

Digitization as a Means of Preservation of Manuscripts: Case study of Osmania University Library

Fazluddin Ahmed

Abstract

This paper presents digitization as a means of preservation of manuscripts. It discusses the hardware and software available for digitization. A complete overview of digitization process has also been explained.

Keywords: Digitization, Digital Camera, OCR, Metadata, PDF

1. Introduction

Traditional libraries were only store house of books. The books were kept in locked cupboards and access to them was limited. The traditional libraries were static in the aspect of development. Open access system and preparation of catalogue cards revolutionized the library and its services. Computers were used in the early sixties by the libraries for their various operations. Libraries currently have experienced technological change in information storage and its retrieval. The medium of information storage has changed from clay tablets, palm leaves to paper and now to electronic and optical media.

The traditional preservation method constitute of all form of direct actions aimed at the life expectancy of undamaged or damaged elements of manuscript like mechanical cleaning solvent cleaning etc .Machine made paper made of wood pulp containing harmful acidic ingredients that caused deterioration was used.

Preservation ensures that people present and future will have access to the information that constitutes the documentary heritage.

In terms of preservation, digital conversion can certainly extend the life of a particular artifact. The use of original can be restricted and a high quality surrogate can be provided. Digitization enhances access to the artifact. as its image can be seen on the web by users all over the world. In addition to this it can be sent for offline viewing using a higher resolution uncompressed master file.

2. Factors Responsible for Digitization:

According to Becker, seven technological advances have created the environment for digitization.

- a. PCs are much more cost effective and powerful in comparison to a few years ago.
- b. Scanning technologies have become cheaper and better, allowing computers to convert manuscripts, pictures and color photos into digital files.
- c. Storage technologies have improved in terms of price and physical space required.
- d. The rapid growth of internet, which provides ever widening access to the digitized information.
- e. The emergence of more widely accepted standard protocols.
- f. The increased use of standards like SGML and HTML.



- g. The emergence of standard image formats and compression technologies making it possible to share images on computers.

Presently, libraries are linked through networks of various types. The changes brought by advancement in technology have been so extensive that it is difficult to assess their total effect but it is clear that libraries are in state of fundamental transformation from manual to electronic, digital and virtual library.

3. Tools for Digitization

The Hardware and Software tools for digitization are as follows.

3.1 Hardware

a. A Computer with the following Minimum Configuration

Pentium 4 class processor running at 2.0 GHz or higher
512MB minimum memory
160GB Hard drive
18" or larger monitor
128MB video card
CD-RW and/or DVD-R optical drive

b. Scanners

Scanners are of various types. Flat bed scanner is the most commonly used.

c. Digital camera

A digital camera uses a light sensitive processor chip to capture photographic images in digital form on a small diskette inserted in the camera or on flash memory chips. Digital cameras snap pictures by providing a real digital bit map to read directly into application.

3.2 Software

The software required is as follows

- HTML editor
- XML editor
- OCR software
- Image editor
- Page layout and design software
- PDF software

OCR:

The most important software is optical character recognition. OCR or text recognition is the process of electronically identifying text in a bit mapped page image or set of images and then creating a file containing that text in ASCII code or in specified word processing format. OCR is performed in order to make every word in the scanned documents computer readable and fully searchable without having to key every thing into the computer manually. OCR technology successfully permits the reading of documents containing a mixture of fonts in different sizes and styles

Image editor

This program allows saving images in multiple file formats, resizing images, cropping etc. The best example of this software is adobe Photoshop.

PAGE layout and design software. This software has more design features than word processors. Adobe page maker is one of the software available for this purpose.

PDF software

PDF is a widely used format .It provides a quick and easy solution for viewing and downloading the online files. This software is freely available on internet. This format's popularity can be gauged from the files available on internet.

4. Osmania University

4.1 History and back ground

Osmania University was established in 1918 and the university library was an integral part of it. In 1919 the library was divided into two sections the occidental and oriental sections with a modest collection of 2124 and 3313 volumes respectively. The library is working in the present building since 3rd August 1963.

4.2 Library system

The present Osmania university library system is complex in nature. It consists of University library, seminar libraries and constituent college libraries. Seminar libraries located in the respective department of various faculties in the university campus and constituent college libraries attached to the colleges.

4.3 Manuscripts available at university library

University library has a rich collection of manuscripts. The total number of manuscripts is 6825. There are two types of manuscripts available in the library one being paper and the other palm leaf. The manuscripts available are in Sanskrit, Kannada, Tamil, Marathi, Hindi, Arabic, Persian Urdu and Turkish language. The condition of paper manuscripts is better than that of the palm leaf. The manuscripts are preserved in specially made cupboards. There are two fumigation chambers available for preservation of manuscripts.. Some manuscripts which are rare and having artistic value are;

a. Bhagvata Scroll

This is a 64ft long scroll having width of 4 inches. It has complete text of the 12 Skandhas of Srimad Bhagvata.it contains miniature paintings of Ganesa

