

# DEVELOPMENT OF AN XML SCHEMA REPOSITORY SYSTEM FOR PUBLIC CONSTRUCTION INFORMATION SHARING

Choung-Houng Wu, Shang-Hsien Hsieh, I-Chen Wu, and Hui-Ping Tserng  
Department of Civil Engineering, National Taiwan University, Taipei, Taiwan  
boyce@caece.net, shhsieh@ce.ntu.edu.tw, kwu@caece.net, hptserng@ce.ntu.edu.tw

Tzung-Yu Wen  
Application Integration Division, Acer Incorporated, Taipei, Taiwan  
robin\_wen@acer.com.tw

**Abstract:** The Public Construction Commissioner (PCC) in Taiwan is sponsoring a project to establish the mechanism for developing information standards, in terms of XML schemas, for public construction work and develop an XML Schema repository system for managing and publishing the standardized XML schemas. Parts of the ebXML (Electronic Business using eXtensible Markup Language) Registry Information Model (eBRIM), Version 2.5, proposed by OASIS (Organization for the Advancement of Structured Information Standards), are adopted in the design of the repository system. This paper reports the progress on the development of the repository system in the project.

**Keywords:** XML schema, Repository, Information standardization and sharing, ebXML, Public construction

## 1. INTRODUCTION

In recent years, Taiwan government has made many efforts in providing government e-services, including government-to-government (G2G), government-to-business (G2B), and government-to-customer (G2C) services to its client users via Internet. For example, in the area of public construction, the government procurement information is now available through the Internet and on-line tendering services will be provided soon. Furthermore, a project has recently been sponsored by the Public Construction Commissioner (PCC) to promote exchange and sharing of life-cycle construction information in Taiwan. There are two major tasks to be accomplished in the PCC project. The first one is to establish the mechanism for developing information standards, in terms of XML Schemas, for public construction work. The second task is to design and implement an XML schema repository system for managing the standardized XML schemas and providing schema search and retrieval services on World Wide Web (WWW). This paper reports the progress on the development the XML Schema repository system in the project.

In the rest of this paper, requirement analysis of the repository system is discussed first in Section 2. Then, the design of the system is presented in Section 3. The Unified Modeling Language (UML) [1-2] is employed in this work for object-oriented analysis and design of the repository system. Section 4 discusses the implementation of the system and provides some demonstrations on the web interfaces

and services of the system. Finally, some conclusions are drawn.

## 2. SYSTEM REQUIREMENT ANALYSIS

The main purpose of the repository system is to store, manage, and publish the standard XML schemas and their corresponding documentation for information sharing and exchange in public construction work. The repository system also stores and manages a data dictionary containing definitions and synonyms of all of the vocabularies used to define the tag names in the standard XML schemas. Each of the tag names is unique and named in English. Its corresponding Chinese tag name and related attributes are also defined in the data dictionary. Furthermore, an XML schema for the data dictionary is developed and standardized in this work.

For development of standard XML schemas and the corresponding data dictionary, a standards establishment committee has been organized by PCC and a standards establishment mechanism has been established. Once the new version of a standard XML schema is approved by the committee, it is registered into the repository system for publication on WWW. The contents (mainly the tag names and their attributes) of a newly registered schema are validated by the system to ensure that they have been uniquely defined in the data dictionary. The system should also provide query and download services for its client users to search and retrieve the standard XML schemas and related information in the data dictionary needed for their work. Furthermore, the

client users may need validation services for checking the validity of their XML documents to a specific standard XML schema. In addition, system administration service is needed for maintenance and monitoring of the system.

Based on the above requirement analysis, the following major services should be supported by the repository system:

- System administration service: This service provides the necessary functions for managing user accounts, maintaining schema data and data dictionary, monitoring system performance, etc.
- Registry service: This service provides the necessary functions for the standards establishment committee to register and publish the approved versions of standard XML schemas as well as the updated data dictionary.
- Query management service: This service provides the necessary functions for searching and retrieving a specific standard XML schema and its associated information in the data dictionary.
- Content validation service: This service provides the necessary functions for validating the new or updated XML schema when it is registered into the system and for checking the validity of a given XML document to its corresponding standard XML schema.

