

December 2017

The Value and Problems of Digital Preservation for Historical Documents in China

Yongsheng Chen

School of Information Management, Sun Yat-sen University, 13678990606@163.com

Huanning Su

School of Information Management, Sun Yat-sen University, suhuanning@126.com

Please take a moment to share how this work helps you [through this survey](#). Your feedback will be important as we plan further development of our repository.

Follow this and additional works at: <https://ideaexchange.uakron.edu/docam>

Part of the [Archival Science Commons](#)

Recommended Citation

Chen, Yongsheng and Su, Huanning (2017) "The Value and Problems of Digital Preservation for Historical Documents in China," *Proceedings from the Document Academy*: Vol. 4 : Iss. 2 , Article 14.

DOI: <https://doi.org/10.35492/docam/4/2/14>

Available at: <https://ideaexchange.uakron.edu/docam/vol4/iss2/14>

This Conference Proceeding is brought to you for free and open access by University of Akron Press Managed at IdeaExchange@UAkron, the institutional repository of The University of Akron in Akron, Ohio, USA. It has been accepted for inclusion in Proceedings from the Document Academy by an authorized administrator of IdeaExchange@UAkron. For more information, please contact mjon@uakron.edu, uapress@uakron.edu.

Because of China's long history, it is a country that has many historical documents. With the development of Internet technology, China has realized the importance of digital preservation for historical documents and for a long time now has been undertaking a large-scale digitization project. This paper will describe this initiative, also summing up lessons and experiences. This kind of research can not only help us to explore the disciplines of the digital preservation, but it can also provide some references for future research and practice.

Background

Documents record almost every aspect of human life, which can be seen as a very important tool for descendants who research the development of human society. From documents, people are able to review human history, society and science. That is the value of preserving documents.

As our life is becoming more and more technologically diverse, the materials and methods people use for recording life are more various than ever before. That is the reason why it is difficult to identify what the document is nowadays. However, despite this uncertainty, the importance of historical documents can not be understated. According to the theory and practice in Chinese archival work, the document can be defined as all kinds of information records formed by state institutions, social organizations or persons performing their statutory duties or processing their affairs (DA/T1, 2000). According to The Archives Law of the People's Republic of China, historical documents can be open to public after forming 30 years. From this point forward, this paper focuses on all historical documents digitized in Chinese archives, not only derived from the different dynasties in Chinese history to the establishment of the People's Republic of China, but also all the historical documents formed 30 years ago from now.

The significance of historical documents is self-evident., in China, ancient people used oracle bones to record their life. The pottery writings discovered in Shandong Province were seen as the earliest writing, dating back to about 3000 BC. After the formation of ancient china, people realized they need something to record and deliver different kinds of information systematically and logically. And in order to adjust to the development of social production and life, especially for managing the country, all were official documents are preserved systematically. Therefore, a lot of dynastic archives were formed (Feng & Zhang, 2006, p. 24). China has preserved a lot of

historical documents so far. As it is known to all, digitizing these documents and then exploring the information will be beneficial for the historical research not only about China, but also about other countries all over the world which have relationships with China. Besides, making good use of historical documents can be also regarded as an effective way of rebuilding and promoting history and culture.

Nevertheless, as time goes by, more and more historical documents are faced with deterioration. These problems make the preservation and conservation of historical documents become vital in rescuing the valuable historical information. Fortunately, digitization provides an efficient way for document preservation, and this enables us to explore the methods and technology of digitization as well as exploit and use historical information in documents.

In China, digitization is mainly described as the process of transformation of analog image to digital image by using computer technologies such as scanning and Optical Character Recognition, and so on. (DA/T31, 2005). The documents in research question of this paper are the official historical documents, which are mostly stored in Chinese central and local archives. In fact, there are still a large number of other kinds of historical documents, such as those of civil organizations or individuals, which are scattered in many different places. And the information contained in these kinds of historical documents is also a precious cultural treasure. This paper mainly talks about the digitization of official historical documents for two reasons. On one hand, it is more convenient to get access to and collect the research data and materials, because as the executors of digitization, Chinese national archives at all levels store most of the official historical documents; they are open to public, so these archives make study feasible. On the other hand, the state archives administration of the People's Republic of China has issued unified guidance on the digitization of historical documents. This has made digitization an important aspect of archival work. And the top-down approach to large scale digitization ensures more general and a greater significance in China, which may also be able to make the study results more representative.

