
Digital Archiving : Indian Scenario

Y V Ramana

Abstract

This paper deals with Digital Archiving and presents the Indian scenario of Digital Archiving.

Keywords : Digital Archiving, Digitization, Muktabodha, Jalva, Sampradaya.

0. Introduction

Before the advent of digital video storage analog recordings on films and video tapes were placed on the shelves with a card catalogue. Whenever the library has to make room for the inclusion of new material, some of the old materials were removed and stored off-site, which is properly viewed as an archive. These analog and physical archives had limited shelf life depending on the medium on which the content is stored. In the digital age archiving is in the digital form - that is migrating the content electronically to new media for the purpose of preservation.

The classic meaning of archiving is :

- ✍ To place the documents in a secure place for presentation.
- ✍ To remove the data or documents from common use and conserve for future use.

The word *Archive* derives from the Greek word *Arkheia* meaning "Public records". The word denotes local, regional and national government collections of documents and other records kept for consultation and research purposes. One of the earliest digital -archives was a use-net news group called Talk-origins, which was devoted to mainstream theories in geology, biology, cosmology, theology etc. Contributions arguing the Darwinian evolutionary line were posed and often sent to bear all the hallmarks of serious research arguments and bibliographic referencing.

1. Digital Archiving

Digitization is sometimes recommended as a preservation strategy for materials that are in fragile formats and which are not in frequent use. Images and audio-visual materials are being stored in digital formats more than ever before, for example, digitizing pages of delicate manuscripts, preserving images of texts which are important for scholars. The primary goal of digital archiving is the storage, retrieval and access of rare documents with the aid of electronic technology for image capturing and reproducing. The aim is to create conditions, whereby the researcher can consult the manuscripts without physically handling the individual items and at the same time the artistic quality of the calligraphy and decoration of the manuscripts could be preserved for future generation. In each and every country of the world the *National Archives department* is responsible for the preservation of important documents of that particular nation.

Digital Archives can be defined as Repositories of digital information, that are collectively responsible for ensuring through the exercise of various migration strategies, that the integrity and long term accessibility of the nation's social, economic, cultural and intellectual heritage is preserved in digital form.

2. Advantages of Digitization

The most promising strategy for saving the intellectual and artistic content of endangered volumes is to transform or convert it to a different medium by copying it to a film or digital form. Digital archiving is a recent work of digital libraries and ensures long term storage and access. By digitizing the documents, lots of space can be saved- a stack of CD's will take much lesser space compared to the space occupied if the same information was to be stored as hard copies. In addition, Electronic - files can be accessed from any where, at any time and at a cheaper price.

3. Migration of Digital Information

Refreshing digital information by copying it from medium to medium and the possibility of maintaining a complex set of emulators describe two distinct points on a continuum of approaches to preserving digital information. It is a periodic transfer of digital material from one hardware & software configuration to another or from one generation of computer technology to a subsequent generation.(preserving digital information:)

4. Creation of Information in Digital Form

Information objects in digital form have to be created, edited described and indexed , annotated, revised, re-created, modified , disseminated and retained for future use, by a complex interwoven community of creators and other owners, value-added services and institutional and individual users. The process of preserving digital information will vary with different kinds of objects like textual , numeric, image, video, sound, multimedia simulation etc.

5. Organization of Data

Before scanning the documents for digitization, one has to organize the data, that is to be archived, so that the process will go on in a smooth way and the scanned data can be retrieved properly. The most important steps which are to be followed are given below:

- ✍ Organize the data that is to be archived in to files.
- ✍ Decide how to organize the scanned file, for example- by subject, by code number, by chronological order, by keyword, by author, by title etc.
- ✍ File names should be clearly identified, so that one can easily recognize the contents of the file.
- ✍ All the files or documents which are to be scanned should be kept in order.
- ✍ All the staples and paper pins, if present, should be removed.
- ✍ The paper folds and creases in the paper should be smoothened out.

Scanners :

Different scanners are available for scanning the documents but one has to select the scanner on the basis of materials, like

- ✍ Vacuum-fed belt scanners – can be used for fragile and odd shaped material.
- ✍ Plane-protected wide-angle scanners – can be used for book scanning.
- ✍ High-volume scanners can be used for – architectural drawings and other oversize materials.
- ✍ Flat-bed scanners can be used for – delicate and fragile materials.

