The Design of Metadata for the Digital Museum Initiative in Taiwan

Chao-chen Chen
Department of Adult and Continuing Education
National Taiwan Normal University
cc4073@tpts1.seed.net.tw

Hsueh-hua Chen
Department of Library and Information Science
National Taiwan University
sherry@ccms.ntu.edu.tw

Kuang-hua Chen
Department of Library and Information Science,
National Taiwan University
khchen@ccms.ntu.edu.tw

Jieh Hsiang
Department of Computer Science and Information Engineering
National Taiwan University
hsiang@csie.ntu.edu.tw

Abstract

Metadata plays a crucial role in a digital library/museum environment. However, the development of metadata is not an easy task. Its formulation starts with analyzing the attributes of collections as well as understanding the user information needs and information seeking behavior. The issue of interoperability also needs to be considered in terms of both semantics and syntax. This paper discusses issues related to the development of metadata in Taiwan. We describe the development process of a Chinese metadata system, Metadata Interchange for Chinese Information (MICI), and an XML/metadata management system, Metalogy. Both were developed under the Digital Museum Initiative sponsored by the National Science Council of Taiwan.

1. Introduction

With the rapid growth of the Internet, research on digital libraries and digital museums has received worldwide attention. Taiwan, in tune with the rest of the developed countries, has also been enthusiastically embracing digitisation effort.

Among the major institutions in Taiwan that are currently digitising their collections of cultural heritage are National Palace Museum, National Museum of History, National Museum of Natural Science, National Central Library, National Taiwan University, and Academic Sinica [1]. Their primary tasks are to digitise their own unique collections and present them on the web in a uniform way. In order to organize this vast amount and variety of cultural data in an easily retrievable manner, it is imperative to design a suitable metadata format. From the user point of view, a digital library/museum should at least contain basic functions of retrieval and browsing, as well as to other related web resources. Usually, a digital library/museum with a large volume of data requires a database management system to manage its metadata, digital objects, and the links, etc. In this paper we describe

an effort in developing a Chinese metadata format, the Metadata Interchange for Chinese Information (MICI), and an XML/metadata management system, Metalogy. Both were developed under the sponsorship of the Digital Museum Initiative, funded by the National Science Council of the Republic of China.

2. The Digital Museum Initiative of Taiwan

In order to alleviate the quality of research in humanities and social sciences, the National Science Council (NSC) of Taiwan launched a program termed "Greeting a New Millennium—A Cross-Century Technology Development Program with Concern for the Humanities" in 1998. The Digital Museum Initiative (DMI) was part of this program, with the mission of building digital data from cultural artefacts and, with which, developing digital museums. DMI has several goals. By digitising cultural contents, it aims at cultural preservation and provides a more modern research environment for social scientists. By building digital museums and making them available over the web, it plans to provide high quality cultural and scientific content for general public and students in Taiwan [2]. By promoting digitisation, NSC hopes to stimulate related research such as digital library and multimedia, and the growth of the content industry. Since building a digital museum requires the participation of content experts, information/library scientists, technologies, and educators, another important goal is also to encourage interdisciplinary research collaboration.

During the first year (September 1998 -- August 1999) of DMI, NSC invited experts and scholars with experience on digital collections to form a collaborative mechanism to promote digital museum research. Under this program, two types of projects were initiated: topic-based projects (digital museums on specific subjects) and technical support projects [4]. In addition, DMI also includes an educational program for training and exchanging experience with digital libraries/museum construction and generally serving as a bridge for the digital library/museum community.

The topic-based projects in the first year include subjects in both local and traditional culture. On the local Taiwanese culture side, there were "The Discovery of Tamsui River" and "Indigenous Taiwanese – The Pingpu People". On natural science and environmental ecology, there were "Taiwanese Butterflies" and "Native Plants and Fishes of Taiwan". There were only three projects on traditional Chinese culture: "Traditional Thoughts and Literature (The Four Books, Lou-Chuang, Poems of the Tang Dynasty)", "An Immortal Palace—Han Dynasty Culture and Burials", and "Firearms and Warfare of the Ming and Ching Dynasties"[5].

The technical support projects in the first phase consist of five different areas: geographic information system, word net, metadata, copyright management, and effectiveness evaluation. Among these five areas, the authors of this paper led a metadata research team, ROSS (Resources Organization and Searching Specification), to study metadata related issues and to design metadata suitable for Chinese material.

The principal investigators of the first year's projects were mainly from the Academic Sinica, the National Taiwan University, the National Tsing-Hua University, and the National Chi-Nan University.