At present, most of the digitization work of official historical documents is carried out in different centers or local Chinese national archives as well as the First and Second Historical Archives of China, although most of the practical work has be outsourced to different companies. Focusing on their own collections, the objects of digitization in different archives have different collecting emphases. The First Historical Archives of China, for example,

makes efforts to digitize the documents formed in the Ming (1368–1644) and Qing (1644–1912) dynasties. While the digitization in the Second Historical Archives mainly focuses on the documents formed in the period of the Republic of China (1912–1949). Besides, national archives at various levels process the digitization of documents formed after the establishment of People's Republic of China (in 1949) in different areas.

The Value of Digital Preservation for Historical Documents

In recent years, in order to protect and use the documents well, China is currently undergoing large scale digitization of historical documents. Digitization of historical documents is good for both document management and use in China.

As we all know, the usability of the physical information carrier and the usefulness of content information are two important aspects of documents, and these two factors are important premises of fully using documents and realizing documents' value at the same time. That is to say, for one historical document in the original environment, if the carrier is too broken to be used, even though the information in that document is important or significant, we are not able to realize the value of it. On the other hand, considering the carrier is perfectly protected, but if the information is not useful, we will not use it either. So, that is the reason why the usability and usefulness are two indispensable elements for the using of documents.

The digitization of historical documents is not only beneficial for the protection of the original paper carrier, but also for exploring and using the information contained in the documents. On the one hand, people need to do conservation or restoration on damaged documents before digitization—that is the protection of originals, and a way to rescue damaged documents to some extent. On the other hand, even though the originals will become damaged someday in the future, we will still be able to make use of the information we preserved by digitization, which is the protection and continuation of historical culture, and that is the most obvious advantage of digitization.

Digitizing historical documents has four leading advantages. First of all, for protecting the carrier of documents, digital preservation for historical documents can provide useful information to society without giving originals, which can protect the originals by decreasing the using and wear of them. As for the historical documents which are deteriorating, digital preservation can

rescue the originals, and it is also an important method to save the information of documents in the meantime.

Secondly, digital preservation can make the historical documents more convenient and efficient to use. Preserved as archives, there is only one original of the historical documents in general; people who use documents are always limited by the time and space before. For example, if a friend from America needs to get some information of a document formed in Qing dynasty when doing his or her research, and if the document has never been digitized, the friend must take 15 hours to go across more than 10,000 kilometers by plane to The First Historical Archives in Beijing. That will cost a lot of time and money. But if we have made digital preservation of the document, the friend can use it in America through the Internet. So, the transformation (from paper to electronic form) and use of digital information can make people free from the limits mentioned above.

Thirdly, digital preservation for historical documents is the basis of the construction for digital archives and digital libraries. It is considered that providing access is more important than storing for public cultural institutions such as archives and libraries. But in fact, effective storing is the basic for effective using, and also the information resources are the foundation for these institutions. Taking digital archives for example, if there are no resources in archives, how can archives serve the public by providing useful materials? So it is obvious that resources play a fundamental role in archives. Resources and services go hand-in-hand—one cannot exist with the other. Digital resources are made up of two parts. One indicates the original digital or electronic documents; the other is formed by the digitization, which can transform the paper documents into digital ones. There is no doubt that digital archives must have digital resources, which means as one part of the resources in digital archives, the digitization of historical documents is vital and indispensable.

Last but not least, for the information resources exploitation, digital preservation is the premise for content analyzing and digging through information technologies. In the past, archivists or document managers were not able to manage the content of historical documents (Wu, 1991). That was not because that they did not want to manage them, but because they did not have enough capability and energy to make good content management due to the fact that they could not be familiar with all the documents' content. However, once the documents are digitized, the information technologies can analyze the content of documents automatically and then organize the documental information in different themes, on which basis the managers can have a rough

idea of the content and then provide the documents for use which are under their management more accurately. What is more, content management can promote knowledge management of documents, which is the inevitable direction of the document management in the future.

The Problems of Digital Preservation for Historical Documents in China

Due to the advantages in digital preservation, many historical documents are being digitized in many different archives in China. Although the digital preservation is beneficial in many ways, certainly, there exist some problems in the process of document digitization in China, and the details are as follows.