Figure 1 shows the major services and applications of the XML schema repository system. All of the services, except of the system administration service, are mainly provided in the form of Web Services [3-4]. A public construction information portal, which is currently being constructed, will provide web pages for the users to have easy access to the query management and content validation services of the XML schema repository system.

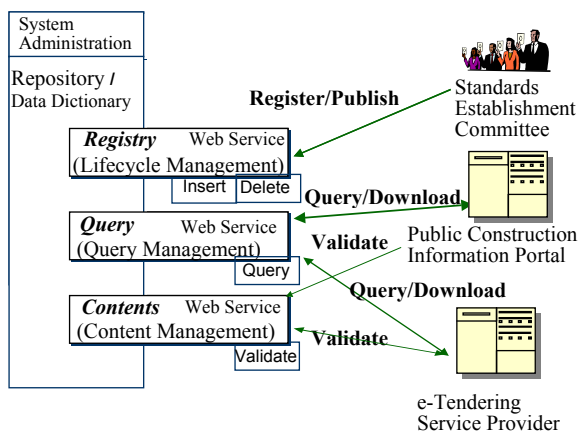


Figure 1. Major services and applications of the XML schema repository system

### 3. SYSTEM DESIGN

Parts of ebXML (Electronic Business using eXtensible Markup Language) Registry Services Specification (ebRS), version 2.5 [5], and ebXML Registry Information Model (ebRIM), version 2.5 [6], proposed by OASIS (Organization for the Advancement of Structured Information Standards), are adopted in the design of the repository system. For designing the services of the system, the following parts in ebRS v.2.5 [5] are adopted in this work:

- Lifecycle management service: It provides a collection of lifecycle management methods for managing objects within the ebXML registry. The client user may submit objects, classify and associate objects, and deprecate and remove objects.
- Query management service: It allows a client user to discover and retrieve information from the ebXML registry. The client user may browse through and issue queries on registry contents.
- Content management service: It uses the content validation service to enforce validation rules upon submitted contents and metadata in a content specific manner.

For designing the classes in the information model of the XML schema repository system, parts of the ebRIM v2.5 [6] are adopted in this work. Figure 2 shows the high level public view of the ebRIM information model. Although the version (i.e., v2.5) adopted here is just a committee approved specification and is not a final release version (which will be v3.0), it seems to be very close to the upcoming final release version. In this work, only the following classes in the ebRIM v2.5 information model are adopted: the RegisterEntry class, ExtrinsicObject class, User class, Organization class, PostalAddress class, TelephoneNumber class, Association class, ClassificationNode class and Classification class.

Figure 3 shows the class diagram for the object-oriented design of the data dictionary. The corresponding XML schema design for the data dictionary has been approved by the standards establishment committee.

Because the relational database is still the most popular and widely used database, it is employed in this work to store and manage the information objects associated with the XML schemas and the data dictionary. The entity-relation (E-R) model diagrams for the schema repository database and the data dictionary database are designed as shown in Figs. 4 and 5.

