

## Discussions on the Ways of Digital Construction of Ancient Books

Jianfei Hu<sup>1,a</sup>, Fang Cheng<sup>2,b</sup>, Qiang Liu<sup>1,c</sup>

<sup>1</sup>Library, University of Jinan, Jinan, Shandong, 250022, China

<sup>2</sup>School of Political Science and Law, University of Jinan, Jinan, Shandong, 250022, China

<sup>a</sup>lib\_hujf@ujn.edu.cn, <sup>b</sup>lib\_chengf@ujn.edu.cn, <sup>c</sup>lib\_liuq@ujn.edu.cn

**Keywords:** Ancient literature; Protection of ancient books; Digitization; Construction ways

**Abstract:** There are large amount of the Chinese outstanding historical and cultural heritage of ancient literature during five thousand years; due to man-made, natural and many other factors, many ancient books disappeared. Therefore, the digitization of ancient books is an effective measure for the protection and utilization of the reproduction of ancient books. Based on the author's study and practical experience, this paper first analyzed the necessity of digitization of ancient books, and then put forward the ways to build digitization of ancient books.

### 1. Introduction

The digitization of ancient books is the use of modern electronic information technology to process the ancient documents, so that they can be converted into electronic data form through the optical disc and net. The digitization of ancient books is one of the important ways to protect and inherit the ancient books. It can not only reflect the original appearance of the ancient books, but also promote the spread of the traditional historical books and culture. It also effectively solves the contradiction between the protection and use of ancient books. The advantage of digitization is that it is convenient to store, and do not need special field space for preservation; moreover, it is easy to carry; and it uses network storage of cloud disk, so you can also be viewed at any time from thousands of miles. High-resolution digital images enable readers to read clearly even on a large-screen display. The digitization of ancient books is the use of modern information technology to convert the ancient books into electronic document form through optical disk, network so that to be preserved and disseminated. It is the new request to the ancient book collation work, also is the development direction in the future.

### 2. The Necessity of Digitization of Ancient Books

#### 2.1 Inconvenient use caused by scattered distribution of the ancient books.

China is one of the ancient civilized countries in the world, and is also a great country of ancient books resources. According to statistics, there are more than 100,000 kinds of rare books and ordinary ancient books in China. They are distributed in large, middle and small libraries and museums all over the country, including Beijing, Shanghai, Nanjing and so on. For various reasons, there are many ancient books in foreign countries, such as the Library of Congress, Harvard Yenching Library, Far East Library, Columbia University Chinese Library, as well as Japan, Korea, Britain and French libraries. These ancient books are the common wealth of all mankind and should be shared by all mankind. However, the scattered collection of these ancient books makes it inconvenient for researchers to consult ancient.

#### 2.2 Digitization of ancient books is an objective requirement for the preservation and use.

As a paper document, the value of ancient book increases with the development of time. At the same time, with the aggravation of paper aging, the problem about how to preserve and effectively use becomes an urgent problem. The problem about how to effectively protect these valuable ancient books without affecting their normal use has become an urgent problem facing the library of

ancient books collection. For many years, librarians have been trying to solve this problem. Many ways have been used, and the most important are two kinds: First, it is to make bibliographic index, summary and other reference books, so that to guide readers to use the ancient books and to reduce the blind use of ancient books; second, in order to reduce the pressure on the original of ancient books, such as Beijing Library and Shanghai Library; many rare ancient books have been copied. Facts have proved that these methods are effective in protecting the ancient books and alleviating the contradiction between protection and use.

### **3. Ways of Digitization of Ancient Books**

#### **3.1 Data formats are waiting to be unified**

Unifying data format may be a topic of change. However, the unity proposed here by the author is not to unify all ancient books and content into unified text or image format. Moreover, it is also not realistic to text all the ancient books in electronic format under the existing conditions. The so-called unification raised by author is to unify the semantic information of a resource object into a metadata format based on general experience of informatization of foreign academic materials in libraries. As a data, metadata is used to describe data, so it is called "metadata". For example, the most common DC metadata (Dublin Core metadata). The standard of metadata contains 15 basic elements and many extension elements, which are used to describe the semantic information of resource objects. These 15 elements are: Title, creator, date, theme, publisher, type, description, other responsible person, format, source, permission, identifier, language, association and coverage range. Metadata can be extended to define a different set of metadata according to each different object; while ensuring openness, it also adopts a canonical syntax. Thus, the description and classification of information can be formatted, so that to create the possibility for machine processing.

#### **3.2 Resource distribution needs census**

Mr. Chen Shuang working in the Institute of History of the Chinese Academy of Social Sciences, has published two articles, including "Overview of the Online Books and History Database" and "Introduction to the Online Full-text Retrieval System of Ancient Books", which are very popular. All major academic websites are reproduced one after another. In the past five years, there are many articles about the electronic or network resources of ancient books like this, but unfortunately, there is not yet a complete report on the survey of digital resources of ancient books. We can not get a comprehensive understanding of the total amount of work done by scientific research institutions, university libraries or individuals, and the digital coverage and distribution of ancient books. The necessity and importance of this work need not be discussed. Like the index of academic materials, it will play a great role in making better use of and building digital resources of ancient books. However, there are also quite a lot of difficulties in this work, mainly as follows: First, the development units are dispersed, and the mechanism of information exchange is not perfect; thus, there is a mutual conservative mentality. Second, the situation of the distribution of digital resources is complex. Because of the problems of format and utilization rate, many resources of early construction are easily forgotten and gradually lost, and the change of network resources or the natural loss of network resources is a common phenomenon. These problems are widespread both at home and abroad. Technically, we have sufficient support. Through web tracking technology, we can relatively overcome the problem of natural dissipation of network resources. Through simple database technology and metadata definition, the integration and continuous updating of resource positioning can be realized. The key is organizational security and financial support. The author thought that it can just gives play to some of the advantages of the existing system to overcome.