After a successful first year, the second year's program, from January 2000 to December 2000, was announced to the general research community through a Request-for-Proposal invitation. It drew wide attention and attracted nearly 90

topic-based proposals, among which 12 were selected and funded, including four continuations from the first year [6]. They are:

1. Treasures of the National Palace Museum
2. The World of Xuanzang and the Silk Route
3. The Discovery of the Tamsui River (II)
4. Native Artist Digital Museum—Yu-Yu Yang Art Research Center
5. Historical Photographs of Taiwan
6. Architectural History of Taiwan
7. Mystery of Human Body
8. Indigenous Taiwanese — The Pingpu People (II)
9. Ancient Texts and Popular Songs of Tang and Sung Dynasties (II)
10. Native Freshwater Fishes of Taiwan (II)
11. Chinese Medicine and Acupuncture
12. Biology-Cultural Diversification of Orchid Island

During the second phase, two technical support projects, on metadata and digital watermarks, were carried out. The ROSS team continued its metadata research and developed an XML/metadata management system—Metalogy.

Since the second phase projects were opened to the entire academia, the principal investigators were from many universities and organizations in addition to the four mentioned in the first phase.

The Digital Museum Initiative is currently in its third year, with 15 topic-based projects cover a wide variety of different topics, including language literature, religious art, folk culture, historical relics, mathematics, biology, architecture, and geography.

3. Resources Organization and Searching Specification

The ROSS (Resources Organization and Searching Specification) research team on metadata studies was actually formed even before the DMI started. It came into existence in 1997, out of the need of a uniform metadata format for Chinese artefacts in the Taiwan University Digital Library/Museum (NTUDL/M) Project. Its research scope contains the following: to understand the history and features of collections at NTU, to study various metadata formats both domestically and internationally, to understand relations among metadata, database and the system framework, and to understand information needs and information seeking behaviour of potential users. ROSS holds that metadata should be able to describe attributes of the collections, to supply users with the mandatory access points, to provide interoperability among different digital libraries and museums to exchange information, and to enhance the quality of cataloguing. Most of the digital collections of NTUDL/M were historical documents. After studying the characteristics of historical documents, ROSS made in-depth studies of the metadata of similar types of collections, including CIMI (Computer Interchange of Museum Information), Dublin Core, EAD (Encoding Archival Description), TEI (Text Encoding Initiative) Headers, and so on. Due to cultural and characteristics differences, however, these metadata formats are not sufficient to describe Chinese special collections. It is therefore necessary to focus on research on the design of Chinese metadata, which becomes the main research goal of ROSS [7].

When NSC lunched the Digital Museum Initiative in 1998, ROSS team was invited to participate as one of its technical support projects. During the first year,

ROSS team was responsible for the metadata development for the topic-based projects "The Discovery of the Tamsui River" and "Taiwanese Butterflies". In the second year its role was expanded to developing a metadata management system, Metalogy, for handling various types of metadata for use by all of the topic-based projects. Currently, Metalogy is used by several organizations, including: National Palace Museum, National Taiwan University, National Chiao-Tung University, and some other libraries. The ROSS team also spent a significant amount of time assisting the National Palace Museum in developing metadata for their contents.

4. Procedures for Designing Metadata

According to our experience, the development of metadata includes at least seven steps.

4.1 Analysing the attributes of collections:

The first step in formulation of metadata is to understand and extract the common features and characteristics of collections. We spent a lot of time discussing with content experts in order to gain better understanding of the attributes of the collections.

4.2 Need assessment of metadata users:

Interview the content experts and potential users to understand their information needs and information seeking behaviours.

4.3 Interoperability Consideration:

During the development of our metadata, special attention was paid to the compatibility with international standards. Thus, we joined The CIMI Consortium (Consortium for the Interchange of Museum Information) since 1998 as a member and were involved in its Dublin Core Testbed Project. We attended the 7th and 8th meetings of Dublin Core Metadata Initiative (DCMI) to keep up with the latest development of the Dublin Core. We also investigated various formats to see which one is most suitable for expressing the complex digital library/museum content. We also adopted XML as the standard format of syntax.

4.4 Semantic design of metadata:

We designed the *Metadata Interchange for Chinese Information* (MICI) format and adopted qualified Dublin Core as its basic structure for metadata semantics. We termed this set of metadata MICI-DC.

4.5 Developing metadata management system:

We developed the *Metalogy* metadata management tool for different types of metadata, including MICI-DC. This system may be used to develop databases for any digital library/museum in different subjects. Utilities of Metalogy include database set-up by the DTD, a metadata editor, an authority file editor, retrieval functions (including both Window and Web interfaces), the importing/exporting of XML files, and others.

4.6 Developing tagging guide and user manual:

In order to make it easier for users to catalogue their collections using MICI-DC, a tagging-guide was compiled with explanations and examples of the 15 elements of DC and their qualifiers.

