



DIGITIZATION OF PUNE UNIVERSITY BULLETIN: A CASE STUDY

D. B. Shinde

Librarian

Arts & Commerce College,
Ashta, District : Sangli

Abstract

Library is a central place where knowledge is preserve for further use. At present libraries have become centers of information dissemination with the help of computer and communication technologies. In a digital world, libraries are emerging as a means of creation and dissemination of knowledge. Information today exists in many forms than just as printed material this lead to libraries into digital era. Library digitization is different from a digital library. Library digitization is the process of utilizing computers, databases, multimedia equipments, networks, video equipments and web technologies to electronically collect, classify, copy, compress, scan, store and transform conventional library information resources.

Today's digital libraries are built around Internet and web technologies with electronic materials as their building blocks. The increasing popularity of Internet and development in web technologies are catalyst to the concepts of digital library. This paper covers concepts of Digitization and Digital Library, Objectives of Study, Research Method and Procedure, Data Analysis & Interpretation and Findings & Suggestions.

Key words: Digital Library, Digitization, E-Resources, Information Technology, Scanning

Introduction

Library is a central place where knowledge is preserve for further use. At present libraries have become centers of information dissemination with the help of computer and communication technologies. In a digital world, libraries are emerging as a means of creation and dissemination of knowledge. Information today exists in many forms than just as printed material this lead to libraries into digital era.



The recent development in information technology has changed the concept of library. One of the most important developments in the library field is digitization of library material. Digital libraries have received wide attention in the recent years allowing access to digital information from anywhere across the world. The dramatic innovation in information and communication technology has emerged which have related a truly global environment for information exchange. The development has imposed certain responsibilities and challenges to library and information professionals. This dramatic revolution in IT especially in digital storage media, the user community is getting more advantages through the electronic media, electronic resources better than conventional method of information system. The concept has brought a phenomenal change in the information collection, preservation and dissemination of information

Concept of Digitization and Digital Library

Library digitization is different from a digital library. Library digitization is the process of utilizing computers, databases, multimedia equipments, networks, video equipments and web technologies to electronically collect, classify, copy, compress, scan, store and transform conventional library information resources. Digital library focuses on the process of making diverse library information resources electronically available, while a digital library is a platform for accessing, collecting, managing, searching and storing distributed digital information resources over the internet and World Wide Web.

According to Lesk, “a digital library is a collection of information that is both digitized and organized and which offers capabilities beyond those of the traditionally library”¹. A digital library is a library consisting of digital materials and services digital materials are items that



are stored, processed and transferred through digital devices and networks. Digital services are services that are delivered digitally over computer networks.

Features of Digitization

1. Seamless access to voluminous information resources.
2. Support multimedia contents for accessibility of information.
3. Provide user-friendly interface to access the information.
4. Provide hypertext linkages to various databases.
5. Provide client server architecture.
6. Support advanced search and retrieval of information.
7. Linkages with other national and international digital library system through networks.

Today's digital libraries are built around Internet and web technologies with electronic materials as their building blocks. The increasing popularity of Internet and development in web technologies are catalyst to the concepts of digital library.

Digital content will be two types: Items created and existing primarily in machine – readable format and materials converted from the traditional format i.e. printed text and pamphlets, images, motions pictures and recorded sound.

Computerization of the library during past few decades has focused heavily on creation of surrogate records of printed documents available in library or for providing computerized services on electronic and optical media. Books and printed record are always important and the libraries will continue to provide book and printed records but the libraries must move and adopt new technologies and should disseminate texts, images and sound available on computer.



