

Metadata and the way of encoding of the collections in Shaanxi History Museum

Yangzeyu Liu

Beijing Institute of Fashion Technology, 100029 No.2 Yinghua Road, Chaoyang District, Beijing, China

202220001021@stu.bift.edu.cn

Abstract. The contemporary advancement of digital museum studies has emerged as a significant subfield within the broader discipline of museum studies and is poised to have a profound influence on the future trajectory of museology. The Shaanxi History Museum is a provincial museum located in Shaanxi, China. It has the distinction of being the first national museum in China to have contemporary facilities on a big scale. Situated in Xi'an, a city with a historical legacy spanning over a thousand years as the capital of thirteen dynasties, the location has a significant cultural and historical heritage, as well as abundant archaeological riches. This paper explores the qualities and limitations of metadata encoding in the context of the Shaanxi History Museum. It does so by analyzing the collection metadata from two perspectives: page search and collection data files. The study also suggests further encoding options about the collection circulation, museum data administration, and display requirements, considering the future trajectory of digitizing museum collections.

Keywords: Shaanxi History Museum, Digitalization, Metadata.

1. Introduction

The proliferation of digital technology across several domains has resulted in the widespread adoption of digitization in everyday life, making it a commonplace occurrence. The facilitation of information accessibility and its subsequent interactive distribution has played a significant role in fostering the advancement of cultural idea transmission. In The 50th Statistical Report on China's Internet Development, it is shown that the scale of Internet users in China has continued to grow steadily. Internet users spend 29.5 hours on the Internet per week, and the rates of browsing Internet news and short videos are 75.0% and 91.5% respectively, with user sizes reaching 788 million and 962 million people [1]. The demand for and use of digital resources is considerable.

In a museum, a public cultural institution, activities such as daily management, exhibitions, research and services, promotion, and publicity are being digitized in line with the development needs of society [2-4]. How to better encode and manage information to promote online interactions in all aspects, as well as to more specifically meet the retrieval needs of different groups of people for museums, are the challenges that museums are facing [5]. However, current research shows that the issue of encoding data about museum collections has not been discussed in more depth and specificity in China [6, 7].

Shaanxi History Museum is a provincial local museum in Shaanxi, China, and the first large-scale national museum with modern facilities in China. Being located in Xi'an, which has a history of more than 3,000 years as the capital of thirteen dynasties, it has a profound cultural and historical background and rich archaeological resources. Therefore, there is a huge scale of cultural relics collection, with a total of 1717,950 pieces and a more systematic and mature exhibition, which can be a good case for the study of digitization of local museums in China [8].

2. Methods

The article adopts the methods of case study, literature survey, content, and local culture analysis. The article will first analyze and summarize the metadata content of the collection data of Shaanxi History Museum, and then compare and analyze it with the relevant documents, programs, and research on the encoding of museum collections, to summarize the characteristics and shortcomings of the museum's own way of digitizing its collections. Secondly, it would trace the causes in terms of local traditions, historical and cultural concepts, and departmental handovers. Taking into account the current needs of different aspects such as visits, research, restoration, etc., and centering on the use of collections, a more comprehensive collection information encoding method is envisioned that is rooted in the local characteristics of the museum itself, and that meets the specific needs of the use of the collections.

2.1. Retrieved metadata presented by the web page

The Shaanxi History Museum is a museum of history and humanities based on locally excavated artifacts in Shaanxi. On the Chinese page of the official website, the collection information is presented in three sections: "Treasures of the Museum", "Collections of Excellence" and "Cultural Relics Data". "Treasures of the Museum" and "Collections of Excellence" sections are on the same page. The "Treasures of the Museum" is a representative collection of the museum, which can best reflect the culture of a certain historical period and has exquisite skills. Using the pictures as an index, click to enter the details page, more information can be seen, including the name of the artifacts, the age, the level of the artifacts, the place where they were excavated, the size, the weight, and the material, whether they are on display or not, as well as the historical background, the utility and other related introductory instructions, and also information on the exhibitions of the collection, and the name and the location of the exhibition halls.

The "Collection of Excellence" presents images of 6,655 items in 55 pages. On the details page, unlike "Treasures of the Museum", not all artifacts in the "Collection of Excellence" are labeled with full descriptions of each metadata. (See Table 1 below for details: information marked with “√” is that of the "Treasures of the Museum" or the "Collection of Excellence", while those not marked refer to the absence of such corresponding information for some collections).

Table 1. Comparisons between Collection of Excellence and Treasures of the Museum.

