

Designing a Multi-level Metadata Standard based on Dublin Core for Museum data

Jing Wan

Beijing University of Chemical
Technology, China
wanj@mail.buct.edu.cn

Yubin Zhou

Beijing University of Chemical
Technology, China
2011200868@grad.buct.edu.cn

Gang Chen

Beijing Gehua Culture Development
Group, China
cg@sach.gov.cn

Junkai Yi

Beijing University of Chemical
Technology, China
yijk@mail.buct.edu.cn

Abstract

Metadata is a critical aspect of describing, managing and sharing museum data. It is challenging to develop a general metadata schema that meets the requirements of different museums due to the large range of data types. The capability of concise description and the simplicity of use need to be considered. In this paper, we report on a finished project that aims to design a metadata schema for museums in China. An extensible metadata standard based on Dublin Core is presented, which includes core of metadata, extension rules and specific metadata. For the core metadata, we introduce terms, definitions, registration rules and detailed examples of description. The principle of choosing terms and refinements is discussed. A specific metadata schema for porcelain is discussed as an extension example.

Keywords: metadata; Dublin Core; museum

1. Introduction and Motivation

With the rapid development of information technology since the 1992, lots of museums adopted collection management systems, digitalized collection data, and provided public. Data sharing and integration among museums became important.

Metadata is defined as “structured data about data”. As a key issue of data standardization and data sharing, metadata for cultural heritage has attracted worldwide attention. A number of organizations and initiatives made great efforts to address this issue. Some published metadata schemas have been widely used and accepted as international standards, for example, Dublin Core (DCMI, 2012), CDWA (Getty Research Institute, 2008), EDM (Europeana Foundation, 2013), CIDOC CRM (CIDOC CRM Special Interest Group, 2011), VRA Core (Visual Resources Association Data Standards Committee, 2007), EAD (Society of American Archivists and the Library of Congress, 2002), and FGDC/CSDGM (Federal Geographic Data Committee, 1988).

China's management system of cultural relics is different from that of other countries. Most cultural relics are owned by the state and under the protection of the state. A state department takes charge of the work concerning cultural relics throughout the country. From 1978, a serial of regulations were published by China's State Administration of Cultural Heritage, which aimed to establish the standard process in registering and compiling files for museum collections. Many government funded projects promoted the work of museum informatics. The project “Cultural Relics Census and Collection Management System Construction” started in 2001, with 48,006 pieces of valuable collections and 1,370,000 pieces of general collections recorded in the database by 2010. In 2012, the project “First National Movable Cultural Relics Census” started, which aimed to investigate, identify and register movable relics through information technology.

Many museums in China established collection management systems and digitized their collections progressively, such as, the Palace Museum, the Capital Museum, and Shanghai Museum. Some museums designed their own data specifications. And several specifications were published by the government, for example, “Data Specification for Museum Collections”, “Standard for Image Archive of Unmovable Cultural Relics”, “Data Specification for the Third National Heritage Sites Census”, and “Data specification for the First National Movable Cultural Relics Census”.

But there is still no national standard for museum data in China. Considering the different management systems, it is difficult to utilize existing metadata schemas without modification. And museums are different in collection types, collection quantities, data quality and the skill levels of staff. So different requirements for metadata need to be considered. The metadata schema should be capable of concise description, be simple to use, and be compatible with the published specifications.

We describe an effort in developing a metadata architecture to address this issue. In the project, we design the core metadata based on Dublin Core, and specific metadata extensions for drawings, porcelain, ancient buildings and inscriptions. For each metadata of these categories, we provide terms, definitions, refinements, registration rules and detailed samples. In this paper, we focus on the core metadata and describe one specific example of metadata extension.

2. Metadata Architecture

Figure 1 shows the metadata architecture, which includes the core metadata, specific metadata and extension rules.

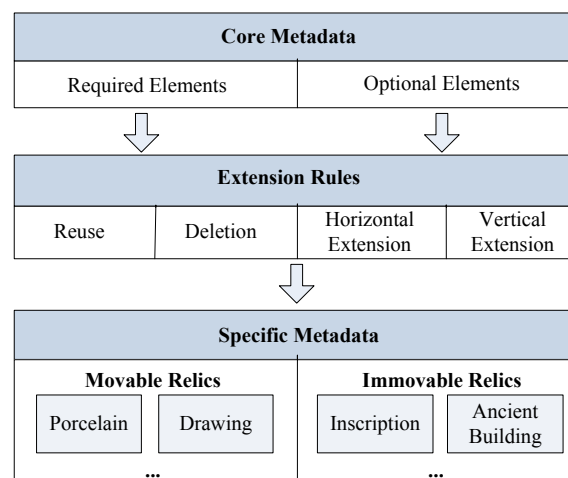


FIG. 1. Metadata architecture

The core metadata is simple and based on Dublin Core Metadata Element Set, version 1.1 (Dublin Core, 2012). This level is used to describe the general core attributes of digital resources. It supports retrieval, integration and data exchange. The elements of this level are easy to use. A museum that has simple data could use it directly. And a museum with a large number of collections and complex data structures can use it as the first stage of a plan. For these museums, entering complete data usually takes several years or even decades. First make it work, and then make it better. This rule is helpful to motivate the staff, get support from other divisions, and gain experience.

