DIGITAL ARCHIVES AND PRESERVATION IN CLOUD FOR THE INTANGIBLE CULTURAL HERITAGE

Chiou-Shu J. Hwang
Department of Culture Creativity, Meiho University,
23, Pingguang Rd., Neipu, Pingtung, Taiwan, R.O.C.
x00002163@meiho.edu.tw

ABSTRACT

Tsz Fung Palace was built in AD 1737 with the history of 277 years, which honors the Chinese Virgin in heaven – Mazu. In order to preserve the local culture assets (tangible and intangible) of Tsz Fung Palace, and to promote the cultural heritage abroad, this project collected the seniors' allusions, videos, historical documents and other information with the temples of art in Tsz Fung Palace, and digitized them into a collection in 9 months, which is similar to a library collection system. In order to design a semi-library system, the project took advantage of librarians' capacity on digitalizing and relevant experience on database management, and set up a rule of rough classification for the current information and documents that may arise in the future. The semi-library system with a classification and catalog system concept, however, it is unable to produce the call-numbers due to the collection of Tsz Fung Palace is not going to check-out for the public. The project team has conducted three steps to build the system. First of all, for each type of data, the team applied information technology to digitize all materials and to preserve all the materials into a database. Second, the team designed the website. Third, the team designed a software program to combine website and database together. Finally, the project team let the collection display on the web page. However, part of the meeting records and budget or funding details, just as the records of purchasing model of library system, will not appear, it is not necessary to be shown for the public. Two-thirds of the Palace almost has been burned on Dec. 3rd. in 1983, which was caused by electrical fire; furthermore, not only the building, but also the documents not much left. Therefore, the digital archives completed not only kept in the server of the Tsz Fung Palace, but also storage in cloud. As a result, the digital archives and collection, which were established in this project, may be served as long as possible. At present, this project has the following benefits: 1. It is helpful for academics research or people's understanding and application, after digitizing the relevant documents of Tsz Fung Palace. 2. After building the database and website, public are easy to consult services of Tsz Fung Palace, and easy to retrieve the relevant information from the back-end of database. 3. Not only academia and related cultural industries can take advantage of such information and material, and to improve the quality of research and production; but also to strengthen the cultural communication between domestic and foreign religion organizations. There are following suggestions for future digitizing archives process: 1. It is better to provide manuals to facilitate the Tsz Fung Palace staffs for conducting data collection and subsequent preservation, because the staffs lack the concept of information and technology. 2. Before archiving in the future, the staffs better to file or aggregate the rough materials by themselves, because they have a better understanding for the historical documents, which sometimes with stories beyond words and films.

Keyword: Digital Archives, Library, Cloud, Preservation, Intangible Cultural Heritage, OAIS

BACKGROUND

Due to the information technology in general and the popularity of Internet, people has not only changed their behavior of communication, but also changed their basic knowledge representation, manner of presentation, and even changed their attitudes and ways of learning and research. The impact of information technology on society and cultural heritage is comprehensive, and has gradually changed people's basic attitudes, perceptions and the way of doing things, obviously, a new comprehensive socio-cultural heritage and digital age is coming.

Besides, decades ago, a fire destroyed Tsz Fung Palace, connected buildings and archives, whether tangible temple buildings, temple's richly ornamented, photos of invisible ritual activities are all gone and nevermore see. This big lost caused the lamentation of local grand old men: how did we not expect to retain it at that time or earlier?

Until nowadays, the temple directors and local seniors have known that information technology is helpful to save of the social and cultural heritage, and is a good tool to attract young people while marketing the religious culture. To sum up, it is a not a good tool but also a trend for using information technology to propose local young people to take the relay baton of cultural heritage.

PURPOSES

It may enhance effectively the accumulation, inheritance and use of the knowledge and experience through the digital archives; therefore, speed up the pace for doing Tsz Fung Palace Digital Archives is a meaningful work.

Benefits arising from this work, including at least the three followings:

- 1. It is in favor of the creation of Tsz Fung Palace and preserving local culture in assets (tangible and intangible) and the new culture.
- 2. Rich in various types of materials can help to promote cultural heritage and marketing at local and abroad.