	Name	Era	Artifact Level	Place of Excavation	Size	Material	Introduction	Related Exhibition	On/Not on Display	Image
Treasures of the Museum	√	√	√	√	√	√	√	√	√	√
Collection of Excellence	√		√		√	√			√	

For example, in "Collection of Excellence", a collection called "Northern Wei Painted Warrior Figurines on Horseback" lacks information on "Place of Excavation", while the description of the information of "Material" corresponds to the content of "Sculpture and Statue", which is not related to the information of the material itself. According to the Specification for registration of cultural relics in the collection of cultural institutions (WWT0017-2013, 6, specification for registration of cultural relics in the collection of cultural institutions), "Sculpture and Statue" should belong to the "Artifact Category". A collection named "Western Zhou perforated sea shells", its "Material" corresponding to the content of the "jade", is inconsistent with the name of the collection itself. In addition, in some of

the collections labeled "On Display", certain exhibit information is missing or does not correspond to a particular collection. For example, a collection titled "Tang Gilt Copper Pave Head" is missing information and the name of the exhibition hall where it is displayed.

This page also provides a search box for users to search by keywords. Three different options, "Era", "Material" and "Exhibition title", are available to define the scope of the search object for more specific queries. Compared with the collection metadata, the search metadata provides limited choices, and does not provide the corresponding "Place of Excavation", "Artifact Level", "On/Not on Display" and other filtering and searching options in the collection information. Therefore, the encoding plan in terms of searching is relatively simple and lacks specificity. In general, the digitized coding is not very different from traditional archival records, except for the coding of "On/Not on Display", which facilitates visitors in deciding their need to arrange their visit to the museum.

2.2. *The way the collection data is encoded*

In the "Cultural Relics Data" section of the web page, the Shaanxi History Museum provides 35 files of public data, including embroidery, leather, oracle bones, porcelain, coins, gold, and silverware, by categorizing the collections based on their materials. These files contain 435,690 entries, accounting for about 25% of the total collection. Among them, "coins" are the most numerous, accounting for about 80% of the total data of the collection, followed by "pottery" with 6%, and "sculpture and statues" with 3%. (See table 2 for details.) These data files contain the following coded information: "Artifact Number", "Name", "Original Name", "Era 1", "Specific Era", "Artifact Category", "Specific Size", "Mass Range", "Specific Mass", and "Mass Unit".

The categories of information included are mostly related to the material properties of the collection. "Artifact Category" distinguishes artifacts in terms of the materials used (e.g., metals, ceramics, etc.). And "Specific Size," "Mass Range," "Specific Mass" and "Mass Unit". "Specific Size", "Mass Range", "Specific Mass" and "Mass Unit" all further present information about the characteristics of the collection itself.

"Name" and "Original Name" provide possible changes in the title of the object. "Name" refers to the title as it is currently used, while "Original Name" may not correspond to it. This may be due to a correction of a former misidentification of the artifact, or a change of name in other historical and political contexts.

Both "Era 1" and "Specific Era" refer to the historical period to which the artifacts are attributed. The division of "Era 1" is more rough, and only limits the data to a certain dynasty broadly, such as "Tang Dynasty (618-907)". The "Specific Era" is literally understood as a more precise division on the basis of the "Era".

However, the corresponding information is not sufficient as expected. In the following chart about the collection of calligraphy, and painting collection statistics, 85% of the data in the "Specific Era" is blank, only about 15% to provide specific dynasty or age information. Therefore, it can be said that this information, which should be progressive in the stratification, has not been well recorded.

In addition to the above documents, the page also provides visualized information on some statistics. Including the percentage of the number of cultural relics collections in Xi'an, Baoji, and Xianyang (the three main cities where the collection of the museum derived from) among the total number of cultural relics in the country and in the province; the number of different percentages of statistics by cultural relic category, by age of cultural relics, by time of accession, by custodial level, by nature of unit, and by subject of collection in Shaanxi History Museum. This information was not used as metadata in the cultural relics data files to record the corresponding information of the collection.

The data coding of the "Collection of Excellence" and "Treasures of the Museum" pages shows a clear difference from the metadata in the "Cultural Relics Data" (as shown in the table below). The collection data file does not contain information about the museum where the collection is located, whether it is on display or not, or the location where it was excavated, whereas the collection

information code used for the web search contains such entries. Meanwhile, the "Artifact Number" and "Original Name" in the collection data file are not displayed on the search page.

Table 2. Two sets of Metadata.

Metadata on Web Page	Name	Era	Artifact Level	Place of Excavation	Size	Material	Introduction	Related Exhibition	On/Not on Display	Image
Metadata in "Cultural Relics Data" files	Artifact Number	Name	Original Name	Era 1	Specific Era	Artifact Category	Specific Size	Mass Range	Specific Mass	Mass Unit

Thus, a comparison of metadata in page retrieval and artifact data shows that the two are not completely unified. This lack of harmonization stems from differences in the purpose of use. Scholar Kiruma has investigated and concluded that museum staff may use two coding systems for the convenience of different work management [9]. The web pages, as a more interactive platform for users, provide more metadata information that fits the needs of users, such as whether the artifacts are on display or not, as well as the pavilions where they are located, etc., which facilitates the visit. Collections counted by category, age, and time of entry into the collection, can make it easier for visitors to have a more comprehensive understanding. While "cultural relics data" files record more detailed information such as coded numbers and original names of artifacts and are used for museum management systems.