Review of Literature

Jagdish Arora in his article has covered the technological evaluation, Cultural Revolution and contents enrichment that led to revolution in growth and development of digital libraries. It also deals with various aspects of building, assessing and organizing digital resources and collection. The describes the steps in the process of digitization technology, formats compression techniques and tools used in digital imaging as well as OCR and advocates for hybrid solution for preservation of digital information. Harsha Parekh has given emphasis on changes in library tasks, policies and practices for a new knowledge base and skills among the professionals. These include, technology handling abilities, managerial skill and better understanding of copyright and legal issues. Swami Chidrupananda has explained in house customization of application package for indexing, scanning, storing, images and retrieving information from articles. This include ideas on methodology and process for preserving old bound periodical and make them easily available to user in user friendly fashion. Dr. B. M. Panage carried out a research on conversation and preservation of library materials with special reference to digitization of rare materials in the year 2003. He studied in detail the traditional methods of conversion and preservation of library materials. This research also presents the review of the state of art of an existing digital library and selected the rare materials from the library and studies the problems there in. Sonali Samudra in her dissertation has studied the conceptual frame works of different library modes, i.e. traditional library, electronic library, Hybrid Library, Digital Library and Virtual library. It also deals with the study of acquisition and preservation of digital resources and an overview of current status of digital libraries all over the world along with India as well review of Process of digitization from some selected libraries in Pune City. Namrata Tiwari in her dissertation has studied on scanning of materials by open book scanner. She studied in detail the scanning of



materials, types of scanners, scanning software, Storage Format. This study is aimed at giving an insight into the working of the scanner.

Objectives

1. To study the need and importance of digitization
2. To study the various methods and procedures of digitization
3. To digitize the Pune University Bulletin issues and observe the difficulties faced while scanning
4. To work out the scanning and cropping process of the issues

Research Method and Procedure

It is an experimental type of research in which printed material of Pune University Bulletin is converted into digital form. The case study method is used for this study. The scanning and digitization process is followed. Four thousand five hundred eighty (4580) pages of eighteen (18) bond volumes of Pune University Bulletin of varying sizes are scanned with the help of Quickscan Software and digitized. During digitizing observation and problems were noted down. After digitizing the digitized images i.e. four thousand five hundred eighty (4580) pages were cropped in Scanfix Pro. As digital library and digitization is a new era of research in library and information science. So there is need to investigate into this area and to study all the literature published. The study is concerned with the latest information and to collect the information.

Digitization of Pune University Bulletin



The study concentrates on the digitizing of Pune University Bulletin in the digital library of Jaykar library. The Pune University Bulletin selected (10) Ten Year Old bound Volumes and (43) Forty three issues which published from September 1979 to 1991 and from 1997 to 2005. During 1979 to 1991's issues which are old volumes. The page was pale and turned yellowish. Some old bound volume issues of Pune University Bulletin were very fragile and delicate and even the print of some of the text was faint and taken a lot of care and was cautions while scanning these issues. The Pune University Bulletin issues were scanned with the help of open book scanner i.e. (Minolta PS 7000) which is available in Jaykar Libray

The details of these books are given in table 1 record of Scanned Pune University Bulletin issues published from 1979 to 1991 and 1997 to 2005.

Table 1: Scanned Issues of Pune University Bulletin

Sr. No.	Period	Issues	Pages
1	October 1979	1,2,3	178
2	August 1980 to July 1981	1,2,3,4,5,6,7,8,9,10,11,12	265
3	August 1981 to July 1982	1,2,3,4,5,6,7,8,9,10	278
4	September 1982 to August 1983	1,2,3,4,5,6,7,8,9,10,11,12	240
5	September 1984 to August 1985	1,4,5,6,7,10,11,12	258
6	September 1985 to August 1986	2,3,4,5,6,7,8,9,10,11,12	277
7	May 1988 to August 1989	1,3,5,7,8,9,11,12	201
8	January 1990 to December 1990	5,6,8,11,12	128
9	January 1991 to December 1991	5,6,8,9,11,12	198



10	March 1997 to December 1997	1,2,3,4	153
11	February 1998 to April 1998	5,6	88
12	January 1999 to December 1999	8,9,10,11	412
13	January 2000 to December 2000	12,13,14,15,16,17,18,19	611
14	February 2001 to December 2001	20,21	244
15	January 2002 to December 2002	22,23,24,25,26,27,28,29,30	375
16	January 2003 to December 2003	31,32,33,34,35,36	392
17	January 2004 to December 2004	37,38,39	127
18	January 2005 to December 2005	42,43	115
TOTAL PAGES			4580

The following procedure was used for scanning procedure.

