information exchanging heterogeneous database system and the scalability of XML. Thirdly, XML document format is simple, which reduces the complexity in the process of programming procedure, lowers the workload of programmers, and XML documents' code is easy as well. In view of this, we choose XML as a middleware to achieve data exchange among heterogeneous database systems [7–9].

In this research, education resource in the database includes image, text, video, music, flash and so on. Database's heterogeneities are mainly the DBMS heterogeneities including Oracle database, IBM DB2, Sybase database, Microsoft SQLserver2000 and MySQL. These already operated database systems display differently in many aspects, including data type definition, data access mode, data manifestation and so on. Therefore different database management systems cannot be connected directly to exchange the information. Oracle database, IBM DB2, Sybase database, Microsoft SQLserver2000 and MySQL have provided powerful support to XML. We transform the communication between various business databases into the data transformation between various business databases and the XML, and take image, video, text, music and flash deposited in various business databases into XML documents and import all the data into the dynamic standard database. Therefore, we adopt the following technical option. We preserve the normal primitive tree data by using various manufacturers mapping rule. The data processing frame is established. The database clusters are controlled by code distributing in the pure application procedure, business process, the database level and the application logic between the storage process, and the data processing frame guiding layer is formed by taking the education resource international standard as the foundation. The dynamic unification education resource database is established based on education resource's multiplicity and the education resource description's multiple perspectives of different discipline. User's personalized request is obtained by cutting out unification education resource database with data processing frame.

3. Education Resource Database System Heterogeneous structure

The design of the system structure is divided into three points: the presentation layer, the database layer and the logic layer, as shown in Figure 1.
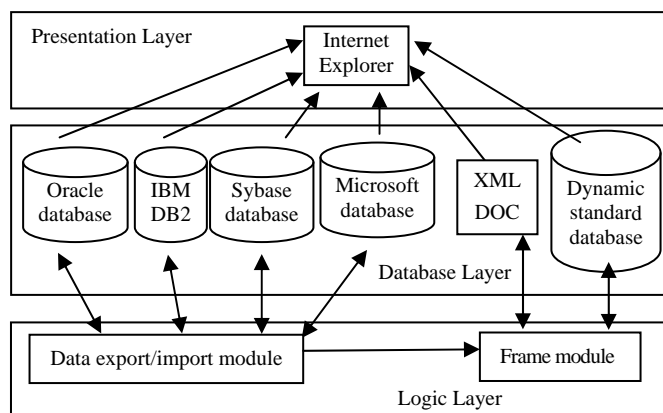
Figure 1. The design of system structure

imaID	Type	Major	Object	Source	Author	TTime	Key Word	Formats	Lenght	resolution	Sweep	Color
01001	animal	Bio-engineering	Graduate Students	Southeast University	Na Tang	2007/05	yeast, bacteria	gif	36	640*480	72	002244
01002	figures	history	Undergraduate Students	Nanjing University	Wei ZHAO	2004/06	background, achievements	jpeg	80	1024*768	300	FF0022
01003	equipment	sports	Undergraduate Students	Suzhou University	Liang-Qian Li	2002/06	function, using methods	jpeg	45	320*240	72	000044
01004	board	computer	Graduate Students	Yancheng Teachers University	Xia Wang	2007/06	CPU, graphics, memory	jpeg	68	1024*768	350	110033
01005	monitor	computer	Undergraduate Students	Wuhan University	Xiaoli Wang	2001/06	LCD, brightness	jpeg	55	1024*768	300	2244CC
01006	natural	geography	Graduate Students	Southeast University	Yao Wu	2000/06	Earthquake, plate movement	gif	38	640*480	72	320C14
01007	plant	medicine	Graduate Students	Southeast University	Hang SUN	2005/02	panda, artificial rearing	gif	74	1280*1024	300	05232B
01008	animal	computer	Undergraduate Students	Peking University	Xia Li	2005/06	lily, photosynthesis	jpeg	38	750*453	72	204A15

Figure 2. Education resource data

The presentation is the browser. Users view the results and the displaying form through browser.

The database layer is composed by Oracle database, IBM DB2, Sybase database, Microsoft SQLserver2000, MySQL, dynamic standard database and the XML documents. We attribute XML document to this layer although XML itself is not a database, since it can be seen as a complete database system with some other tools.

Logic layer consists of two modules: data export/import module and the education resource data frame module. Data export module's function is to export various business databases data and produce corresponding XML documents. Data import module's function is to analyze XML documents and import the data to various business databases. The education resource data frame modular controls all database clusters, and it is responsible to establish the dynamic unification education resource database.