4.7 Providing training courses:

In order to promote Taiwan's museum community to better understand metadata, we organized a number of workshops for museum and library workers to understand the importance and processes of digitisation and metadata. After the completion of

MICI-DC and Metalogy, we also offered a series of workshops and courses to teach the inner workings of our entire methodology. Some of these courses were conducted through the education program part of MDI.

5. Metadata Interchange for Chinese Information -- MICI

MICI-DC has been used to catalogue various types of collections including historical documents, old maps, photographs, calligraphies, archaeological artifacts, and Buddhist scriptures/paintings. Users may choose DC's 15-element set and qualifiers and adjust the orders of these elements according to their needs. Local qualifiers are added to appropriate elements based on the attributes of collections. Individual institutions may also define their own qualifiers based on the attributes of their collections. This provides a better way to capture the characteristics of the diversity of cultural heritage and user needs. Our formalism is compatible with international standards and, in the meantime, gives users great flexibility in meeting local needs.

The details of MICI-DC is listed as follows:

Metadata Interchange for Chinese Information (MICI-DC)* Last Modified 2000.07.03

Element	Qualifier		
Type	Aggregation Level	Item/ Collection	
	Original / Surrogate	Original/ Surrogate	
	Cultural / Natural	Cultural / Natural	
	DC Type	interactive resource	
		Dataset	
		Event	
		Image	
		Sound	
		Service	
		software	
		collection	
		text	
	Local Level		
Format	Medium		
	Extent (size, duration)	Quantity	
		Dimension	Name
			MeasurementsUnit
			Position
Title	Main		
	Subtitle		
	Alternative		
Description	Acquisition	Method	
		Source	
		Price	
	Physical Description	Illustration	
		Color	
		Material	
		Attachments	
		Form Whole Object	
		Part Of Object	
		Scale	
	Abstract / Synopsis		
	Place		
	Collection Or Site Information	Locality	

Element	Qualifier		
		Name	
		Date Gathered	
		Field Number	
		Method of Collection	
		Type Of Site	
		Coordinates	
		Coordinates of Object	
		Phonomena	
		Accompanying Object	
		Cultural Layer	
		Geological Period	
		Age	
		Environmental Details	
	Seal Type	From Artist	Artist Inscription Seal Locality
			Artist Inscription
			Artist Seal
		About Colophon	Colophon Locality
			Colophon Writer
			Colophon Seal
			Colophon Full Text
		About Label	Label Locality
			Label Writer
			Label Seal
			Label Full Text
		About Loose Leaf	Loose Leaf Writer
			Loose Leaf Seal
			Loose Leaf Full Text
		From Collector	Collector Seal Locality
			Collector Seal Inscription
	Inscription	Series Number	
		Position	
		Category	
		Style	
		Content	InscriptionContent Full Text
			InscriptionContent Image
	Decoration	Series Number	
		Position	
		Category	
	Transcription		
	Mount	Volume Cover	
		Protective Covering Case	
		Book Case	
	Release	Edition Name	
		Binding	
		Style Form	Border/ Column
			Center Boundary/ Row
			Block Heart
			Frame Mark
		Lines Per Page	
		Font	
	Exhibition	Exhibition Name	
		Exhibition Size	
		Object Description	Exhibition Description
			Recommendation
			Web Description

Element	Qualifier		
	Condition		
	Grade		
	Notes		
Subject	Subject Descriptor	Primary Subject	
		Secondary Subject	
		Other Subject	
		Situation	
		Function	
		Technique	Series Number
			Position
			Category
		Style And Movement	
		Personal Name	
		Corporate Body	
	Keywords		
Creator	Personal Name		
	Dynasty		
	Birth Place		
	Corporate Body		
	Role		
Contributor	Personal Name		
	Dynasty		
	Birth Place		
	Corporate Body		
	Role		
Publisher			
Date	Cataloging Date		
	Created		
	Issued		
	Acquired		
	Modified		
Identifier	CallNumber		
	AccessionNumber		
	URI		
Source			
Relation	Is Reference Of	Reference Work	
		Collection Catalogue	
		Research Material	
	Has Part	Part Title	
		Part Creator	
		Part Contributor	
		Pagination	
	Is Part Of		
	Citation		
Language	Cataloging Language		
	Item Language		
Coverage	Spatial	Place Of Use	
		Scope Of Coverage	
		Place Of Event	
	Temporal	Period Of Use	
		Date Of Event	
Rights	Owner Name		

Element	Qualifier
	Owner Country

* MICI-DC was developed before “Dublin Core Qualifiers” was announced

6. Metadata Management System -- Metalogy

In order to expedite the cataloging of metadata, the ROSS team developed a general-purpose metadata management system called Metalogy. Metalogy allows a user to input the DTD of a metadata type and generates, automatically, an interface for the user to catalogue metadata, build thesaurus, and other activities related to managing digital archives. If the user chooses to use an existing metadata type instead of designing his own, Metalogy also provides a menu of built-in metadata types such as Dublin Core and MICI-DC. In order to provide maximal flexibility in managing data, Metalogy adopts XML as the basic format for data formulation. We give a more detailed description of Metalogy as follows.