1. Selected one bound Vol. at a time for scanning
2. Paced the document face up on the book table, aligned the top edge of the book with the stopper on the book table and the spine of the book with the center mark checked that the original is positioned with the corresponding size guide.
3. In 'My Computer', clicked on New Volume (E:) and created folder. Selected the appropriate subject and made anew folder, then wrote period of the issue in the folder. Double clicked on the title and made (2) two folders of OTIFF and PTIFF.
4. Stated Quickscan 1.0 version Software to scan the book placed on the book table from file, selected the scanned Minolta PS 7000 clicked OK and then went to option and did the 'Scanner setting'. In Scanner setting there were many things that has to be considered such as dots per inch went to more selected split (left page than right page). Then clicked the image edit label and on frame masking, finger masking, centering, center-line erase and front panel then clicked OK. Now after this selected



the page sizes of the book. There were various sizes of the book marked on the book table such as for portrait A6, B6, A5, B5, A4, B4, and A3 and for landscape A5, B5, A4, B4, A3, B3 and A2 select the appropriate size of the book accordingly.

5. Then again from file selected 'Scan batch to file' and then clicked on the subject and title. A File name was there in that typed for (4) Zero's and double clicked on OTIFF and single clicked on 'use schema' and clicked OK and then started scanning.
6. Pressed 'start' key on the left side of the scanner. The pressed start key turned red and scanning was started. When scanning completed, the start key turned green and the next scan was started.

The next procedure followed after scanning was cropping. The scanned materials were cropped to remove the black spots, stains or the marks already present on the pages of the original issue. At the time of cropping, problems did not occur as the procedure was followed carefully. The software used for cropping was Scanfix 4.21. The following procedure was used for cropping.

1. Start the Scanfix 4.21 version Software.
2. Clicked on 'Start' and 'Batch process' then in input file clicked on directory and selected New Vol. (E:), selected the subject double clicked on that and then double clicked on OTIFF, then clicked on open and clicked on OK.
3. After this in 'output file' typed four question marks and dot TIF (???? TIF). Selected directory, selected the same subject and title and made new folder wrote PTIFF in that folder and clicked on open and 'New output format', then clicked on OK.
4. After clicking on OK, warning was appeared in that clicked on 'disable this warning in the future' and clicked on OK.



5. The last step was to click on the report statistics, pause on completion, show % done and show images and lastly clicked started batch. Then the cropping process was started.

Findings

1. In Scanning, the images of the original pages are captured by the camera as it is.
2. During scanning the error occurs due to carelessness of human being. The following problems are faced by the investigator during scanning the document.
 - A) Bind Size for (portrait A6, B6, A5, B5, A4, B4, and A3 and for landscape A5, B5, A4, B4, A3, B3 and A2) were not detected. Therefore investigator had to reset the book or change sheet. This type of error occurs due to improper setting of the original book and if the original book touches the binder. The other reasons are that the torn pages lying outside the book and if the steel mirror remains dirty.
 - B) Pages not scanned properly due had binding and if more numbers of pages are loaded on one side than the other. Sometimes the print of the original book is faint and fade which results in improper scanning of pages
 - C) Disordering of pages may occur at the time of deleting the scanned pages which are not properly scanned, while deleting the pages the cursor is kept on the left side than the order of pages may get disturbed.
 - D) Computer hangs in between during scanning it may happen when Quickscan Software runs for a longer time and wrong keys are pressed.
3. When any error occurs in the scanner, an error message appears on the display, the start key turns red and scanning stops.
4. 150 pages/hour are scanned by the researcher if no error occurs. It is also found that sometimes due to errors researcher could do hardly 35 to 40 pages/hours.



5. The scanning speed depends on the book size, the portrait size A6 or B6 can scan one page in five seconds (5 Sec.) and landscape size A5 or B5 can scan one page in seven seconds (7 sec.)
6. The scanning, cropping processes are being carried out on the volumes. After the cropping, cropped pages are saved in OTIFF folder.
7. The scanfix 4.21 version software is used for cropping. This software is so adjusted that expect photo, figs., drawings, logo's and text all the marks, black spots and stains on the pages should be eliminated.
8. The scanner provides two images modes, 'Text' and 'Photo'. In case if the original book has only photo and no text, photo mode should be selected for scanning. When the original issue has both text and photographs, one should determine which be given higher priority. When text mode is selected clearer characters can be obtained and photographs can be scanned in both the image mode i.e. Text and Photo.
9. It was found that there is decrease in memory size occupied by each page after scanning, cropping ranging from 9.19% to 58.3%.