The specific metadata is used for data sets of particular type or domain. It is designed by analyzing existing archives and possible data requirements coming from museum management.

The extension rules are used to extend metadata to meet the actual requirements of a specific museum. Rules and implement approaches need to be provided to guide users in customizing metadata.

3. Core Metadata

3.1. Approach

The Core Metadata consists of an elements set and qualifiers. It is a vocabulary of nineteen properties for use in digital museum's collection description. "Core" means its elements are generic, and usable for describing a wide range of museum data.

Taking into account versatility, scalability, and interoperability, we design the core based on the Dublin Core Metadata Element set. In addition, data specifications and published standards in China are considered. Existing data are stored in the database or on paper are analyzed. We also consider the data elements adopted by the cultural relics census. Using this approach, we adopt eleven elements from Dublin Core and add eight elements and qualifiers.

Element qualifiers make the meaning of an element narrower or more specific. Following the practice of Dublin Core Qualifiers, there are two classes of qualifiers, element refinements and encoding schemes. The element refinements include object qualifier, basic qualifiers, and composite qualifiers.

1. **Object qualifier.** The metadata should be capable of describing movable and immovable cultural relics. But these two types of relics have great differences. This qualifier is used to describe the range of an element.
2. **Basic qualifier.** It is the basic unit of qualifier. It cannot be extended.
3. **Composite qualifier.** It consists of basic qualifiers and/or composite qualifiers. For example, the copyright of the image has a composite qualifier including three basic qualifiers—owner, copyright restriction, and copyright description.

We define each element and qualifier by nine properties, which are name, identifier, version, definition, repeatability, data type, required status, domain, and qualifier.

3.2. Element Set

The element set of the core metadata includes nineteen terms. We adopt terms from Dublin Core Metadata Element Set, version 1.1 with the exception of *language*, *contributor*, *publisher* and *source* (DCMI, 2012). For collections in China, the *language* element always has the value "in chinese", so we don't adopt it now. The *contributor* element and the *publisher* element of a collection are the same as its keeper, which is included in the element *rights*. So we don't adopt *contributor* or *publisher*. We ignore the element *source* for it has no value for a collection. Table 1 shows the correspondence between the core metadata elements and the Dublin Core metadata elements. Many of these terms have basic constraints.

We describe standard vocabularies for some elements. The value "Yes" of the Encoding Scheme column of Table 1 on the following page indicates that vocabularies for the element are provided. For example, the *grade* of the movable cultural relics includes the values "grade one", "grade two", "grade three", "not determined", and "normal". These terms are defined in the standard "Grading Standard For Cultural Relics" published by China's Ministry of Culture.

TABLE 1: Alignment of the core metadata element set and DC element set.

Term	Comment	Refinements	Encoding Scheme
Name	DC: Title	Registered Name, Alternative Name	
Identifier	DC: Identifier		
Type	DC: Type		Yes
Date	DC: Date		Yes
Subject	DC: Subject		
Description	DC: Description		
Creator	DC: Creator		
Coverage	DC: Coverage	Geographic Coordinate, Scope Coordinates(Measure point number, Measure Point Coordinates, Adjacent Measure point), Geographic Name	
Right	DC: Rights	Ownership Type, Affiliation	Yes
Relation	DC: Relation	Image, Reference, Component	
Material	DC: Format	Material Type, Specific Material	Yes
Acquisition		Approach, Enter Scope, Enter Date	Yes
Grade			Yes
Measurement		Dimension(Length, Width, Height), Weight, Distribution Area, Protection Scope Area, Building Area, Construction Control Zone Area	
Conservation		Residual Level, Conservation Status, Status Assessment,	
Quantity			
Condition		Use Unit, Subordination Unit	Yes
Environment		Natural Environment, Humanities Environment	
DamageCause		Natural Cause, Man-made Cause	Yes

4. Extension Rules

Because of the large range of museum collections, it is hard to use the core metadata to meet the description of each item. So we design the extension rules to generate more specific metadata. And we provide the design of four specific metadata, which includes terms, definitions, registration rules and detailed examples.

There are four classes of extension approach:

Reuse. It refers to adopting existing elements or refinements of the core metadata. It includes complete reuse and partial reuse. The reuse class indicates adoption without modification. Partial reuse adds some restrictions.

Deletion. Refers to deleting elements or refinements that are useless in this level.

Horizontal extension. Refers to adding a new element.

Vertical extension. Refers to adding refinements according to the extension rules.

5. Metadata for Porcelain

The specific metadata for porcelain is an example of how the extension rules are applied. Table 2 shows how the specific metadata for porcelain is extended from the core metadata. It includes sixteen elements. The followings are examples of four extension rules with the porcelain metadata:

1. **Reuse.** The element name and its two refinements (registered name and alternative name) from the core metadata are included in the specific metadata. It is complete reuse. The element grade from the core metadata is included in it too. But the value range of the element grade is changed, so it is part reuse.
2. **Deletion.** The element coverage has three refinements in the core metadata. We delete one refinement (scope coordinates) in the specific metadata for it is useless for porcelain.