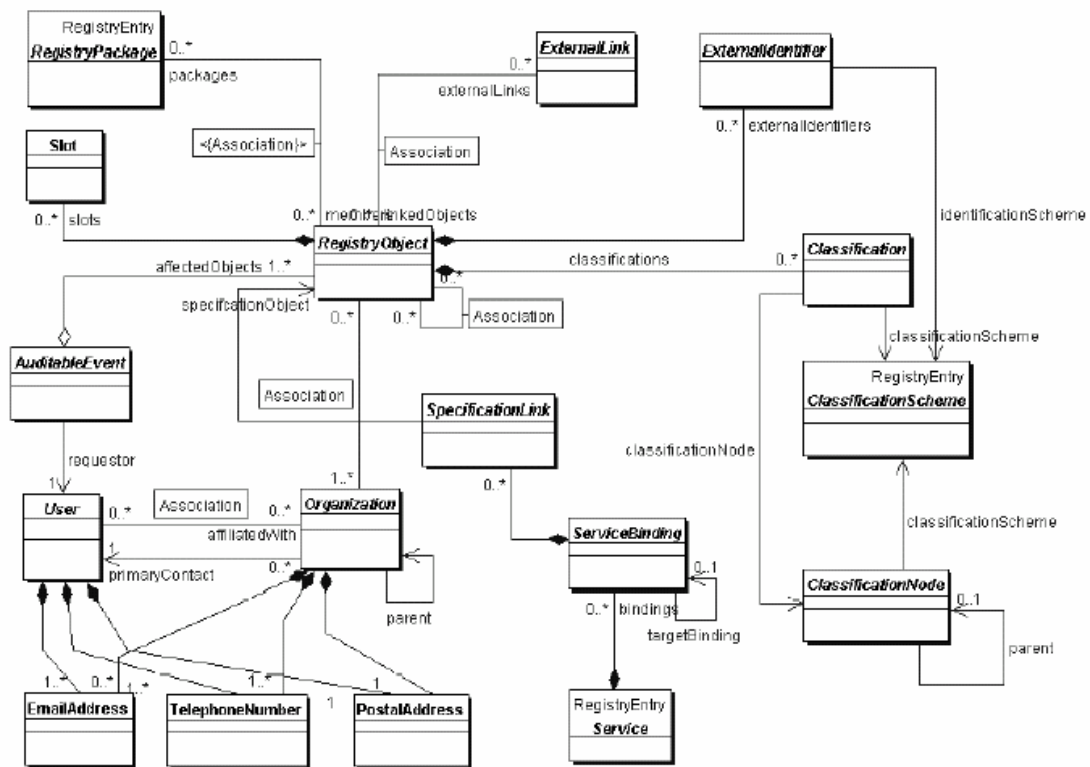


Figure 2. The high level public view of the ebRIM Information model [6]