3. The digital archives, which included seniors' experiences, videos, historical documents and temples of art, are established through information technology can be saved as long as possible.

LITERATURE REVIEW

The project team members are requested to have digitized infrastructure cognitive and practical experience, which included electronic database build experience, understanding the religious believers habits of using information, the Internet penetration, academic research experience, and the service capabilities and marketing ability in industry application, and so on.

Taiwan's current application of information technology is among the world's top spear, therefore, the above capabilities and experiences are mature and enough to support this project. In additionally, there are two fields' documents and literature, which are digital archives and cloud preservation, are applied in this project.

The Development of Digital Archives in Taiwan

The first digital archive project started from the American memory Pilot Program in 1990, which was a four-year project and promoted by the Library of Congress. The Congress Library owned rich historical documents, photographs and audio-visual materials, and digitized these material into hundreds CDs, which were distributed to 44 selected schools and libraries (Tsai, Huang Chiu, 2007).

There are about three phases for the relative development of digital archives projects in Taiwan:

Phase ONE-from the "Digital Museum Project" in Aug. 1998, which was set up and executed by the National Science Council (Hereafter referred as NSC). The original purpose was to promote the knowledge of science and culture via multimedia and internet, and anticipated to end this project by 2001; however, it was last in 2002 by change the project title as "National Archives Digitization Plan", which mean to combine the most important and precious cultural relic from the following museums: National Palace Museum, National Library, National History Museum, National Museum of History, Taiwan Historica, Museum of Natural Sciences, National Taiwan University and Academia Sinica. The modified project has be oriented as a preservation of cultural assets, and the construction of public information system through establishing a national level digital archives database, and consequently, the goal became to exquisite popular culture, integrate the information technology and humanities, and to promote the development of industrial and economic (Yao, 2007). Phase TWO-the NSC enhanced the "National Archives Digitization Plan" became to the "National Level Digital Archives & Technological Program", which was

separated into two steps to execute, from 2002 to 2008. In the meantime, the

"National Level Digital Learning Plan" was switch on and executed from 2003 to 2007. From 2002, in order to execute the "National Level Digital Archives Program", the NSC opened four kinds of plans for selection: content development, technology research and development, creative bonus, creative learning. In the same period of time, Executive Yuan proposed the White Paper "Challenge 2008: Focuses of National Development Plan", among this Challenge 2008 plan, the "TWO Trillion Binary Plan" was gotten most attention. The TWO trillion means the industries of biotechnology and digital content. The digital content industry with the development of the knowledge economy served as several indicators of economic significance, and is regarded as powerful tool for upgrading the traditional industries transited into high value-added industries. In this phase, in order to execute the "National Level Digital Archives Program", the NSC also conducted a comprehensive digitized collection of work for all types of knowledge resources, which included important domestic academic institutions, museums, public and private collections of cultural institutions.

<u>Phase THREE</u>-it started the combination of the above several national projects into the "The National Science and Technology Project of Digital Archives and Digital Learning", from 2008 to 2012. Until nowadays, the digital related projects in Taiwan have been developed for around 15 years, and have modified and enhanced through reflecting to policy and environment of Taiwan. The following chart may display the development clearly.

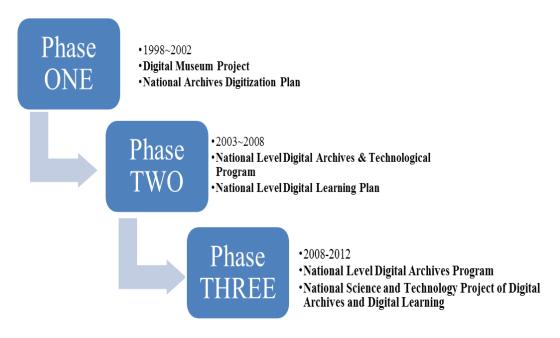


Figure 1: Historical Development of Digital Archives In Taiwan (by Project Team)

Adopted from: Ma & Lin, 2013

The Content of Digital Cultural Archives

In order to execute this project for Tsz Fung Palace, first of all, thinking about the importance of knowledge economy and creative industries, the project team classified and analyzed the theses and dissertations in Taiwan by referring the researches of Ma and Lin (2013), and other related literature. The project team found that the above information may be categorized into the following five kinds of theme.