3. **Horizontal extension.** There is no horizontal extension.
4. **Vertical extension.** The element name has a new refinement (original name). We add it because many original names of porcelain collections are revised in order to conform to the naming rules published by the authority. The revised name is the registered name of a collection. But sometimes the original name is well known. So we need to record it too.

TABLE 2: Specific metadata for porcelain.

Index	Term	Refinements	Extension
1	Name	Registered Name, Alternative Name, Original Name	Complete Reuse+vertical Extension
2	Identifier		Complete Reuse
3	Type		Part Reuse
4	Date	Manufacture Date, Use Date	Vertical Extension
5	Subject		Vertical Extension
6	Description	This term has 17 refinements.	Vertical Extension
7	Creator	Name, Gender, Native Place, Birth, Death, Creator Description	Vertical Extension
8	Coverage	Geographic Coordinate, Geographic Name	Deletion+Complete Reuse
9	Right	Ownership Type, Affiliation	Complete Reuse
10	Relation	Image, Reference, Component	Complete Reuse+Vertical Extension
11	Material	Material Type, Specific Material	Complete Reuse
12	Acquisition	This term has 12 refinements.	Complete Reuse+Vertical Extension
13	Grade		Part Reuse
14	Measurement	Dimension (Length, Width, Height), Weight	Deletion+Complete Reuse
15	Conservation	Current Condition, Natural Damage, Physical Damage, Remarks, Citations	Complete Reuse+Vertical Extension+Deletion
16	Quantity		Complete Reuse

5. Conclusions and Future Work

This paper introduces a project aimed to design an extensible metadata standard for museum data in China. We consider the capability of concise description and the simplicity of use. We present a standard including core metadata, extension rules, and specific metadata. The core metadata is based on Dublin Core and is easy to use. It includes nineteen elements and refinements. There are four extension approaches that are reuse, deletion, horizontal extension and vertical extension.

In the future, we plan to develop a metadata management system, which will help museums to customize the metadata element set for their application. We also plan to enhance the use of standard vocabularies and make them compatible with the international standards.

Acknowledgements

We would like to express our thanks to the Capital Museum for providing data samples. We also gratefully appreciate the contribution of Hu Chui from the Palace Museum, Xiao Fei from the National Museum of China, Shen Guihua from China Cultural Heritage Information and Consulting Center for the discussion on the metadata. This work is financed by the State Cultural Relics Bureau within the project “*A Study of a Metadata Standard for Museum Data in China*”.

References

- China's State Administration of Cultural Heritage. (2001). The Data Specification for Museum Collections. Retrieved April 1, 2014, http://www.nach.gov.cn/art/2008/7/9/art_343_3636.html.
- China's State Administration of Cultural Heritage. (2005). The Standard for Image Archive of Unmovable Cultural Relics. Retrieved April 1, 2014, http://www.sach.gov.cn/art/2008/7/8/art_343_3633.html.

- China's State Administration of Cultural Heritage. (2007). The Data Specification for the Third National Heritage Sites Census. Retrieved April 1, 2014, from <http://pucha.sach.gov.cn/tabid/64/Default.aspx>.
- China's State Administration of Cultural Heritage. (2012). The Data Specification for the First National Movable Cultural Relics Census. Retrieved April 1, 2014, <http://www.wenwu.gov.cn/kydwwpc/>.
- CIDOC CRM Special Interest Group (SIG). (2011). Definition of the CIDOC CRM. Retrieved April 1, 2014, from http://www.cidoc-crm.org/definition_cidoc.html.
- DCMI. (2012). DCMI Metadata Terms. Retrieved April 1, 2014, from <http://dublincore.org/documents/dces/>.
- Europeana Foundation. (2013). EDM Definition. Retrieved April 1, 2014, from <http://pro.europeana.eu/documents/900548/770bdb58-c60e-4beb-a687-874639312ba5>.
- Federal Geographic Data Committee. (1998). Content Standard for Digital Geospatial Metadata. Retrieved April 1, 2014, from http://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/base-metadata/index_html.
- Getty Research Institute. (2008). Categories for the Description of Works of Art (CDWA). Retrieved April 1, 2014, http://www.getty.edu/research/publications/electronic_publications/cdwa/.
- Metadata Architecture and Application Team and Taipei National Palace Museum.(2007). The Information Requirements Specification for Taipei National Palace Museum. Retrieved April 1, 2014, from <http://www2.ndap.org.tw/eBook08/showContent.php?PK=51>.
- Society of American Archivists and the Library of Congress, (2002). Encoded Archival Description. Retrieved April 1, 2014, <http://www.loc.gov/ead/>.
- Visual Resources Association Data Standards Committee. (2007). VRA Core 4.0 Introduction. Retrieved April 1, 2014, from <http://www.loc.gov/standards/vracore/schemas.html>.