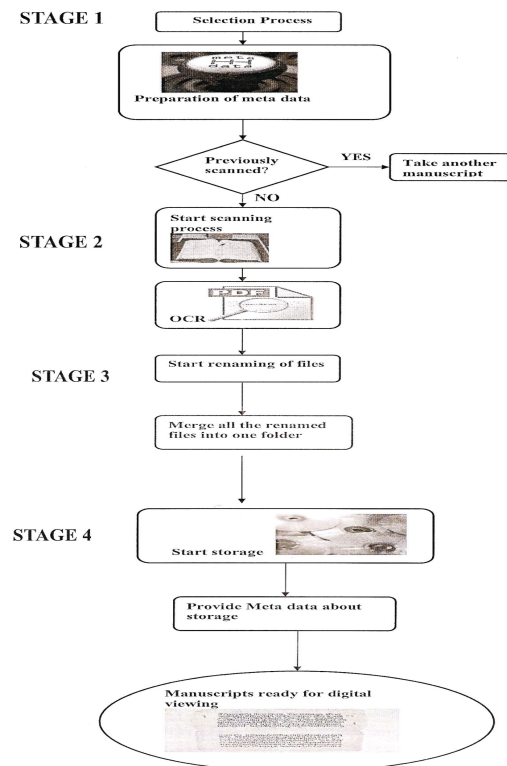
Saraswathi and the 10 Avatars of Vishnu. Original colors are used in this painting.

b. Mahabharata manuscripts :

There are two manuscripts of the Mahabharata of two different sizes. Each page contains miniature paintings of the characters mentioned in the text on that particular page



Digitization workflow



Osmania University signed a memorandum of standing (MOU) with Juma-al-majid centre for culture and heritage, Dubai for digitization of manuscripts available at university library. It was decided that the total collection of the manuscripts (paper and palm leaf) be digitized.

4.4 Selection Process

It was decided to get the total manuscripts digitized. Firstly the work of paper manuscript was started. The manuscripts which were in deteriorative condition were given treatment. The palm leaves manuscripts were mechanically cleaned and citronella oil was applied to the leaves. This helped for getting clear image of the leaf.

4.5 Preparation of Meta Data

A register was maintained for the manuscripts sent for digitization details of the manuscript such as author, title, folios, scribe date etc were noted in this register

A separate register has been maintained for each language after which the manuscripts were sent for digitization. The digitization centre also maintains Meta data which was filled up in the Performa. The digitization process starts with the first image as Meta data of the manuscripts.

Preparation of Meta data for various languages like Sanskrit, Kannada, Arabic, Marathi, and Turkish required the help of respective language experts. The multilingual nature of university library has helped in deciphering of title, author and other information. All the information has been written in Roman English. The palm leaf manuscripts have accession number mentioned on them. The authorship and title can't be read at the time of digitization. There were some complete palm leaf

Digitization as a Means of Preservation of Manuscripts---

manuscripts which do not even have accession number. They were also digitized and kept separately.

4.6 Scanning



The scanning work was done with the use of Sony cyber shot DSC_H5. Digital camera. This camera offers a sensor resolution of seven mega pixels along with the a new Carl zeiss vario- tessar 12xoptical zoom lens with 58mm threading and super steady shot optical image stabilization.

The H5 can shot as high as ISO 1000. The specifications of camera are as follows:

| | |
|--------------------------|--|
| Sensor Effective pixels. | 1 / 2.5 " Type LCD7.2 million |
| Image sizes | 3072 x 23043072 x 2048 (3:2)2592 x 19442048 x 15361632 x 1224640 x 480 |
| File format | JPEG |
| Lens zoomCarl | 36-432 mm equiv12 x optical zeiss vario-tessar F2. 8-3.7 |
| Digital zoom | 2 x (24 x total)up to 57 x smart zoom (dependent on selected resolution) |
| LCD Monitor | 3.0 inch TFT230 k pixels |
| Connectivity | USB 2.0 high speed AV out |
| Storage | Memory stick duo/ pro duo compatible32 MB internal memory |
| Power (optimal) | 2xNIMH batteries AC adapter |

The camera was fixed on the Hama professional camera stand. The stand has 4 bulbs of 150 watts each to focus on the image to be digitized. A mirror was fixed to check the clarity of the image available in the camera. Fixing the mirror helped to work even by sitting on the chair. The manuscripts were kept on the stand and the size of the image was fixed in the camera. After checking the image in the mirror, it was captured the number of images captured is limited by the size of the memory of camera. The memory card that had been used is of 1 GB. When the memory is full the images were transferred to the computer using the USB data cable. The images of the manuscripts were stored in the computer. Three persons were engaged for digitization work and a person was dedicated for storing the images in the computer and for carrying out the OCR. The images were verified and each of them was stored.

4.7 Renaming Files

The stored digital images in the computer for each manuscript were renamed. Renaming is done by using the accession number of the manuscript. Due to non-availability of author and title name, accession number was used for identifying a manuscript.

4.8 Merging files

When the files were ready after renaming, a new folder was created and the files were merged in this folder up to 700MB containing manuscripts varying from 5 to 6 in numbers.

4.9 Storage

When the folders containing manuscripts were ready, storage process is started. These manuscripts were written on CD. A detailed list of contents of CDs was prepared which act as an index to the cds.

4.10 Availability

The CDs along with the list is been handed over to the library by the digitization centre. The cds are checked with the list provided. The manuscripts will soon be available on web opac. The work has been started. University library is using New Gen Lib software for its library operations and web opac. The data entry for Sanskrit manuscripts (paper) has been started the availability can be checked by logging into library's opac.

5. Future

A new terabyte server has been purchased. The juma almajid centre has also provided two hard discs in which the total digitized manuscripts are made available. With the use of terra byte server the manuscripts will be available to users via intranet.

References

1. **Anderson, cokie g.** starting a digitization centre. Oxford: chandos publishing, 2004
2. **Arms, William y** .Digital libraries.London: MITpress, 2000
3. www.sony-asia.com/cybershot.
4. **Becker, j** The Guts of the Digital library information technology services maintaining system LC. information:262

About Author

Mr. Fazluddin Ahmed, Library Assistant,
University Library, Osmania University, Andhra Pradesh

E-mail: fazaluddinahmed@yahoo.com