- 1. Strategy Research: Discussing the operation, management and development strategies of digital archives project, and other related discussions.
- 2. Technology Application: Discussing the hardware practice, software planning, and design and production in digital archives, and other related discussions.
- 3. User Research: Discussing users 'operations and experiences of hardware and software, needs and satisfaction while using digital archives, and other related discussions.
- 4. Content Resolving: Discussing the material, content, and content application of digital archives and other related discussions.
- 5. Digital Teaching: Discussing all kinds of teaching and learning by applying the results of digital archives, and other related discussions.

The most frequent appearance of discussions is "Technology Application" of the above five categories, which implies that the technology-oriented implementation is the main trend in the related education of digital archives in Taiwan. The second frequent appearance of discussions is "Strategy Research", which concerns the commercial values, management, licensing, and marketing of the digital archives.

Secondly, from the content and framework of theses and dissertations, there are several categories classified.

- 1. System: Discussions on digital archives of planning, implementation, model, and strategies facets; there are the largest numbers of theses and dissertation in this category.
- 2. Content: Exploring digital content of international digital archives projects.
- 3. Technology: Discussing relative issues about the hardware and software production methods of digital archives.
- 4. Learning: Discussing the ways and Feasibility for applying the content of digital archives in learning or teaching; there are the minimum numbers of theses and dissertation in this category.

Additionally, there are few theses and dissertations concerned about the user-oriented issues, which may be separated into three types-user research, content parsing and digital education, although the amount paper numbers of these three types are relatively few. From those theses and dissertation, it is obviously that most scholars and researchers in Taiwan mostly focused on the construction of digital content,

hardware and software technology, and strategy use of digital archives, not on the contents of users' behavior, demand, and interaction of digital archives.

In short, the government of Taiwan has completed a preliminary organizational change. For example, in the past years, the affairs of cultural and creative industries were assigned to several relevant ministries under the Executive Yuan. Now, these affairs are managed systematically by the newly established Ministry of Culture. It is a good sing for the future development of digital archives, which means the results of digital archives may be applied with a richer and more in-depth continuously in the cultural and creative industry. For this project, the team keeps the thought of, in steading of a shelved material collection, through the digital archives of cultural assets to enhance the study, research, and as well as in entertainment, business value-added applications.

Cloud Preservation

Thinking about the long-term preservation, maintain hardware, and expanded capability yearly of content, the service of cloud preservation seems offers a solution of preparing storage devices and other issues for digital archives, because these issues can be resolved by the cloud storage providers.

The Choose of Cloud Storages

From the view of long-term preservation and access for digital archives, the main concept of cloud preservation means that the saved resources on the internet in general as the cloud. When the users upload or download the digital information through the internet (the cloud), there is no necessary for them to know where a hard disk is located, and how resources are allocated. The users enjoy cloud storage service which is shared by the cloud providers through the internet, for the users, there is also no need to invest a lot of infrastructure and the cost of hardware storage devices may be reduced. When these cloud infrastructure services (Infrastructure as a Service, IaaS) start to appear, the adoption of data storage and duplication through cloud technology has become an option (Michele & Sandy, 2010). Therefore, through relevant and proper procedures of choosing cloud preservation, the problems in the service, technology and people's lack of respect can be solved while designing the digital archives.

From the view of digital archives to see the cloud preservation service, the institute of digital archives may be viewed as an end, which means the end-user of cloud services. The end-users do not need to understand the infrastructure about cloud, to handle the corresponding professional skills and knowledge, and to build and maintain equipment directly, but need to pay based on the usage. The concept may be presented by the following table.

