

Towards Digital Archive Systems: Architecture and Design of Digital Museum Archive

Hiroyuki Kawano

Nanzan University, Aichi 4890863

Abstract In the information society, the volume of born-digital contents and digitized contents from original resources is rapidly increasing. Therefore various types of digital contents, such as text, images, movies and other data types, are stored into digital archive systems and key technologies are long-term preserving storage, database management system, information retrieval, web service and others. The one of major application programs is CMS (Contents Management System), some types of CMSs focus on specific functions in order to administrate digital contents and support a work flow in museums. In this paper, using several CMS softwares, we also develop a digital gallery of "*the Anthropological Museum of Nanzan University*" in order to organize various collections with metadata and photographs. Firstly, we consider metadata attributes based on the guideline of museum objects, and we propose XML metadata formats like URI/RDF/MODS. Secondly, we design the metadata schema of museum collections, we also validate our proposed schema by using the MSXML parser. Thirdly, we construct the prototype archive system by using Joomla and the package of gallery and next we provide the annotation function in the system, which is based on Drupal and related modules. In the archive system, we have 173 digital contents with 74 metadata attribute values. At present, we are trying to implement the similar image search for 40,000 photographs stored in the system.

Keywords Digital Archive, Digital Museum, CMS, Annotation, Similar Image Search

1 Introduction

Various types of digital contents, such as text, images, movies and other data types, are stored into digital archive systems, and key technologies are long-term preserving storage, database management systems, information retrieval, web services and others[6, 7, 8].

For instance, a first important academic result of museum digital archives is "National Digital Archives Program (NDAP)" (http://www.ndap.org.tw/index_en.php). At present, many famous museums (<http://2k.si.edu/>, <http://www.louvre.fr/>, <http://www.hermitagemuseum.org/>) are making efforts to provide digital museums in order to display their own collections.

In order to administrate and search museum digital objects, various CMSs (Contents Management System) of database applications play an important role[5]. CMS supports a museum work flow in order to index, search, manage and archive various kinds of digital contents totally. Dspace (<http://www.dspace.org>) is one of famous softwares in order to construct open digital repositories. For example, the system of "Chinese Virtual Olympic

Museum" has been constructed by Dspace platform[11]. The mashup technology is also becoming important in order to integrate different information resources in CMSs by using standard APIs.

In this paper, using some CMS softwares and related modules, we develop the digital museum archive of "*the Anthropological Museum of Nanzan University* (<http://www.nanzan-u.ac.jp/MUSEUM/>)" in order to organize various museum collections with metadata and photographs.

In Section 2, in order to design metadata schema, we focus on the "International Guidelines for Museum Object Information" (<http://icom.museum/guide.html>)[3]. We propose metadata attributes based on the guideline of museum objects, and we design XML metadata formats like URI/RDF/MODS (Metadata Object Description Schema, <http://www.loc.gov/standards/mods/>)[10]. We define the metadata schema of museum collections, and we validate our proposed schema by using the MSXML parser[9].

Next, we develop the prototype archive system by using Joomla and some modules described in Section 3, and we construct the museum archive based on Drupal and related packages of gallery and annotation modules, which are presented in Section 4. In the digital museum archive, we have 173 digital contents with 74 attribute values in the metadata schema.

2 Metadata formats in anthropological museum

In order to design metadata schema of museum contents in the anthropological museum, we mainly focus on the "*International Guidelines for Museum Object Information*" (<http://icom.museum/guide.html>) and "*CIDOC Conceptual Reference Model CRM models*"[2]. This guideline only provides the conceptual framework for managing museum objects, therefore we have to propose metadata attributes of information management and technical attributes of photographs and movies, which are based on other metadata formats. In this section, we consider other metadata such as Dublin Core of Dublin Core Metadata Initiative[1], MPEG-7 and MPEG-21 for multimedia contents[4].

CRM model defines the relationships of various entities, such as "*Man-Made Object*", "*Production*", "*Time Primitives*" and others. In Dublin Core, fifteen basic elements are selected and standardized, we choose following elements in our system; "*description*", "*date*", "*type*", "*format*", "*language*", "*coverage*" and "*rights*". We also consider metadata attributes of photographs based on MPEG and JPEG standards, and we design our XML metadata formats.