6.1. Features and structure of Metalogy (version 1.0) System

Metalogy, an XML/metadata management system, was developed under the Digital Museum Project funded by National Science Council over a one year period. This system may be used to develop databases for any digital museum, digital library, or digital archive on different subjects. Functions of Metalogy include database set-up via DTD, metadata editors, authority file (thesaurus) editors, retrieval facilities (including both Window and Web interfaces), and importing/exporting of XML files. It includes the following features.

- The system schema is decided according to the DTD employed.
- Metalogy allows the co-existence of different types of DTD.
- Metalogy is capable of retrieving different formats of data at the same time.
- Metalogy allows the user to adjust the element format and access restrictions according to the system schema.
- Metalogy provide a user-friendly interface for the user to define hyperlinks, index, and the retrieval and display of elements.
- Metalogy allows the import and export of data according to their DTD format.
- Metalogy is capable of determining whether the imported data conforms to the designated DTD format and to check for duplication or conflict of data.
- Metalogy can handle structured elements, multimedia objects, and texts.
- Metalogy contains management functions such as access control and transaction log.
- Metalogy provides web searching utilities, which allow external users to retrieve information from the database via Web interface.

Structure of the Metalogy system is shown in Figure 1.

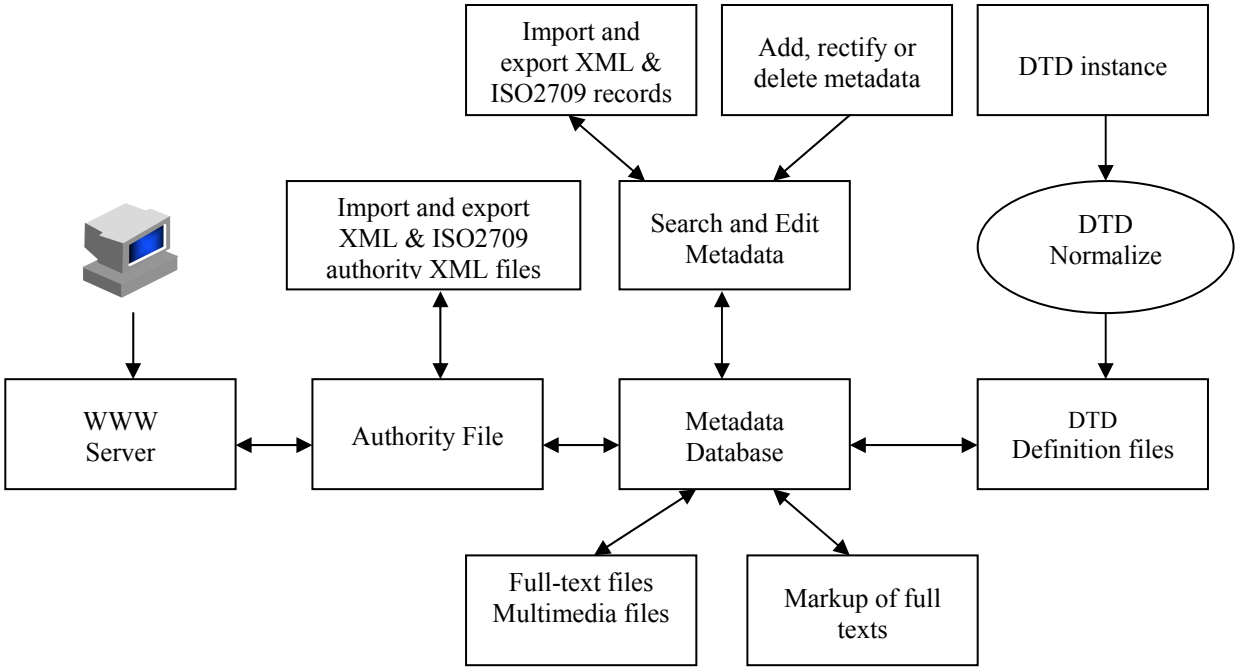

Figure 1: Structure of Metalogy System

6.2 Features and functions of Metalogy

Metalogy was developed in Delphi 5.0. The Web searching facilities use ASP. The back-end database system can be either Oracle or SQL server.