#### **3.3 To establish the digital alliance of ancient books**

The establishment of a digital alliance of ancient books, including scientific research units, libraries, professional development companies and experienced groups and individuals, will not

only help to implement large-scale public welfare projects such as the digital construction census of ancient books, but also play a leading role in industry development, avoid duplication of construction, uniform industry norms simultaneously. As we all know, in the construction of digitization of ancient books, it is a bottleneck that the existing character set is not enough. Whether it is the extended specification of Chinese character inner code or the square large character set, there are too many useless characters invented. Why is this happening? The main reason is that the designer does not know the actual demand of GB2312 with more than 6,000 Chinese characters which can meet the needs of the general literature. However, it is difficult to have a clear understanding of which characters are needed to deal with ancient books without actual experience of digitization of ancient books. The establishment of the industry alliance can focus on experience and work out a relatively stable and reasonable character set standard. In the process of document information processing, each step will involve standard issues, such as the unification of variant characters, the organization and classification of information, keyword determination and other links. Digital ancient books are essentially "public information" products as pointed out by Mr. Li Duo, whose purpose is to enable more people to use them. At this time, we cannot use the simple market thinking of "who is bigger and stronger is the standard". Instead, the public idea of sharing and winning should be followed. Over the years, the call for an industry alliance has been unfulfilled. It has something to do with the lack of this kind of public concept.

### **3.4 To set up a leading group on the digital construction of ancient books**

It is the trend of the times to realize the digitization of ancient books resources, but the ancient books in China are numerous and distributed in libraries and other institutions of different systems and departments. Therefore, the digital construction of ancient books needs the coordination of different departments. The establishment of the leading group for digital construction of ancient books is the organizational guarantee to carry out the digitization construction of ancient books. China's National Ancient Books Sorting and Publishing Planning Group has organized more than 10 large libraries to collaborate in compiling the catalogue of Chinese Ancient Books with the aim of collecting all kinds of ancient books in China. On the basis of compiling the General Catalogue of Chinese Ancient Books, the author planned to build the database of Chinese Ancient Books Bibliography as a useful supplement to the bibliographic data of ancient books of more than 10 libraries. From 2008 to 2010, the "Joint Catalogue of Ancient Chinese Books" was initially formed, actively utilizing the achievements of ancient books protection and serving the public. This act will be a good beginning of the organized and unified planning of the digitization of ancient books, and will help to speed up the digital construction of the ancient books. All the collectors of ancient books in China should take an active part in providing standardized bibliographic data of ancient books to the total bibliographic database of ancient books in China, and on this basis, they establish their own standardized bibliographic database of ancient books.

### **3.5 Talents cultivation for digital construction of ancient books**

At present, there are very few talents engaged in ancient books collation, especially those who have knowledge of ancient books and skilled talents who are skilled in using modern information technology. With the retirement of the older generation of staff, this problem is particularly prominent. Therefore, it is urgent to cultivate a group of new talents who can understand ancient books and use computer technology skillfully in the digital construction of ancient books, which is also the need to carry out the deep service of ancient books in the future.

## **4. Conclusions**

Through the digitization of ancient books, we can construct a large database of ancient books. We can digitize the ancient books and the cultural achievements of the past dynasties, so that to form a broad and profound database, and realize the functions of information query and research of ancient books. The ancient books that handed down for five thousand years and loaded with heavy Chinese civilization, are a huge spiritual legacy left by the great ancestors and condensed the

wisdom of the Chinese nation from generation to generation. It is an indispensable duty for us to inherit and develop Chinese culture, and the digitization of ancient books will have a long way to go.

### **Acknowledgements**

The study was supported by National Book Collection Committee Funded Projects: The Compilation of Criminal Cases Revision (No. SZ14036)

### **References**

- [1] Li Xiaoping, Shen Yang, Zhang Chuanli, Cai Jimei. A Preliminary Study on the Digital Construction of Medical Documents of Ethnic Minorities in Southwest China [J]. Chinese Journal of Library and Information of Traditional Chinese Medicine, 2017, (03): 39-42.
- [2] Xu Jin. A Preliminary Study on the Construction of Digitization of Ancient Books [J]. Expensive Photo Journal, 2014, (04): 39-41.
- [3] Yang Yongjun. Discussion on the Digital Construction of Ancient Books [J]. Journal of Economic Research, 2012, (27): 239-240.
- [4] Huang Weixia. Digitalization of Ancient Books and Construction of Digital Library [J]. Information Science, 2010, (08): 1269-1271+1277.
- [5] Wu Guiying. Practice and Thinking on the Digitalization of Ancient Chinese Medicine Literature [J]. Journal of Medical Informatics, 2010, (04): 54-56.
- [6] Gong Yajun. Some Thoughts on the Construction of Digitization of Ancient Books and Documents [J]. Modern Intelligence, 2008, (12): 95-96.
- [7] Luo Lili. Discussion on the Digitization Construction of Ancient Books in Guizhou [J]. Theory and Contemporary, 2008, (06): 40-41.