In our system, we also need multilingual attributes, so we define multilingual fields based on MODS metadata schema. A <Object_Name_Information> has child elements <Title_Information>, which have <lang> and <type> attributes. The description in the attribute <lang>, we use ISO 639-2b, *jp*: Japanese, *eng*: English and so on. Finally, we use following 22 information groups and design detail metadata schema based on some examples of the museum collections.

- Acquisition Information
- Condition Information
- Deaccession and Disposal Information
- Description Information

- Image Information
- Institution Information
- Location Information
- Mark and Inscription Information
- Material and Technique Information
- Measurement Information
- Object Association Information
- Object Collection Information
- Object Entry Information
- ObjectName Information
- ObjectNumber Information
- Object Production Information
- Object Title Information
- Part and Component Information
- Recorder Information
- Reference Information
- Reproduction Rights Information
- Subject Depicted Information

For example, Condition Information has categories of “Condition”, “Condition Summary” and “Condition date”, we extend these categories for administrators and we add several attributes, “Receiving Date”, “Transportation”, “Receiver” and “Return Date”. In the information group of “Image Information”, we need more detail attributes, such as Title, Image Reference, Keyword, Comment, Medium, Film Instrument, Artist, Film Date, Width, Height, Size, Size Unit, Film Place, Film Latitude Longitude, Photo (including horizontal and vertical resolutions and bit depth etc.) and so on. Initially, we propose and design a metadata schema with 102 attributes, and we validate our proposed schema by using the MSXML parser[9], which provides extension functions of DOM, SAX, XSLT and others.

3 Prototype system of museum archive

In Fig.1, we present the conceptual architecture of the museum CMS, the user send a query to the server and the contents management system retrieves various data, such as text, photograph, metadata and layout from different databases, the system dynamically creates a web page based on a user’s preference. The granularity of information depends on the grant of users, such as administrators, researchers, guests and so on.

By the way, there are various CMSs, such as Drupal, Joomla, MovableType, XOOPS and many others, we develop the prototype archive system having photo gallery by Joomla and related module softwares. We evaluate the comparison of accessibility between museum CMS and the original museum web site, the performance of CMS is better than the original web site from various view points, such as contrast of colors, stylesheet of layout design and control, design of link navigations, W3C standard descriptions, consistency of hyperlinks and etc. However, our system does not satisfy the conditions of “Double-A” of “Web Content Accessibility Guidelines 1.0”[12].

In order to improve our prototype system, after increasing of volume of digital contents, we reconstruct digital archive by using Drupal presented in Fig.2. The top page of

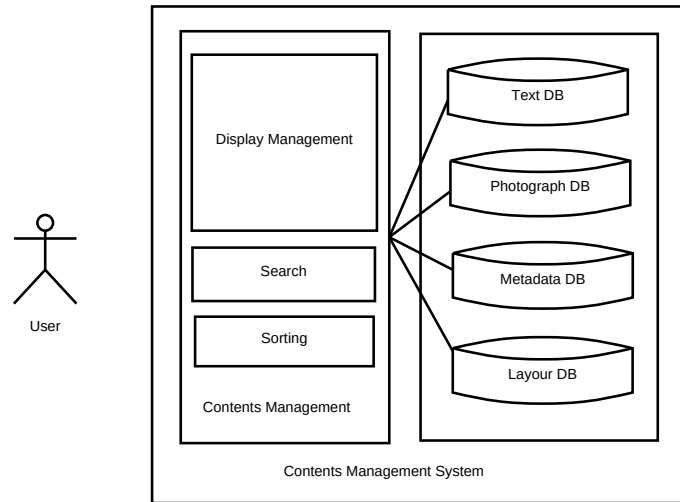


Figure 1: Conceptual framework of museum archive

gallery shows a few collections having a part of attribute values.

4 Implementation of museum CMS

In order to provide the annotation function in the system, we compare various tools and web sites, such as PhotoStuff, M-OntoMat-Annotizer, Taggify and so on, but it is so difficult to integrate other different modules in the specific CMS. So, we reconstruct the museum archive based on the combination of Drupal and related packages of Gallery2 and Fotonotes annotation module.