Digitizing Without a Good Plan

Firstly, people who are responsible for digitization in China pay the closest attention to the amount and speed of digitization, and so they might not make correct value judgements about the historical documents before digitization. That is to say, archival digitization projects in China have not been well-planned, meaning that, for instance, documents of greater value have not been processed preferentially.

You may wonder why China is in such a hurry to digitize. That is because of the requirements of the State Archives Administration of the People's Republic of China, which is an executive agency leading and managing all the Chinese national archives. The director of the State Archives Administration Minghua Li read out the "Thirteenth Five-Year plan" Outline of National Archival Career Development for National Archival Work in a conference. There are some basic requirements for the archival resources' construction. One of these requirements is that the proportion of digitized collections in provincial archives must reach 30–60%, while the ratio in prefecture-level archives must turn to 40–75%, and the percentage in county-level archives must get to 25–50% from 2016 to 2020 (Li, 2016). Another request of the archival collections of the State Archives Administration is that the rate of digitization in our country must be quadrupled, from 2010 to 2020. From the amazing expectations expressed above, you can see the quantity is the first and the most important target to evaluate the accomplishment of digitization. In order to accomplish the aim, all the national archives are

digitizing the collections with fast speed and in large numbers, which may cause the scanning quality to be low.

However, digitization without regarding to the relative importance or urgency will not only cause a waste of resources, but also lead to the documents of high value or close to destruction not being processed preferentially. In fact, the funds and resources for digitization are limited, so we need to make a good plan to maximize the use of all kinds of resources. But how to make a better plan before digitization to avoid this problem? According to Zhao (2007), there are two basic principles we should follow. The first one, called Value Principle, is to think about the value of the historical documents, and the other one is the Use Principle, which is related to the using frequency of historical documents.

The Value Principle means the historical documents with precious value should be digitized firstly, on the contrast, the documents which are not so important could be digitized last. This principle is also related to the special collections in archives. Unlike libraries, the collection resources in every archives in China are different, because the collections in archives are mainly from particular administrative areas, which can reflect the politics, the culture and the development of the place. So these collections are always rare or even the only existing copy. It is easy to understand that such documents need to be digitized with priority, or once they were broken or lost, we may not be able to find them again. Digital preservation can help us to protect those precious cultural treasures at least from an information perspective.

The Use Principle means the historical documents which are used frequently should be digitized firstly, while the documents which are not used often can be digitized later. There are two advantages to obeying the principle. Firstly, documents with a high use ratio likely means they are of high use value too. Digital preservation can meet more people's use requirements so that it can improve the documents' use ratio, because it can provide access for different people in different places simultaneously. People using historical documents need not come to the physical archives to search for originals anymore. On the other hand, some damage to originals is inevitable when people use documents. The more use, the more damaged the documents will be. But digital preservation can provide information without necessitating the use of originals by the public, which can protect the historical documents to a large extent.

Neglecting Exploring Information Resources after Digitization

The second problem of taking too much energy in the digital preservation of historical documents is that it will lead people to neglecting other aspects which limit the promoting of using efficiency, such as the description, content management, information exploitation, and so on. The most important purpose of digital preservation is to provide them for social use. However, digital preservation is not the only thing we need to do to promote the document use.

In order to make document utilization more efficient, we still have a lot of things to do, combined with the modern information technology. The digital practices in China place emphasis on the scanning of historical documents. However, digitization is only one aspect of making documents usable; content information management after digitization is more crucial for document use. Nowadays, information science technology helps archives to dig out and analyze the content of historical documents; and as archivists, we need to research or design the methods for providing the accurate retrieval and access to public users, and then accomplish it with information technology to meet more use requirements. For example, after digitization, the content of the documents can be transferred to other editable carriers through Optical Character Recognition techniques, and then archivists can use information mining and analysis or other big data technologies to analyze the public use needs and the document content, and then accomplish the editing and research of documents in different themes automatically. This work will become an important area in archival work.

Another thing archivists neglect when digitizing paper historical documents is the preservation of other forms of archives, such as audiovisual documents. Compared with paper, the information in audiovisual documents is more likely to be lost because the carriers of audiovisual documents are more unstable, and the life expectancy of these magneto-optical carriers is also shorter than paper (Zhi, 2016). So digital preservation of historical documents should not only pay attention to the digitization of paper documents, but also to the protection of the original digital documents, including the safety and long-term preservation of these kinds of historical documents.