6. Method of Scanning

After deciding the type of scanner and type of material to be used, one has to decide whether to scan the documents as images or with OCR (Optical Character Recognition). If the document is scanned with OCR, the files need to be saved in the form of application for easy editing. The benefit of OCR is that the documents are converted to an editable format. If the document is to be archived to maintain the integrity of the records for preserving, they should be scanned as images.

The documents should be scanned at a minimum resolution of 150 ppi, through 300 ppi, for future use or print. The choices available for saving the documents as images are :-

- ✍ BMP- Bit Map is the native image format for the windows operating system and will provide high quality images.
- ✍ TIFF- Tagged Image File Format is very detailed and it is supported by many operating systems.
- ✍ JPEG- Joint Photographic Experts Group format comprises the data, but it is enough for normal use.

7. Storage Method

When all the documents are scanned, they should be saved on CD zip disks or floppy disks or tape drives. For storing the files on the disks, labeled folders should be created, the disks as well as the disk cases should be labeled.

In digital archives there may be many levels of storage on the basis of expected use for the retrieval. Seldom-used material may be stored offline, usually in tape media, which is less expensive. The data which is in high demand and where retrieval time is at a premium, online storage in magnetic media may serve best. An intermediate solution is near-line storage, where the data can be stored on optical or tape media and loaded in jute box, which is more responsive to users demand than offline storage.

The most promising strategy for saving the intellectual and artistic content of endangered volumes is to transform or convert it to a different medium by copying it to a film or digital form. Paper documents can be transferred to microfilms, DVD's (Digital Video Data) and CD-ROM's (Compact Disk Read Only Memory). This process is known as Archiving. Digital archives can use any of the above three methods.

8. Retrieving Digital Archives

Digital archives must maintain the information in a form, so that the users can access the information over the net through search engines, print, audio etc. If the archives are saved on CD's, floppies or tape mediums they should be checked periodically for content redundancy. If the digital archives need to be accessed in a distributed network environment, the digitized data should be connected to networks using proper protocols and bandwidth suitable for retrieving the information.

9. Archiving in INDIA

Preservation of Indian culture and literature is one of the oldest traditions of India. Thousands of years ago, Indian philosophers, scholars and poets have expounded their knowledge through treatises, sacred songs, commentaries on scriptures etc. which were recorded on specially treated palm leaves and birch barks, in the form of engravings on brass and copper plates etc. for future use. Before the invention of the printing press, the texts were handwritten by the craftsmen so that the universal knowledge can be preserved for future use. This proves our theory that Archiving has existed since long in India.

a. Indian National Archives

Indian National Archives department, which is located in Delhi is making efforts through various conservation methods and photo duplication processes to preserve the important documents of Indian cultural art, political and economical and literature. Department of Indian National Archives is using a unique method of rejuvenating and repairing the old documents with the help of "Cellulose acetate foil and tissue paper", which is known as "Solvent of Hand lamination".

The Conservation Research Laboratory is a department of Indian National Archives has been successful in retrieving palm leaves that have gone dry since the last three decades. The National Archives is trying to ensure the longevity of documents by microfilming and reprographing.

b. Preservation of Indian Art

C-DAC – Center For Development of Advanced Computing is trying to digitally preserve and disseminate Indian Art in the form of paintings. By using HP (Hewlett Packard) equipment to capture, store and make available via internet to the researchers and art lovers. In this project, using comprehensive suite of H.P. Technologies, the scientists capture images in extraordinary details using digital cameras to store and process the images on workstations and to develop new colour pallets and drivers to make the individual works and print realistic reproductions on large prints. By digitizing Indian art in this way not only the art is preserved but is also made available to a larger audience.

c. Muktabodha Institute - Archiving of Indian Manuscripts

Muktabodha institute is working in collaboration with libraries and institutions around India to make digital images of specific, rare manuscripts and texts. The institute recognizes the importance of saving the spiritual knowledge and preserved around 871 works in microfilm format. Around 262 of these works were converted to digital image format. The institute is launching the "On-line digital library" website with 74 volumes of the Kashmir series of texts and studies. These volumes were originally published in Sanskrit language. The main contribution of this institute is in the preservation and accessibility of photographic images of rare manuscripts and religious scriptures, so that the valuable information is more easily available to scholars around the world. At present Muktabodha discontinued the practice of microfilming the manuscripts and started producing digital images directly from the manuscripts. Since the main objective of the institute is in the dissemination of old religious literature by digitizing and preserving them, the institute decided not to face the hassels of microfilming and adopted the process of digital imagery.