Table 1: Characteristics of Elements in Digital Could Preservation (by Project Team)

Items	Explanations
Cloud	A virtual service provider, which offers online storage, centralized
	management, common share, with functions of virtual storage, remote
	backup, reliable and safe, energy-saving and carbon reduction, and so
	on.
End	An organization or unit which retrieves the saved content from the
	cloud
Preservation	A process of uploading local or host content to the cloud for upgrading
	service quality

The Futures of Cloud Preservation

In general, there are several attributes of cloud preservation (Connor, Corrigan & Bagley , 2011):

- Information Pool and Multi Users: several users can use the same storage
 infrastructure together. The cloud services may be assigned the cloud
 information pool according the needs of clients. In most of situations, the
 clients do not need to understand or control the real location of specific
 resources. Even in the private cloud for some specific institute or
 organization, the multi units or departments may share the infrastructure
 and other applications.
- 2. Expandable and Flexible: the cloud (virtual) storage capacity can be expanded immediately for unlimited purposes and needs.
- 3. Through the connection protocol: the cloud preservation needs be connected through standard protocol, which includes HTTP, FTP, XML, SOAP and REST.
- 4. Reduction facility: it may reduce the cost for client in Hardware construction costs (CAPEX) and maintenance costs (OPEX).
- 5. Pay online and on demand: Pricing based on the classification of users demands. Usually based on the rate per GB of upload and download or monthly how much storage space is to use.
- 6. Sharing and Collaboration: the cloud content is allowed to be access from multiple locations or multiple users via internet.
- 7. On-demand self-service: the clients may self-manage the ssigned storage services through management console.

The Types of Cloud

There are three levels in the typical cloud construction: infrastructure, platform and application (Chen, 2010). According to the three levels of construction, and consequently there are three kinds of service-public cloud, private cloud, and mix cloud. The clients may choose different cloud according different needs.

- 1. Public Cloud: The type of service means a cloud which is managed by one individual supplier or multiple suppliers, and under this cloud, all the users may share all the resources of the individual supplier or sometimes, multiple suppliers (Peter & Timothy, 2011).
- 2. Private Cloud: This cloud is constructed by business organizations, and only offers preservation service for inner users or clients. The business organizations own the decision-making power for entire preservation environment, and may improve service according to their needs (Goh, Wu, & Sun, 2010).
- 3. Mix Cloud: This kind of cloud combines and offers the functions of public cloud and private cloud. In this cloud, the clients usually upload the noncritical information to the public cloud, but the mix cloud supplier has to manage the noncritical resources based on consumer requirements.

 (Retrieved from http://opencloud manifesto.org/Cloud_Computing_Use_Cases_Whitepaper-4_0-ChinaT.Chin ese_translation.pdf.)

FRAMEWORK AND EXECUTION

The framework of this project is designed by basing on the concept of OAIS, the researcher's LIS (Library and Information Science) professional background, and the above relative literature.

Based on OAIS (Open Archival Information System)

The researcher for designing and executing this project is a LIS expertise, and fully understands there are many large collection agencies have referred the OAIS model and conducted the digital collection, such as the Florida Center for Library Automation (USA), Networked European Deposit Library (Europe), PANDORA project (Australia), NDL project (Japan), and so on. These are important indicating projects which stated clearly the use of OAIS model while digital archiving.

In Taiwan, there are also several large libraries have referred to OAIS and Designed their own collection systems (hereafter referred as IRS-Institutional Repository System), such as National type E-Repository IRS, Taiwan University IRS, Chiao-Tong University IRS, Tsinghua University IRS, Tamkang University Baseball IRS and Chinese Culture University IRS, and so on.

OAIS is a conceptual model of collection system, and may be applied into any type of information, especially for Long-term preservation and suitable for information access and retrieval. The OAIS model can describe the informative construction of institutional environment, functions and workflow. The institutions or organizations which applied OAIS may be match the international standards for using common reference mode and sharing the conceptual structure and definitions, and of course, more easier to share resources and experiences with other institutions and organizations Tsai, Su, & Chiu, 2012).

OAIS provides a functional model, which includes information about the kinds of activities undertaken by each function. The seven functions outlined by OAIS are: Ingest, Archival Storage, Data Management, Administration, Access, Preservation Planning, and Common Services.

A useful diagram of the seven OAIS functions and the transition from SIP to AIP to DIP is given below (Retrieved from http://www.paradigm.ac.uk/index.html, the Introduction to OAIS).