On the other hand, however, it is also possible that due to the complexity of management or information disclosure, there is no standardized coding between departments. As Tim Sherratt referred to what Mike Jones argues, the separation of museums and archives, the specialization of museum departments, and the complexity of responsibilities and handovers have contributed to the fragmentation and disconnection of collection information [10].

The institutions of the Shaanxi History Museum, the Institute of Data Management, the Display Management, and the Collection Management are all responsible for the preservation and management of collection information to a certain extent, but they belong to three different departments (the departments marked in green in the institutional setup in the figure below).

The Data Management Section is under the Department of Information and Materials, which coordinates the overall information management work. The Display Management Section is under the Department of Display and Exhibition, which is responsible for the storage and maintenance of the collections and the implementation of the exhibition projects provided by the Exhibition Planning Section. The Collection Management Section is under the Collection Management Department, which is responsible for the management and preservation of the existing collections, as well as the inclusion and integration of new collections and their information in cooperation with the Collection Acquisition Section.

The three departments have a different focus on the inclusion and use of collection information in their respective work. Therefore, this information presents different contents. And other departments also inevitably need the information collected by these departments in their own work.

Therefore, just as in the museum web page the way of data encoding of "Collection of Excellence" and "Treasures of the Museum" are not exactly the same as the "cultural relics data" in the collection of data entries presented in its metadata, there are also gaps in the data collection, the work of the handover that may lead to the duplication or missing of information or the inconsistency. These are issues that need exploring and optimizing.

3. Differences and Deficiencies

3.1. *Inadequacy of the categories of basic archiving information*

Through the comparison of the way of coding in *Specification for registration of cultural relics in the collection of cultural institutions* (WWT0017-2013, 1-2), it can be understood the characteristics of the coding of the collection of the Shaanxi History Museum's own collection coding as well as its shortcomings.

The first is the definition of "Name" and "Original Name". Is "Original Name" referring to the name determined by the archaeological excavation with written records or other bases, the name named by archaeologists for the cultural relics after they were unearthed, or the name told by the donor? What is the origin of the "Name" itself? Is it a standardized name given to the object after it has been added to the collection? Are there clear criteria for such a definition? This issue has not yet been addressed in the coding of collection information in the Shaanxi History Museum. On the other hand, the retention of "Name" and "Original Name" can provide information on the variations in the dynamic process of naming collections (e.g., differences in the naming of collections by different groups of people, such as archaeological excavations, collectors, or scholars). This will facilitate subsequent research to understand the context of the collection more vividly.

When corresponding to the coding method of the "Specification for registration of cultural relics in the collection of cultural institutions", it will also be found that the cultural relics data files released by Shaanxi History Museum are not in full accordance with the metadata content of the "Specification for registration of cultural relics in the collection of cultural institutions".

The basic information of the cultural relics in the collection in "Specification for registration of cultural relics in the collection of cultural institutions" should have a general registration number; name (including the original name); age; texture; category; number; size; quality; level; source; intact condition; preservation status; date of entry. The above data are not all publicized in the documents of cultural relics data and the page search of the official website of the museum.

3.2. *Inadequate coding information for digital display and interaction needs*

Besides the basic archival information, as digital technology becomes more widespread, there is a need for metadata that can be adapted to new ways of display and management as part of the encoding of the collection. Charlotte Berry mentions the concept of the "relational museum" in Mike Jones' book *Artefacts, archives, and Documentation in the relational museum, Archives and Records*. In this type of museum, the management of information about the collection does not only depend exclusively on a specific department and leaves the other departments that are outside of the specialized information uninvolved [11].

In the concept of the "relational museum", collections are coded to provide a complex, processual, multifaceted, and integrated description of the historical and cultural context of the objects, rather than relying on the standards of a dysfunctional and purposeless stereotypical documentation system. Such an encoding would then need to focus on more of the museum's interactions with public events in the community, with visitors, scholars, archaeologists, and other diverse populations and their needs, so that the information is more explicitly engaged in the process of cultural dissemination.

Therefore, how to maximize its convenience and efficiently present the exhibition and circulation of collections; how to use the space of display to construct the relationship between audience, history, and works in digital coding; how encoding can collaborate and cooperate with museology and digital management of museums? These are the contents that museums need to consider in the future coding scheme for digital applications.

For example, in the circulation of the collection, to track and feedback on the status of the collection, the following metadata can be used to label the status of the collection, such as "on display", "maintained", "on loan", "in storage" etc...And provide further status details in the sub-tags, such as "on display", "on display location", "on display duration".