When we construct the museum archive, we consider the stability and extensibility of the system, then we select Drupal (drupal-5.10.tar.gz) and Japanese module (DN_ja_resource_kit_5__7_009.tgz). In order to support annotation functions of photographs, videos and other content, we select several modules of “Gallery2” (gallery-2.2.6-developer.tar.gz), “Fotonotes” (fotonotes-5.x-1.1-beta6.tar.gz) and other tools into the Drupal. The detail environment is as follows:

- Ubuntu 7.04
- Apache 2.2.3
- Database (MySQL 5.0.38)
- PHP 5.2.1
- Drupal 5.10
- Gallery Developer 2.2.6
- Fotonotes 5
- Image module 5

At present, we store 173 digital contents into our museum archive system, and we define 74 attributes, a part of attributes are shown in Fig.3.

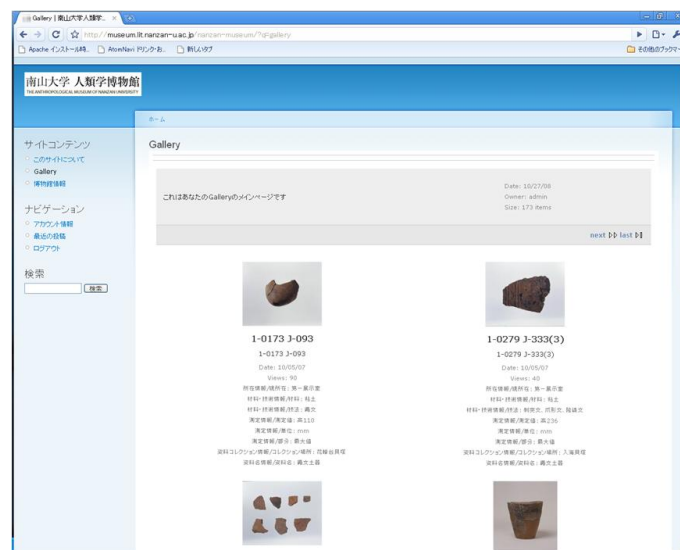


Figure 2: Top page of museum archive

5 Conclusion and future works

In this paper, using Joomla, Drupal and related modules, we introduce a prototype of digital gallery of "the Anthropological Museum of Nanzan University". In order to organize various collections with metadata and photographs, we propose metadata attributes and design XML metadata formats. At present, we provide the museum CMS, which has 173 digital contents with 74 metadata attributes, for research staffs in the museum. We are also implementing the annotation function in the archive system. In this year, we are storing various photographs and implementing the similar image search function of those photographs stored in the system. The museum collection consists of more than 40,000 Andean ethnographic photographs including the landscape, portraits, ceremonies and others.

Acknowledges

A part of this work is supported by "the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), 19500098, 2009" and "2009 Nanzan University Pache Research Subsidy I-A-2". Under the open research project and collaborative work with staffs in the anthropological museum, the archive system was constructed by students, Mr. M. Iwata, Mr. S. Kato and Miss C. Ohara, Miss T. Kusakari, Miss K. Ichikawa and Miss R. Kawai, who graduated from Department of Information and Telecommunication Engineering in Nanzan University.

References

- [1] Dublin Core Metadata Initiative, "Dublin Core Metadata Element Set, Version 1.1," 2008. (<http://dublincore.org/documents/dces/>)