Some archives have misunderstood digital preservation, thinking that digitizing historical documents is a means of protecting only the originals. Some archives seal up the originals and only provide access to the digital copy because they think it is enough for users to use content information. So for the sake of extending the life expectancy of the originals, they do not allow people to see

or use the originals anymore once the documents are digitized. Nevertheless, this kind of mindset regarding document management is extreme. The famous librarian Paul N. Banks from the School of Information of University of Texas at Austin put forward Ten Rules for Document Conservation and Preservation. One of these says, “The physical medium of a book or document contains information” (McCrary, 2000). Obviously, as copies of originals, the digitized documents can not contain all the information contained in the originals. At the very least, the originals’ paper quality can reflect the development and history of the papermaking, the culture of paper use and so on. In fact, there are still a lot of people demanding for access to the originals for a variety of reasons. So as information providers, archivists in China managing historical documents should meet the users’ demands as much as possible. They need to solve the problems faced in the process of providing access rather than avoiding them by forbidding access.

Conclusion

According to the latest statistics from the State Archives Administration of the People’s Republic of China, the digitization of collections of all archives in China in total was 22.43 million gigabytes at the end of 2016 (State Archives Administration, 2017). And the digital preservation of historical documents is still going on.

The relationship between digitization and access needs to be addressed. In China, public access to historical documents depends on an “opening appraisal” procedure. If a document does not pass this appraisal, it will not be accessible, even if it is digitized. In China, the opening appraisal is an important procedure in archival work because it is related to the safety of the documents forming institutes, even to the country, and archivists should always be cautious about it. So, although the policy requires the archive should be opened for public access after 30 years of its formation, in practice the opening appraisal lags well behind, which causes the public to not be able to use the historical documents in a timely manner. If we want to improve our access to historical documents, we must overcome these difficulties. Only if the historical documents are accessible can digital preservation realize its potential value and meaning. Once this happens, people all over the world can use them on the Internet after providing some necessary information and submitting an application.

What’s more, due to some limits, this paper only refers to the official historical documents, while there are many other individual or organizational

historical documents. Maybe the digitization of these historical documents is different from the official documents because the management and resources are quite different between the two forms of documents. Further research can be undertaken to learn about the digital preservation of other kinds of documents, so that researchers can summarize more representative and comprehensive disciplines of digital preservation for historical documents.

There are some problems in the large-scale digital preservation of historical documents in China that need to be identified and corrected. However, the value seems to outweigh the problems. Because for the sake of history and culture, the most important thing is to preserve and keep them. Then, users can have chance to do more things with the historical documents.

References

- DA/T1. (2000). General terminology for archives work. Retrieved from <http://www.saac.gov.cn/xxgk/site2/20150604/00e04ce0897016da317b01.pdf>
- DA/T31. (2005). Digital technical specification for paper files. Retrieved from http://www.hada.gov.cn/html/News/303_67687.html
- Feng, H. (冯惠玲), & Zhang, J. (张辑哲). (2006). 档案学概论 [Introduction to archives]. Beijing: Chinese People's Publishing House (中国人民出版社).
- Wu, B. (吴宝康).(1991). 中国档案学研究在发展中 [The Archival Science in China is in developing] . 档案学通讯 [Archives Science Bulletin] ,1991(6):3-7.
- Li, M. (李明华). (2016). 中国的数字档案资源建设 [The construction of digital archive resources in China]. 中国档案 [*China Archives*], 2016(10), 14–15.
- McCrary, E. (2000). In memoriam: Paul N. Banks, 1934–2000. Abbey Newsletter, 24(1). Retrieved from <http://cool.conservation-us.org/byorg/abbey/an/an24/an24-1/an24-102.html>
- State Archives Administration of the People's Republic of China. (2017, October 17). 2016 年度全国档案行政管理部门和档案馆基本情况摘要 [Summary of the basic information of the national archives administration and archives in 2016]. Retrieved from http://www.saac.gov.cn/xxgk/2017-10/17/content_208010.htm

- Zhao, J. (赵晶莹). (2007). 因地制宜对馆藏重点档案优先数字化 [Key documents of the archives are preferably digitized according to local conditions]. 黑龙江档案 [*Heilongjiang Archives*], 2007(4), 19–21.
- Zhi, L. (智璐璐). (2016). 对当前档案信息数字化的反思 [Reflections on Digitization of Current File Information]. 档案与建设 [*Archives and Construction*], 2016 (8): 66, 86–87.