The current version of Metalogy has incorporated the following functions:

a. Load a DTD to set up a database
If the user has already designed a metadata type and its associated XML DTD, then he can simply import it into Metalogy. The corresponding database and cataloging display interface will then be set up automatically. If the user wishes to use an existing metadata, he can then choose from a menu of metadata types provided and maintained by the system. Figure 2 provides an example (in Chinese).

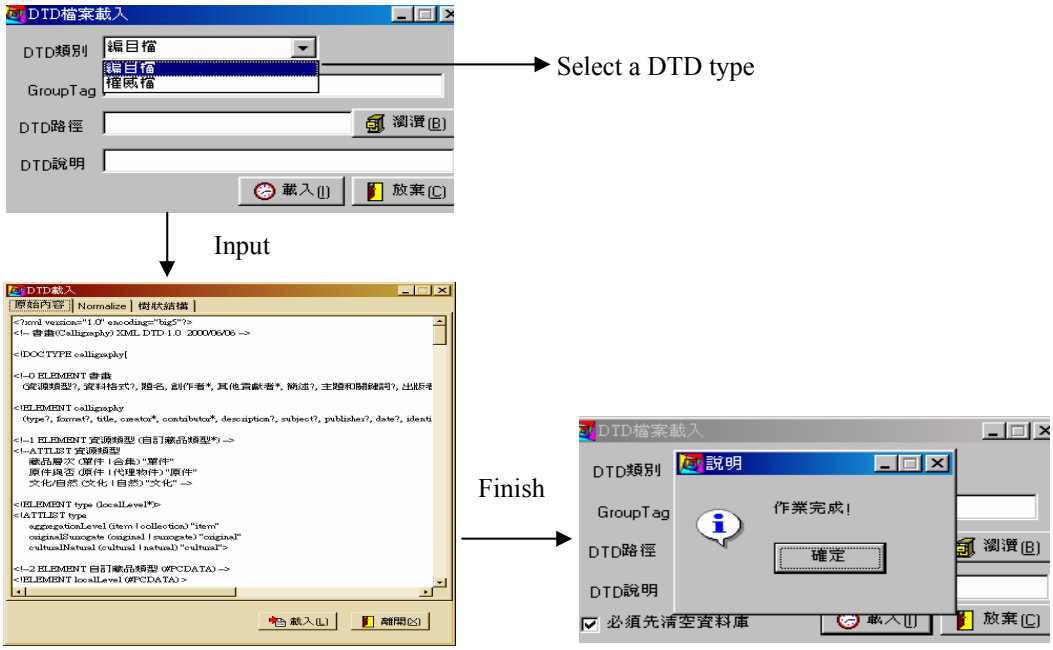

Figure 2: Input DTD

We emphasize that since the data structure of Metalogy is built entirely on XML DTD alone, it can accept any form of metadata, be it a bibliographic file or authority. Neither is the language of the metadata an issue. (The current interface is in Chinese, but the metadata can be designed in any language as long as the character type is accepted.) Once the DTD is fed into Metalogy, it will be transformed into schema in the database, as shown in the following Figure 3.

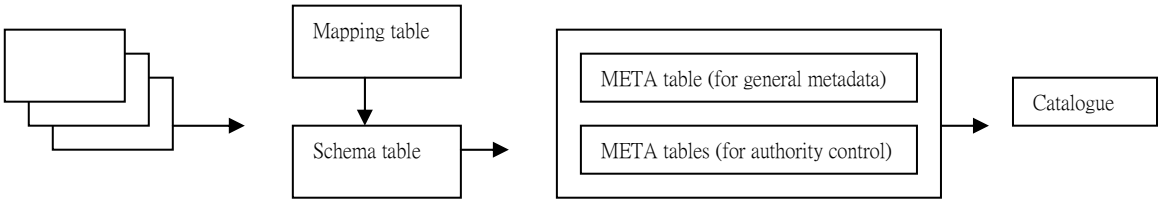

Figure 3:The referencing relationships of tables in importing XML Schema

b. Define the database schema
The XML DTD, as it is defined, does not allow type declarations such as data format, conversion specification, maximum length of input character, authority control and index file. Consequently our current system cannot perform a complete type-checking of an imported DTD. Therefore although the database schema, when importing a DTD, will be generated by the system automatically, the user may need to check manually for any inconformity.

c. Catalogue metadata
After choosing the metadata type and loaded the DTD, the user may add, modify, or delete records in the database. When editing a record, the user may duplicate or delete an element, change its authority control, link to external multimedia objects, or extend its sub-elements based on the mapping. The user can also invoke the query and retrieval interface to check on the current status of the database or any record being maintained.