4. Implement of Heterogeneous Database of Education Resource

Oracle XSU draws the XML documents to DOM, and decomposes the XML documents into a group of sub-documents by using XSLT. IBM DB2 XML Extender establishes the mapping through DAD between the database table and the XML documents' structure, storing by DB2 CLOB data type. Microsoft SQLserver OPENXML uses sp_xml_preparedocument as storing process, gaining a XML documents handle through translating XML documents into the internal DOM expression. Sybase uses XML documents type ResultSet to describe a XML documents metadata and the actual line data. The selected data is shown in Figure 2.

Document Object Model is a set of standards set by W3C, which provides an interface parsing the document. Various program languages achieve these interfaces in accordance with the DOM standards, and the parser is

```xml
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
-<image>
   -<imageinformation>
      <imaID.> 01001</imaID>
      <Type> animal</Type>
      <Major> Bio-engineering</Major>
      <Object> Graduate Students</Object>
      <Source> Southeast University</Source>
      <Author> Na Tang</Author>
      <TTime>2007/05</TTime>
      <Key Word>yeast,bacteria</Key Word>
      <Formats> gif</Formats>
      <Lenght> 36</Lenght>
      <Resolution>640*480</Resolution>
      <Sweep> 72</Sweep>
      <Color> 002244</Color>
      <Address>..\\DataSharing\src\image\1.gif</ Address >
      </imageinformation>
-<image>
   -<imageinformation>
      <imaID.> 01002</imaID>
      <Type> figures</Type>
      <Major> history</Major>
      <Object> Undergraduate Students</Object>
      <Source> Nanjing    University</Source>
      <Author>Wei Zhao</Author>
      <TTime> 2004/06</TTime>
      <Key Word> background, achievements</Key Word>
      <Formats> jpeg</Formats>
      <Lenght> 80</Lenght>
      <Resolution> 1024*768</Resolution>
      <Sweep> 300</Sweep>
      <Color> FF0022</Color>
      <Address> ..\\DataSharing\src\image\2.gpeg</ Address >
      </imageinformation>
```

Figure 3. Education resource spanning tree

given to parse the documents. The parser establishes a tree in memory through reading XML documents. The tags of XML document, tagged text content and entities correspond to a certain node of the tree in the memory. It's easy to deal with XML documents, to read, traverse, modify, add and delete the documents through operating the node of tree in memory. We use DOM parser to do XML programming for an application, which can easily handle XML documents by operating the node of tree in memory to obtain the data needed.

DocumentBuilderFactory class is responsible for creating the instance. DocumentBuilderFactory class calls its newInstance () method to instantiate a Document-BuilderFactory object. Factory object calls newDocumentBuilder () method to return a DocumentBuilder object. And finally the builder object calls newDocument () method to achieve instantiating Document interface.

```
DocumentBuilderFactory factory =
DocumentBuilderFactory.newInstance();
DocumentBuilder builder =
factory.newDocumentBuilder();
Document doc = builder.newDocument();
doc.setXmlVersion("1.0");
Element root = doc.createElement("image");
doc.appendChild(root);
```

The CreateXML.xml documents are produced as follows:

```
File file = new
File("F:/temp/MatPrj/WebRoot/CreatXml.xml");
if( !file.exists()||!file.isFile()){
   new FileOutput-
Stream("F:/temp/MatPrj/WebRoot/CreatXml.xml");
   file = new
   File("F:/temp/MatPrj/WebRoot/CreatXml.xml");
   }

StreamResult streamResult = new StreamResult(file);
Source inputSource = new DOMSource(doc);
   TransformerFactory transformerFactory =
   TransformerFactory.newInstance();
   Transformer transformer =
   transformerFactory.newTransformer();         transformer.transform(inputSource, streamResult);
```

The XML documents are produced, and the data is imported to XML documents, its education resource spanning tree is shown in Figure 3.