Suggestions

During digitizing process various errors were encountered due to carelessness of human being and not being cautious at the time of scanning therefore it is suggested that one should be very much cautious and careful while digitizing the material in order to minimize the errors. In order to avoid and solve these errors, the some following suggestions are:

1. To avoid the error 'Bind size of the book is not detected reset book or change sheet', then the book should be kept properly on the size marked on the book table and book



table should be adjusted. This error may also occur when the steel mirror is dirty so the steel mirror should be cleaned before scanning.

2. Pages are not scanned. This can be solved by making the binding loose and the pages should be pressed and supported by hand. If the pages are loaded on one side then the book should be supported with the help of other means so that level of both the sides can be maintained. If the print of the original book is faint in the case the contrast and brightness should be set on the 'scanner settings.
3. Disordering of pages, this error one can avoid by being carefully shifting the cursor to right hand side which avoids the disordering of pages (New pages will appear before the previous pages.
4. There is need to development OCR software which can recognize the various variations, fonts and old scripts because Digitization is done mainly for the preservation of old and rare books which are in old scripts and old fonts.

Conclusion

The recent development in information technology has changed the concept of library. One of the most important developments in the library field is digitization of library material. Digital libraries have received wide attention in the recent years allowing access to digital information from anywhere across the world.

Digitization is very important in today's world due to the special features of it. These features are such as Seamless access to voluminous information resources, Support multimedia contents for accessibility of information, Provide user-friendly interface to access the information, Provide hypertext linkages to various databases, Provide client server



architecture, Support advanced search and retrieval of information, Linkages with other national and international digital library system through networks.

In this study an attempt is made to digitize the Pune University Bulletin and it is found very beneficial to the library of Pune University. While making this digitization various errors were encountered due to carelessness of human being and not being cautious at the time of scanning therefore it is suggested that one should be very much cautious and careful while digitizing the material in order to minimize the errors. In this regards some important suggestions are given. This paper shows the need of digitization of libraries.

References

Arora, Jagdish “Building digital libraries overview: an overview” *DESIDOC Bulletin of Information Technology*, 21 (2001) P.13-14

Besser, Howard. “Procedure and practices for scanning”, (<http://sunsite.berkeley.edu/imaging/database/scanning/index.html>)

Cyric, Jiji & Deshmukh G.R. “Digitization of Libraries: In modern era” *ILA Bulletin* 38 (2002) P.68

Kaul, H.K. “Digitization of knowledge Resources: The 21st Century 2005 (2005) P.172-189

Lahiri, A. “Digitization a new mandate” *Information today and Tomorrow* 20 (2001) P.1-2

Lee, Stuars D. “Digital imaging: a partial hand book” *London Library Association* (2001) P.3

Pan, Rosalind and Higgins, Richard “Digitization project at Durham University Library: An overview” *Program electronic library & information system ASLIB* 35 (2001) P.355-368

Panage, B.M. “Conservation and preservation of library material with special reference to digitization of rare materials” *Theses* (2003)

Paul, Conway “Digital technology made simpler: In preservation of library and Archival material a manual (1999)



Parekh, Harsha “The digital movement: challenges and opportunities for librarians *IASLIC Bulletin* 48 (2003) P.46-52

Samudra, Sonali “Digitization in libraries: An art *Dissertation* (2003)

Shoaf, Eric C. “Preservation and Digitization: Trends and implication” *Advanced in librarianship* (1999) P.203-238

Swami, Chidrupananda “Indexing, Scanning and Storing images of Articles” *IASLIC Bulletin* 48 (2003) P.195-205

Tiwari, Namarta “Scanning materials by open book scanner” *Dissertation* (2004) P. 7-8

Singh, Surdev “Digital libraries and digitization” *Ess Ess Publication*, New Delhi (2011) P. 1, 316-325.