10. Commercial Archiving Organizations

Many commercial organizations have come up and made it easy to have any time anywhere, to access to information. These organizations provide electronic publicizing, content management, digital archiving etc. for the E-Publishers, individual purposes and institutions like libraries etc. Some of the organizations of India are the following:

a. DPS (Digital Publishing Solutions)

The Digital Publishing Solutions (DPS) is a Pune based organization, which has digitized so far 3.5 million pages and made digital archiving a reality. It scans the data from paper, photographs, negatives, microfilm, microfiche, transparencies and then the data digitized using the OCR and stored the digitized data in E-book format. Now the data is sold online using DX Reader, online XML reader. This organization is a approaching the libraries interested in preserving the rare books, manuscripts etc. In India DPS is working for 15 libraries, National research Institute, Indian national science academy etc.

b. JALVA

Mumbai based Jalva Media provides solutions for media management, indexing , content archiving data distribution , with partnership of other companies like mediasite, Ibeam Broadcasting, Madge.web, microsoft, Real and Apple. Jalva and SGI are offering better technologies like MPEG2(moving picture experts group 2) which allows archiving data in DVD (digital versatile disk) by providing longer storage like and cross-platform compatibility.

c. SAMPRADAYA

SAMPRADAYA Which means *handling down of tradition*. It serves as a resource centre by documenting the various aspects of the South Indian (Karnatic) music. It was founded in 1980, by two students of South Indian music Michael, Nixon of Veena and Ludwig, Pesch of Flute, under the guidance of late Savitri Rajan with the purpose of documenting Sampradaya Sangeetam. The main contribution of this organization is to bring to light and preserve this music tradition by documentation, archiving of recordings, music books, journals, manuscripts etc. This organization is conducting in-depth taped interviews of eminent Carnatic musicians, and organizing seminars and workshops on important topics of South Indian Music. Sampradaya Audio listening library is only one of its kind in South India, where the public has free access to music recordings and archival collections. It has 8 separate music listening stations.

11. Conclusion

Preserving or archiving digital information is not new and has been explored at a variety of levels over the last five decades. Archivists are responsible for governmental and corporate records to convert the information in digital form for future generations. In the face of rapid technological obsolescence and to overcome the problem of media fragility, archivists have adopted the technique of “refreshing” digital information by copying it onto new media. Refreshing by digital information by copying will work as an effective preservation technique only as long as the information is encoded in a format that is independent of the particular hardware and software needed to use it and as long as there exists a software to manipulate the format in current use.

12. References

1. ARELMAT, Electronic Document Management (EDM)
2. <http://www.sum.uio.no/research>
3. Canadian Heritage Information Network, Preservation of Digital Information
4. <http://www.chin.gc.ca/english/digital-content/preservation/index.html>
5. PTFS – Digital Archiving Solutions
6. http://www.ptfs.com/services_digi_con.html
7. <http://www.chennaionline.com/bandstand/sampradaya.asp>
8. <http://www.vidhyanidhi.org.in/asp/search/metadata.html>
9. <http://www.sandiego/edu/theo/risa-l/archive/msg04939.html>

About Author

Mrs. Y. V. Ramana is serving as an Assistant Librarian, at Vellore Institute of Technology (Deemed University). She received her master's degrees in Ancient Indian History and Sociology from Kanpur University. Subsequently she did her Bachelors Degree in Library Science from Lucknow University and Masters Degree in Library and Information Science from Annamalai University. She has been working as Assistant Librarian in several institutions since 1990 and her field of interest is Digital Libraries. Email: yvramana58@yahoo.com