The data processing frame is established to control the database clusters, and the dynamic uniform education

WU.dbo.image table

imaID	Type	Major	Object	Source	Author	TTime	Key Word	Formats	Lenght	resolution	Sweep	Color
01001	animal	Bio-engineering	Graduate Students	Southeast University	Na Tang	2007/05	yeast, bacteria	gif	36	640*480	72	002244
01002	figures	history	Undergraduate Students	Nanjing University	Wei ZHAO	2004/06	background, achievements	jpeg	80	1024*768	300	FF0022
01003	equipment	sports	Undergraduate Students	Suzhou University	Liang-Qian Li	2002/06	function, using methods	jpeg	45	320*240	72	000044

WU.dbo.text table

TxtID	TxtType	TxtMajor	TxtObject	TxtSource	TxtAuthor	Txt Time	TxtKey Word
02001	Economics Paper	International economic	Graduate Students	Economics	Li Zhang	2007/05	Bubble economy, Economic globalization
02002	Economics Paper	China's economy	Undergraduate Stu-dents	Economics	Ke Ban	2004/06	China imported inflation, Real estate and living
02003	Management Paper	Business Management	Undergraduate Stu-dents	Management	Xiao-ming WU	2002/06	Enterprise Project Management

WU.dbo.vedio table

VedID	VedType	VedMajor	VedObject	VedSource	VedAuthor	VedTime	VedKey Word
05001	Computer	An Introduction to Computer	Graduate Students	Southern Yangtze University	Kequn Wang	2006/05	computer, chip
05002	Computer	Computer Application	Graduate Students	Suzhou University	Liang Ke	2004/06	computer, NET,J2EE

WU.dbo.flash table

FlaID	FlaType	FlaMajor	FlaObject	FlaSource	FlaAuthor	FlaTime	FlaKey Word
03001	Computer	Software Engineering	Graduate Students	Shenyang Polytechnic University	Xing Liu	2005/05	computer, C++
03002	Education	Chinese Language & Literature	Underraduate Students	Central China Normal University	Chuan Zhao	2004/06	literature, education

WU.dbo.music table

MusID	MusType	MusMajor	MusObject	MusSource	MusAuthor	MusTime	MusKey Word
04001	Education	Chinese Language and Literature	Graduate Students	Southeast University	Keyi Zhang	2007/05	literature, education
04002	Education	Chinese Language and Literature	Undergraduate Students	Nankai University	Guosheng Huang	2003/06	literature, education

Figure 4. Dynamic uniform education resource database

resource database is also established as shown in Figure 4 according to the multiplicity of education resource's manifestation and the multiple perspective of the education resource description of different discipline based on international standard of education resource.

The personalization description of education resource is obtained by cutting the dynamic unification education resource database using the data processing frame according to the user's personalized request, as shown in Figure 5.

```
Imageinformation 0
   imaID          01001
   Type           animal
   Major          Bio-engineering
   Object          Graduate Students
   Source         Southeast University
   Author         Na Tang
   Ttime          2007/05
   Key Word       yeast, bacteria
   Formats        gif
   Length         36
   Resolution     640*480
   Sweep          72
   Color          002244
   Address        ..\\DataSharing\src\image\1.gif
```

Figure 5. The personalization education resource result in client after cutting out

5. Conclusions

The paper defines a XML document which describes the database structure. We fill the information of the education resource database into self-explanatory XML documents in order to let users create database friendly. And a scheme is made as designing a middleware between the database and its outside. A data processing frame is advanced to process each kind of actual problems which are brought by the complexity, the dynamic and the personalization of the education resource data. The exchanging requests of internal or external education resource data are submitted to the middleware with XML. The interaction of specific education resource database is realized by the middleware and the results are fed back to the requester through XML. All the details in the process have been shielded to achieve transparent sharing access of heterogeneous education resource database.

REFERENCES

[1]   Oracle XML-SQL Utility, http: // www. oracle.com /technology/ index.html. Last accessed on July 27, 2009.

[2]   IBM DB2 XML Extender, www.ibm.com/ software/ data/ db2/ extenders/xmlext. Last accessed on July 27, 2009.

[3]   XML Perspective, In control with FOR XML Explicit. SQL Server Magazine, http://msdn.mcrosoft. com/ library/ periodic/. Last accessed on July 27, 2009.

[4]   A. P. Sheth and J. A. Larson, "Federated database systems or managing distributed, heterogeneous, and autonomous databases," ACM Computing Survey, Vol. 22, No. 3, pp. 183–236, 1990.

[5]   Q. Feng, H. Q. Lv and H. Feng, "The connection of Heterogeneous databases," Computer and Information Technology, Vol. 9, 2001.

[6]   X. Wang and S. M. Wei, "Java technology-based distributed heterogeneous database Web access technology," Computer Engineering and Applications, pp. 135–138, 2000.

[7]   D. Martin, "XML High-level Programming," Beijing Machinery Industry Press, pp. 77–103, 2001.

[8]   D. Motton, "XML programming technology [M]," Beijing Machinery Industry Press, pp. 134–200, 2001.

[9]   X. H. Dong [EB/OL], http://www.XML.org.cn, Applications to build XML.