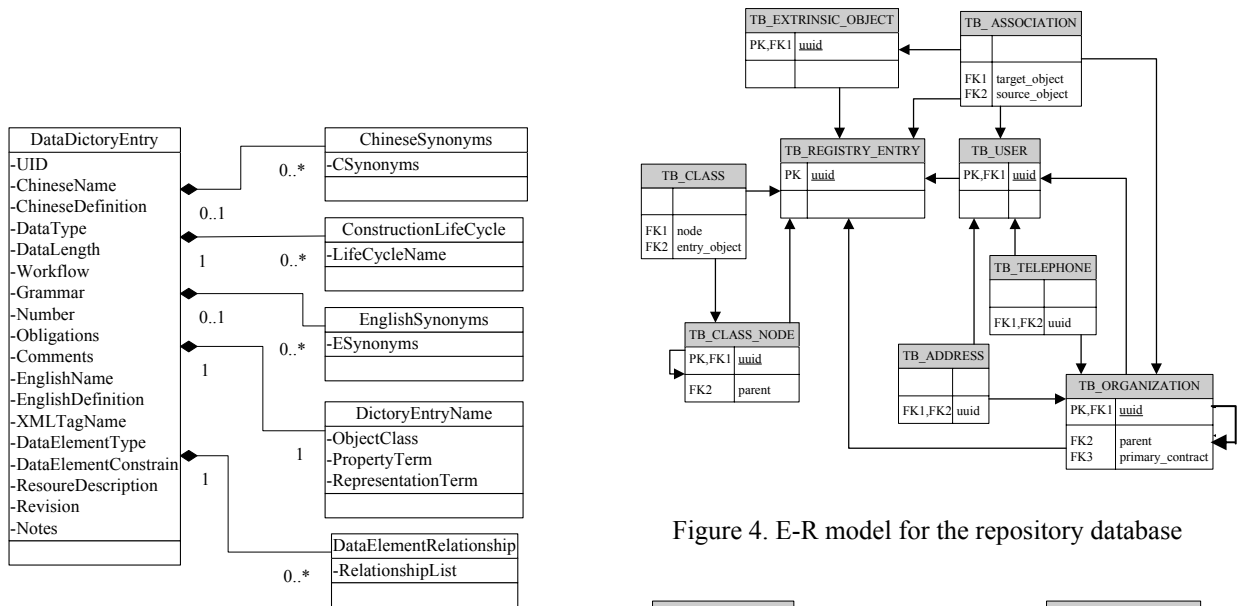


Figure 4. E-R model for the repository database

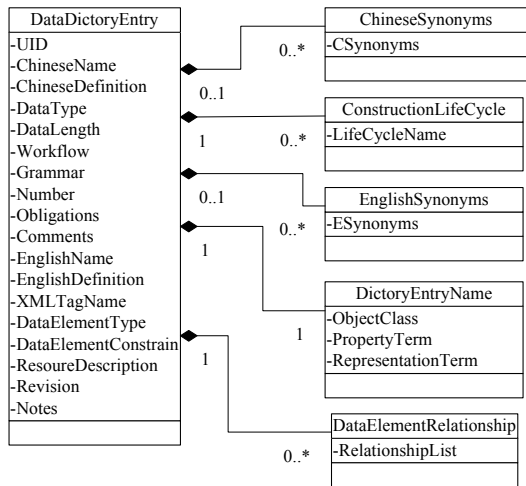


Figure 3. Class diagram for the data dictionary

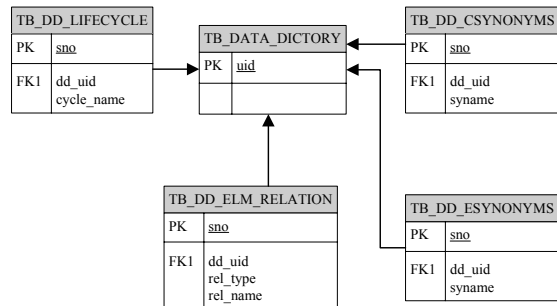


Figure 5. E-R model for the data dictionary database

## 4. SYSTEM IMPLEMENTATION AND DEMONSTRATION

The implementation of the XML schema repository system requires two servers: database server and web server. This work employs Informix v9.4 as the database server and Resin v2.1.10+SLL as the web server. The operating system used for running both servers is the RedHat v8.0 Linux. The programming language used is the Java programming language. In addition, the Java WebService Module (Axis v1.1), the Java Upload Module, and FullText-Search for Informix are used to facilitate the implementation of web services, uploading service, and searching service for the repository system.

Most of the basic service functions for the repository system have already been implemented although more work is still needed to complete the system implementation as designed in Section 3. Figure 6 shows the home page of the repository system. At present, all of the system interfaces are in Chinese. The web page shown in Fig. 7 illustrates the results of a conditional query on the standard XML schemas in the repository. Figure 8 demonstrates the query result, expressed in tree format, on a specific XML schema. The query result can also be expressed in the table format. Figure 10 shows the query result of the data dictionary information.



Figure 6. The home page of the repository system

## 5. CONCLUSIONS

The progress of an on-going project on development of an XML schema repository system has been reported in this paper. The system is designed to

promote and facilitate information sharing and exchange for public construction work in Taiwan. The standard XML schemas and the data dictionary managed by the system are two essential components to achieve the goal. They are also important assets of the construction industry because they represent the results of accumulated knowledge and experiences in the public construction work. It is hoped that the repository system can provide good services for maintaining and querying standard XML schemas and help to promote the use of information standards in Taiwan's construction industry.

## ACKNOWLEDGEMENTS

The work presented in this paper is sponsored by the Public Construction Commissioner (PCC), Executive Yuan, Taiwan, Republic of China. In addition, the authors would like to thank the Application Integration Division of Acer Incorporated for their efforts on implementing the XML schema repository system.

## REFERENCES

- [1] G. Booch, I. Jacobson, and J. Rumbaugh, *The Unified Modeling Language User Guide*, Addison-Wesley, 2001.
- [2] G. Booch, I. Jacobson, and J. Rumbaugh, *A Brief Guide to the Standard Object Modeling Language (UML Distilled Second Edition)*, Addison-Wesley, 2000.
- [3] Newcomer, E., *Understanding Web Services – XML, WSDL, SOAP, and UDDI*, Addison-Wesley, 2002.
- [4] Cerami, E., *Web Services Essentials*, O'Reilly & Associates, 2002.
- [5] OASIS/ebXML Registry Technical Committee, "OASIS/ebXML Registry Services Specification v2.5 – Committee Approved Specification," <http://www.oasis-open.org/committees/regrep/documents/2.5/specs/ebrs-2.5.pdf>, June 2003.
- [6] OASIS/ebXML Registry Technical Committee, "OASIS/ebXML Registry Information Model v2.5 – Committee Approved Specification," <http://www.oasis-open.org/committees/regrep/documents/2.5/specs/ebrim-2.5.pdf>, June 2003.