d. Establish thesaurus and authority file

The process for building the thesaurus and authority file is the same as cataloging the metadata records. The authority control is built automatically by the system according to the database schema. Metalogy provides maximal flexibility for the user in terms of authority control. The user can decide whether authority control is required for a given metadata type. If the same user uses Metalogy to build different databases based on different metadata types, he can choose to have authority control on one database but not on the others. This differs Metalogy from the design of conventional library automation systems.

e. Manage multimedia objects
If the user already has a file or files of multimedia objects with simple description, they can be incorporated into the metadata catalogue by linking. This can be done either as single records individually or in batch mode.

f. Search
Metalogy provides search facilities with a single metadata attribute or several attributes, exact or fuzzy, and Boolean combinations. We refer to Figure 4 for an example.

Figure 4: Search Function

g. Search on authority files
Search can also be performed on the authority files. It is done the same way as searching using metadata fields.

h. Import XML files
Data exchange between Metalogy and external systems is done via XML. Metalogy can import XML files from an external system as long as it recognizes (or after it imports) the DTD. In order to avoid duplication of data, Metalogy allows the user to set conditions to check on specific attributes before importing XML files. An example is shown in Figure 5.

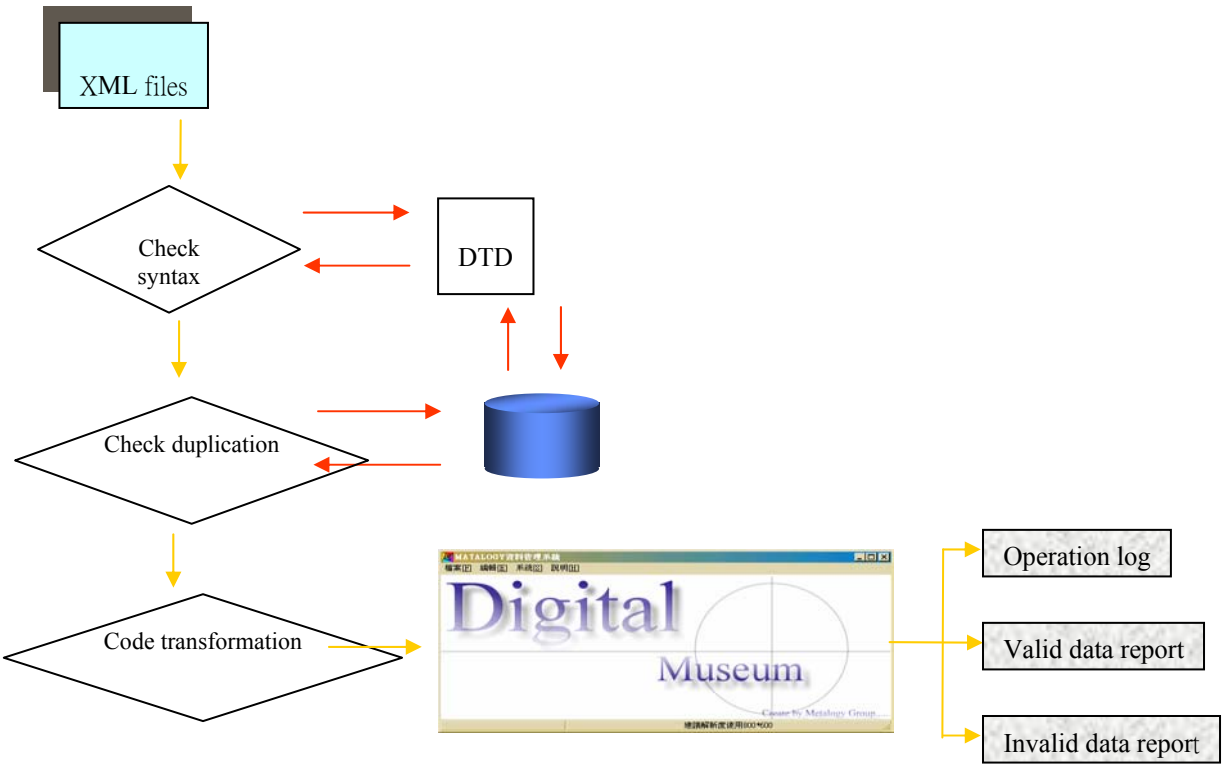

Figure 5: Import XML files

i. Export XML files

Metalogy may exchange metadata with other systems or export well-formed XML files for user access. The user can set, a priori, the metadata attributes that are to be exported or not to be exported. The user can also specify the records to be exported via search. Figure 6 shows an example.