Table 3. Possible Arrangement of Status Tags.

Condi- on of the Collecti- on	On Disp- lay	On Displa- y- Exhibi- tion Name	On Displ- ay- Locat- ion	On Displ- ay- Date	Restora- tion	Restorati- on-Date	Restora- tion- Project	Restoration- Institute/Dep- artment	On Lo- an	On Loan- For Resea- rch	On Loan- For Resea- rch- Date	On Loan- For Exhibi- tion
	1/0					1/0				1/0		

The presentation of such information not only facilitates visitors to better arrange their viewing schedules to meet their needs but also helps museums manage their collections and ensure that the work and handover of the collections between different departments is organized.

At the same time, scholars and organizations interested in studying the collection can also learn the location and status of the collection more conveniently through the real-time status of the collection code information, so as to facilitate the appointment for research.

Such metadata integrates and shares information from museum to museum, from museum to academic institution, and from archaeology and restoration department to department. It allows departments that were originally separated from each other to participate in the presentation and production of the collection's information. Thus, a more comprehensive data coding is constructed at the management level, which is applicable to various activities for nowadays' needs.

Digital coding of the collections based on the display space requires not only dimensional information about the collection itself, but also metadata categories that encode real-time information about its display needs. It could be the data that includes the distribution of visitors in the pavilion, the duration of the exhibition; and the indexing of historical and cultural related research, other exhibitions, and different media information related to the collection; also shows the ways of displaying and viewing the exhibits according to the materials, techniques and preservation methods of the exhibits themselves, and so on.

In that case, the characteristics of the exhibits themselves and the needs of visitors to the exhibition are considered in terms of space, visual effect, preservation methods, and other aspects. While fully and specifically understanding the historical and cultural background of the collection, a rich and specific exhibition context can be better designed to improve the experience of the exhibition and the interactive effect of cultural dissemination.

4. Conclusion

In general, the digitization of the collection information of the Shaanxi History Museum has a systematic representation on the page. However, the standardization and concreteness of information on details, such as year and source, need to be completed. The reasons for this problem need to be further discussed. On the one hand, due to the differences between traditional Chinese historiography and modern archival standards, there are certain difficulties in standardizing the dating and naming of specific works. At the same time, the traceability and handover of the works, as well as the complex social background at the time, make it difficult to trace and restore the information, which also hinders the optimization of the collection of information. This is also a possibility for diversifying the information in the collection in the future. Obviously, just labeling the upper and lower limits of the year is not enough. As for the era of collections, it is necessary to consider how to create a set of plans on the periodicity of works that can be rooted in their local characteristics. It's not about forcing a definitive truth. Rather, it is a multi-faceted presentation. There can be more than one truth. The further improvement of metadata coding based on digital exhibition and interaction needs is yet to be implemented by a larger number of staff with more professional and technical skills, more complete

equipment supports, and further deepening of information interaction between working departments, museums, and all walks of life.

References

- [1] Xie B 2008 Multimodal computer-mediated communication and social support among older Chinese internet users. *Journal of Computer-Mediated Communication*, **13** pp.728-750
- [2] Karp I 2006 *Museum frictions: Public cultures/global transformations*. Duke University Press
- [3] Cameron F, Kelly L 2020 *Hot topics, public culture, museums*. Cambridge Scholars Publishing
- [4] Murawska-Muthesius K, Piotrowski P. 2015 *From museum critique to the critical museum*. Farnham: Ashgate
- [5] Rosenstein C 2010 When is a museum a public museum? Considerations from the point of view of public finance. *Int J. Cul. Policy* **16** (4) pp.449-465
- [6] Bollo S, Zhang Y 2017 Policy and impact of public museums in China: Exploring new trends and challenges. *Mus. Int.* **69**(3-4) pp.26-37
- [7] Wang W, Fu M, Hu, Q 2020 The behavioral pattern of Chinese public cultural participation in museums. *Sus.* **12**(7) p.2890.
- [8] China Internet Network Information Center 2023 "The 50th Statistical Report on China's Internet Development", CNNIC, Accessed September 3, 2023 https://www.thepaper.cn/newsDetail_forward_20105580
- [9] Kimura A 2023 The elements of museum digitization procedure: a case study of Lithuanian museums, *J Baltic Studies*, DOI: 10.1080/01629778.2023.2185271
- [10] Sherratt T 2022 Mike Jones explores the connections to be made in museums, *Hist. Australia*, 19:1, 205-206, DOI: 10.1080/14490854.2022.2028571
- [11] Berry C 2022 Artefacts, archives, and documentation in the relational museum, *Arc. Rec.* **43** 212-216, DOI: 10.1080/23257962.2022.2045919