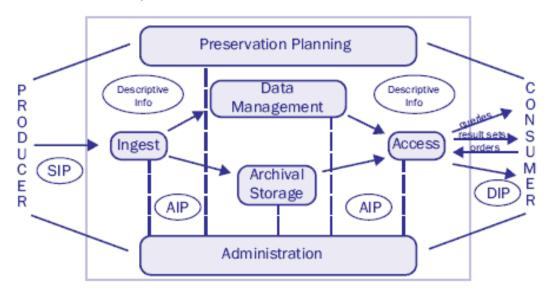


Figure 2: Seven OAIS Functions

http://www.paradigm.ac.uk/workbook/introduction/oais-functional.html

A Concept of Executive Framework

The purpose of this project is going to establish a digital data collection and a maintenance center for Tsz Fung Palace, and to help the internal workers to gather and preserve the relevant collection, such as Documents, photos, videos, audio, picture, interviewing record with seniors, historical files, and meeting minutes. All of the collection is going to save and upload cloudily in data management module,

which is designed as well by the project team according to the concept of OAIS. Following is the framework.



Figure 3: The Concept of Executive Framework

Designed by the Project Team

Additionally, this framework is designed to execute for classifying information, subheading storage and information sharing with external services and expansion goals, through adopting the integrative mode of data and service on both local host server and cloud service.

Execution Stages and Periods

In order to complete a series of systematic data collection and information digitization and management, it is necessary to transfer all archives into a same digital file, such as the old VHS tapes need to be converted into an electronic file format. After finishing the digitalizing job, in the sake of good future, the documents, photos, videos, audio, picture, interviewing record with seniors, historical files, and meeting minutes are shelved in a suitable range which must be accepted and understood by the workers of Tsz Fung Palace. There are four steps of executing this project:

First Step:

Sorting out and listing systematically all of the archives

Second Step:

Purchasing the relative hardware devices

Third Step:

Conversing the documents and photographs scanned files, audio files into digitized files.

Fourth Step:

Establishing a digital collections system and database, which may save the aggregated data and upload into both local server and cloud.

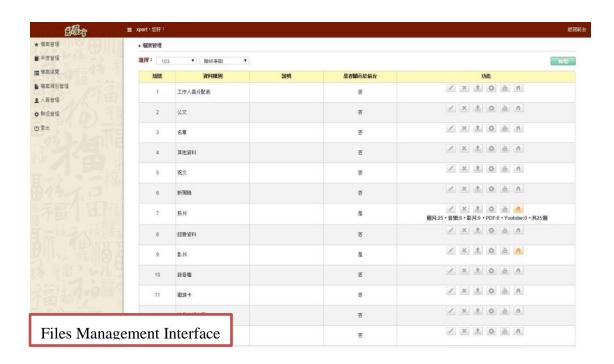
Additionally, this database of digital archives is not only designed to execute for classifying information, subheading storage and information sharing with external services and expansion goals, but also through adopting both local host server and cloud service to set up the integrative mode for both data preserving and service. It is worth mentioning, For the purpose of easy reading, all the scanned material will be transfer into e-book through software, which is the Characteristic of this project.

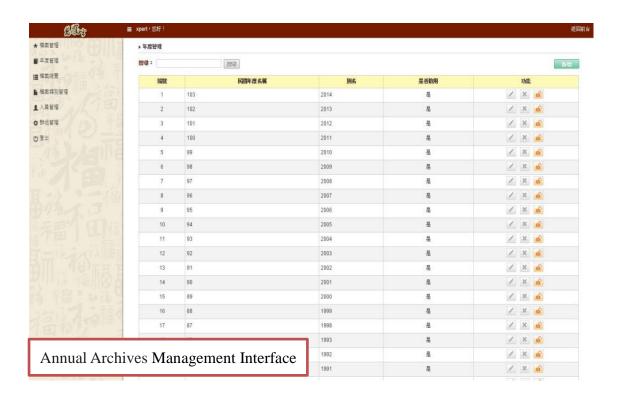
RESULTS AND CONCLUSION

Through the four steps by following the framework, the digital archives preservation has been established. In the future, the workers in Tsz Fung Palace may follow the operation manual for saving almost all of the files, materials, and information, both in local host server and cloud. The project team also designed a website with a database for them to show these archives on line for sharing to the religious people around the world. However, since not all of the archives are necessary to be shared on the website, there is a mechanism on the database system for the internal workers to choose open on public or save in private only.