2009/07/30

南山大学 人類学博物館
THE MUSEUM OF ANTHROPOLOGY, NANSAN UNIVERSITY

ホーム > Gallery
1-0397 J-030

1-0397 J-030

ダウンロードを付ける

Date: 10/05/07
Size: 
Full size: 1000x809

first previous next last



Photo Properties
summary details

Color Space	sRGB	Date/Time	Fri Oct 5 20:14:20 2007
Custom Fields			
取得情報/方法: 移写			
取得情報/日付: 不明			
取得情報/取得元: 日本考古学研究所→考古学研究所→南山大学人類学研究所付属陳列室			
状態情報/状態: 修復済あり			
状態情報/注釈: 従元年月日不明 復元部表面は2006年彩色			
状態情報/日付: 2006.10.02			
記述情報/物理的記述: 口縁に1対の大小の把手と1対の同様の把手が付いている。口縁部には渦巻状の唐紋文と爪形文が施されている。			
記述情報/基本現状: なし			
国際情報/タイプ: ポン			
組織情報/経路名: 南山大学			
組織情報/住所: 愛知県名古屋市昭和区山屋町10			
組織情報/国名: 日本			
所在情報/所在地: 第一展示室			
所在情報/取得日付: 2006.10.02			
所在情報/タイプ: 展示中			
配号・細印情報/細印文: E C78			
配号・細印情報/タイプ: 注記			
配号・細印情報/細印説明: 『E』は「亀山」, 『C』は「C地点」と思われる			
配号・細印情報/技法: 注記			
配号・細印情報/位置: 口縁内面			
配号・細印情報/言語: 英, ローマ字, アラビア語			
材料・技術情報/材料: 粘土			
材料・技術情報/技法: 唐紋文, 爪形文, 網文			
測定情報/測定値: 口径271 高205			
測定情報/単位: mm			
測定情報/部分: 最大値			
資料関連情報/関連場所: 千代田市川市橋井町			
資料コレクション情報/コレクション場所: 南山大学			
資料コレクション情報/コレクション日付: 1940.0201-29			
資料コレクション情報/コレクター: G.Groot, H.Shong			
資料コレクション情報/方法: 撮影			
資料受入情報/受入所有者: 南山大学人類学博物館			
資料受入情報/受入日付: 不明			
資料受入情報/受入番号: 不明			
資料受入情報/受入理由: 考古学研究所			
資料名情報/資料名: 網文土器			
資料番号情報/資料番号: 1-0397 J-030			
資料番号情報/日付: 不明			
資料製作情報/製作日付: 縄文時代中期			
資料タイプ情報/タイプ: 唐紋式			
物品・構成情報/構成数: 1			
物品・構成情報/物品・構成説明: 口縁部のみ 4突起あり			
記録者情報/記述者: 佐藤隆雄			
記録者情報/記録日付: 2007.10.15			
記録者情報/典拠: 『亀山県史』『南山大学人類学博物館紀要第19号』『日本考古学研究所台帳』			

Figure 3: An example of detail output with metadata attributes and values

- [2] ICOM/CIDOC CRM Special Interest Group, "Definition of the CIDOC Object-oriented Conceptual Reference Model, Version 3.2," 2001.
(http://cidoc.ics.forth.gr/docs/cidoc_crm_version_3.2.rtf)
- [3] International Council of Museum, "Information and Documentation - A Reference Ontology for the Interchange of Cultural Heritage Information," ISO 21127:2006, 2006.
(http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=34424)
- [4] ISO/IEC JTC1/SC29/WG11, "MPEG-7 Overview," 2004.
(<http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>)
- [5] J.-S. Hong, B.-H. Chen, J. Hsiang, "XSL-based Content Management for Multi-presentation Digital Museum Exhibitions," ECDL 2001, pp.378-389, 2001.
- [6] H. Kawano, "Web Archiving Strategies based on Web Log Mining Patterns," 2004 CORS/INFORMS International Meeting INFORMS, TC18, 2004.
- [7] H. Kawano, "Steps Towards Better Digital Archives," Digital Archive Workshop on Human Civilization, Digital Archive Forum with APAN and APNG, China, 2007.
- [8] H. Kawano, "Reputation-based Contents Crawling in Web Archiving System," Lecture Notes in Operations Research, World Publishing Corporation, Vol.8, pp.317-324, 2008.
- [9] Microsoft, "MSXML 4.0 (Microsoft XML Core Services)," 2006.
(<http://support.microsoft.com/kb/269238/ja>)
- [10] MODS Editorial Committee, "Metadata Object Description Schema (MODS)," (<http://www.loc.gov/standards/mods/>)
- [11] A. Stuart, "Capturing China's Knowledge: Ancient Terracotta Warriors, Scientific Discoveries, even 2008 Olympics to Go Online", 2007.
(http://www.hpl.hp.com/news/2007/oct-dec/dspace_china.html)
- [12] W3C, "Web Content Accessibility Guidelines 1.0," 1999.
(<http://www.w3.org/TR/WCAG10/>)