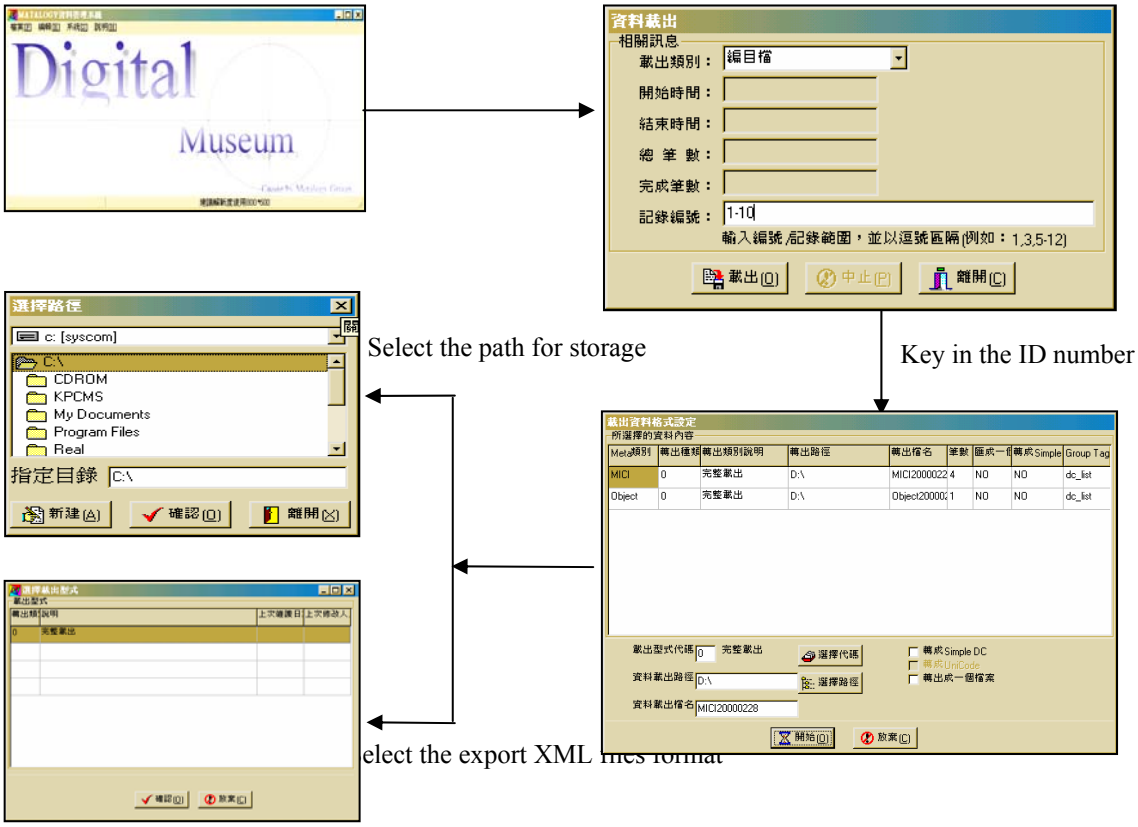

Figure 6: Export XML files

j. Access control

Metalogy provides management functions for establishing basic information and access restrictions on users. When initializing the system, the user needs to give a user name and password, the system will then allow the user to use Metalogy after verification.

k. Message management

Metalogy provides a number of error messages such as “unauthorized usage”, “overwrite existing record”, “wrong password”, etc. They can be edited or modified according to specific user need.

l. Web search facilities

Metalogy provides the same search facilities for external users to search via the Web as for searching from within the database. A Web interface was built for this purpose. See Figure 7 for an example.

Figure 7: Web Search

6.3 Metadata and DTD instances

While designing a digital library/museum/archive, one must design the metadata based on its data. In addition, while using Metalogy, one should use XML/DTD to express metadata. Since designing Metadata requires an in-depth understanding of the characteristics of the target objects (which can be rather time-consuming) and a good understanding of the interoperability issues, it is desirable to provide ready-for-use metadata types. We have incorporated several such metadata types and DTD in Metalogy. Originally designed for cataloging objects of the National Palace Museum, they include metadata types and DTD for calligraphies, artifacts, scriptures, exhibitions, references, name authority files, title authority files, geographical names thesaurus and time thesaurus. They also came with a tagging guide and examples.

6.4 Current status of Metalogy

In order to promote the metadata usage and design ability of Taiwanese research community and companies, we offer the source code of Metalogy free of charge to the public. Not only do we allow them to use, we also encourage them to build and improve upon our system. Metalogy has been extensively used in numerous digital archiving projects. Its users include the National Central Library, National Taichung Library, the Cultural Centers of Ilan County, Taipei County, Hsinchu County, Kaohsiung County, Taichung City, and various academic departments in different universities. The domains of application include Taiwanese theatrical recordings, ceramics, Hukka cultural artifacts, old photographs, puppet shows, indigenous clothing, sculpture, genealogical records, and materials about physical therapy.