In this project, the anticipated goal was achieved which is adopting technology to digitalize, preserve, and retrieve archives in cloud. Moreover, the digital archives may be displayed on line, which satisfied the workers of Tsz Fung Palace for the purpose of preaching. The following figures are parts of the results, which are the website on-line screen, the database backend screen, and the e-book screen.

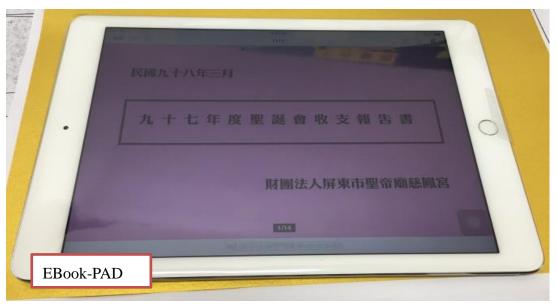












SUGGESTIONS

After executing the entire project and observing some current situations in this temple, the team has several suggestions.

- 1. The manual is necessary and is customized only for internal workers of Tsz Fung Palace. Not only due to their inadequate concept of information technology, but also concern their working environment. This manual may help them to carry out the follow-up data collection and preservation.
- 2. It is a very important pre-work for internal workers of Tsz Fung Palace to compile the documents and variety materials. Because the project team has no idea for the historical activities and relative picture of Tsz Fung Palace, and cannot help them to classify the documents. However, classify of document or archive is the most important step for establishing the digital archives.

Since digital preservation is the current trend for archiving, it is better for adopting technology to save historical material, no matter Tangible or intangible. In general, the digital archives are to strengthen the domestic units to communicate with each other and share the collections. In additionally, the preserved archives may benefit to academy and industry to improve the quality of research and production, even better, it may stimulate the energy of creativity. To sum up, the digital archives are a worthwhile work that may benefit the future descendants and entire human.

References

- Cai, Y., Huang, K. L., and Qiu, Z., 2007. Introduction to Digital Collections Technology. Taipei: Taiwan University Publishing Center.
- Cai, S., Su, W., and Qiu, Z., 2012. Digital Collections Combination of Cloud Storage to Explore. *Science and Technology Museum*, *16* (4), 81~103.
- Chen, Y. 2010. Cloud strategy. Taipei: World Magazine.
- Cloud Computing Use Cases 2010 Discussion Group. *Cloud Computing Use Cases White Paper, 4th. Ed.* Retrieved from: http://opencloud manifesto.org/Cloud_Computing_Use_Cases_Whitepaper-4_0-ChinaT.Chinese_t ranslation.pdf.
- Connor, D., Corrigan, P. H.& Bagley, J. E. 2011, *Cloud storage: Adoption, practice and deployment. Storage Strategies NOW.*, Retrieved from:

 http://www.cleversafe.com/images/pdf/ Cloud-Storage-Outlook-Report.pdf

- Executive Yuan, 2002. Challenge 2008: National Development Plan focused on the White Paper. Retrieved from: www.ndc.gov.tw/m1. aspx?sNo=0001539&ex=1&ic=0000015#.VO7OunyUd7r
- Goh, S., Wu, Z., and Sun, S. 2010. Exploring the Cloud Storage in the Long-Term Preservation of Digital Resources, *Library and Information Technology*, 6, 1-6.
- Ma, R. P. and Lin, R. 2013. Theses and Dissertations in the Department of Taiwan's cultural and creative industries related to the Digital Archives themed content research papers, *Design Journal*, 18(2), 1-24.
- Michele, K. & Sandy, P. 2010, Using cloud infrastructure as part of a digital preservation, Educause Quarterly Magazine, 33(2). Retrieved from: http://www.educause.edu/ero/article/using-cloud-infrastructure-part-digital preservation-strategy-duracloud.
- Peter, P. & Timothy, G. 2011, *The NIST definition of cloud computing (draft)*.

 Recommendations of the National Institute of Standards and Technology,
 Retrieved from:
 - http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf
- Yao, L. 2007. Digital Archives National Science and Technology Program- a Review of Total Result. *Digital Archives National Science and Technology Program Newsletter*, 6(4).