Figure 7. Results of a conditional query

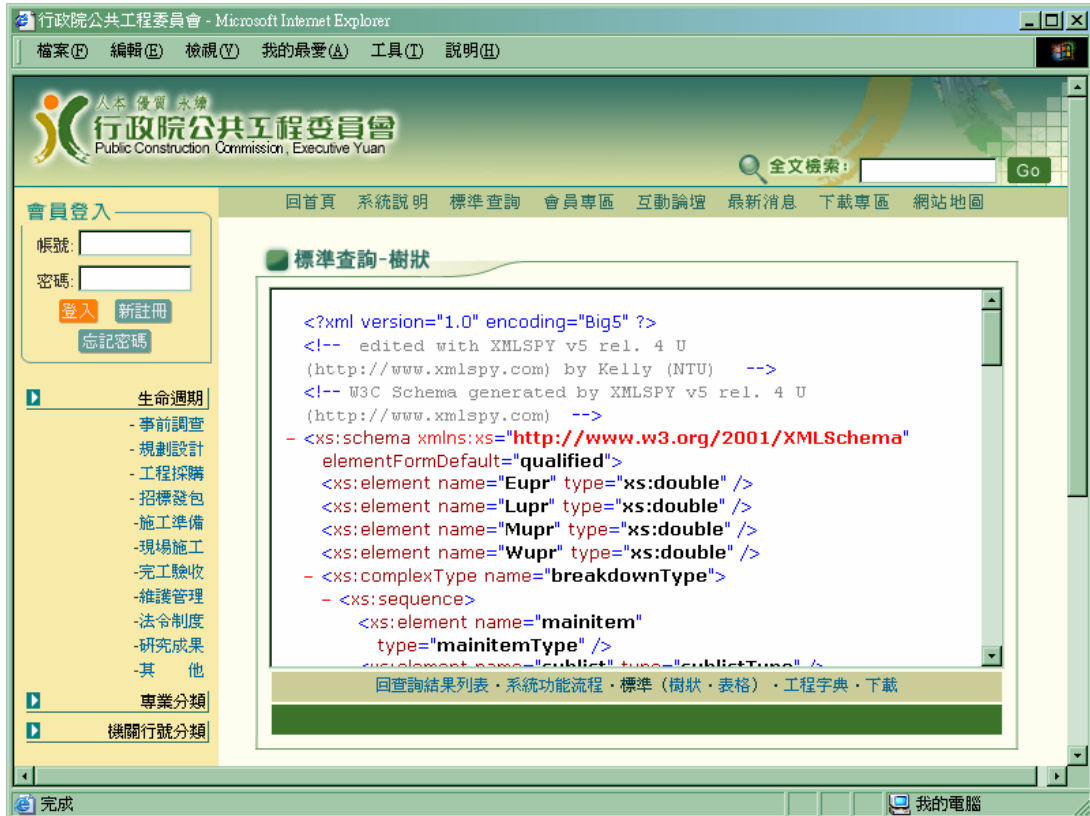


Figure 8. Query result (in tree format) on a specific XML schema

行政院公共工程委員會 - Microsoft Internet Explorer

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)

行政院公共工程委員會  
Public Construction Commission, Executive Yuan

全文檢索:  Go

回首頁 系統說明 標準查詢 會員專區 互動論壇 最新消息 下載專區 網站地圖

**會員登入**

帳號:   
密碼:   
登入 新註冊  
忘記密碼

▶ 生命週期  
- 事前調查  
- 規劃設計  
- 工程採購  
- 招標發包  
- 施工準備  
- 現場施工  
- 完工驗收  
- 維護管理  
- 法令制度  
- 研究成果  
- 其他

▶ 專業分類

▶ 機關行號分類

**標準查詢-工程字典**

屬性名稱	說明
UID	XXXXXXXXXX
中文名稱	工項編碼
中文名稱定義	按公共工程施工綱要規範所編列之工作項目編碼
中文同義詞	項目編碼 工程項目編碼 作業編碼 契約項目編碼
資料型態	string
資料大小	20
營建生命週期階段	跨多個生命週期
工作流	
詞彙型態	欄位名詞
出現次數	1..1
資料項權重	Mandatory Element
註解	
資料元素關係	
英文名稱	Work item No.

完成 我的電腦

Figure 9. Data dictionary information query