7. Strengths and Weaknesses of Metalogy

The greatest strength of Metalogy is its generality and flexibility. However, the same features also become its weakness. Generality often comes at the cost of system effectiveness. Metalogy allows the co-existence of different databases and provides flexible search/retrieval, thus it has to accommodate the co-existence of different index files. As the amount of data become greater and greater, the index files also grow in variety and sizes. This may affect the system response time.

When we started to design Metalogy, a main purpose was to provide novice users (non-programmers) a simple tool to build data quickly. Our goal was to provide a tool so easy to use that a user, without any knowledge of databases, can install the system, choose the target features, build data and the associated database, and provide retrieval facilities on the Web for other users to search through his data. After announcing Metalogy to the public and starting to have users from all sectors of the society, we started to realize that our original intention was too naïve. The users still need to have basic database concepts such as access points, plus some knowledge of XML, metadata, authority control, and client-server architecture to fully utilize what is offered in Metalogy. Indeed, a completely user-friendly system is still a goal yet to be accomplished. We do feel, however, that for an IT worker with minor background in database management systems, a half-day training is sufficient to become an expert in using Metalogy. This provides significant saving in both time and money for many agencies interested in building large digital archives.

8. Other Lessons Learned in Building MICI and Metalogy

During the past four years, the ROSS team gradually built up its experience with metadata design through increasingly more complicated objects and more elaborated scopes of collaboration. Like other pioneering projects, the ROSS team suffered through the initial period of trying to find a common language and an appropriate working model among people from different disciplines. We feel that several factors contributed significantly to the success of our effort. First, the team members were bond by a strong sense of mission and urgency to preserve the indigenous Taiwanese cultural heritage. Second, the team's early participation in the CIMI metadata testbed project gave us a jump-start and hand-on experience on metadata development. Third, during the process of metadata developments, the ROSS team members held frequent discussions with content experts to better understand their information needs and seeking behaviours. Furthermore, discussion meetings were held periodically for all staff to report their work progress and difficulties encountered. In addition to official meetings, we also organized social activities, which played an important role to build personal friendship; it helped people get acquainted and enhanced the willingness of cooperation.

Acknowledgement

The work described in this paper was partially sponsored by the National Science Council of the Republic of China under the NSC grant number NSC88-2745-P-002-007 & NSC-89-2750-P-002-013.

References

[1] Chang, San-Cheng (1999). Digitize national collections (In Chinese). In Science and Technology Advisory Group of the Executive Yuan of the Republic of China (Ed.), The 9th Executive Yuan of the Republic of China（SRB）meeting (In Chinese) (153-182). Taipei: Editor.

[2] Hwang, Jenn-Tai (1999). Greeting a New Millennium -- A Cross-Century Technology Development Program with Concern for the Humanities (In Chinese), National science council monthly 27(7), 715-718.

[3] Wang, Mei-Yu (2000). Achievement of National Science Council Digital Museum Project (In Chinese), National Science Council monthly 28(4), 249-253.

[4] NSC, National Science Council (1999a). National Science Council Digital Museum Project, call for proposal, 2000 (In Chinese) [Announcement]. Taipei. Retrieved Dec. 25, 1999 from the World Wide Web: http://www.nsc.gov.tw/announce/89digi_museum.html

[5] NSC, National Science Council (1999b). The general situation of National Science Council Digital Museum Project (In Chinese). Retrieved Dec. 25, 1999 from the World Wide Web: http://www.nsc.gov.tw/y2k/dml/880209DATA2.html

[6] DMPE, Digital Museum Project Extension, National Science Council (1999). National Science Council Digital Museum Project summary: Digital Museum Project Extension, National Science Council (In Chinese). Retrieved July. 11, 2000 from the World Wide Web: http://mars.csie.ntu.edu.tw/~dlm/plan/1st/intro12.htm

[7] Chen, Hsueh-hua, Chen, Chao-chen, & Chen, Kuang-hua (1999). Metadata Interchange for Chinese Information. In Ching-chih Chen (Ed.), IT and Global Digital Library Development (pp.65-74). West Newton: MicroUse Information.

[8] DMPO, Digital Museum Project Office, National Science Council (2000). An Introduction of National Science Council Digital Museum Project (In Chinese). Retrieved July. 11, 2000 from the World Wide Web: http://dm.ee.ntu.edu.tw/projects.htm