Impact Factor: 3.1028(UIF)



e-Library Science Research Journal



Vol ume - 5 | Issue - 8 | June - 2017

# DIGITIZATION OF LIBRARY RESOURCES AND THE FORMATION OF DIGITAL LIBRARIES

Miss. Avinashi M. Mahant<sup>1</sup> and Dr. Mahesh Darbar<sup>2</sup> <sup>1</sup>Ph. D Research Scholar , Rai University , Ahmedabad.



# ABSTRACT

A n electronic library (also referred to as digital library or digital repository) is a focused collection of digital objects that can include text, visual material, audio material, video material, stored as electronic media formats (as opposed to print, micro form, or other media), along with means for organizing, storing, and retrieving the files and media contained in the library collection. Digital libraries can vary immensely in size and scope, and can be maintained by individuals, organizations, or affiliated with established physical library buildings or institutions, or with academic institutions.[1] The electronic content may be stored locally, or accessed remotely via computer networks. An electronic library is a type of information retrieval system.

KEYWORDS: electronic library, digital objects, electronic media formats.

# COMPONENTS

## The components of a digital library are:

1 Infrastructure 2 Digital Collection 3 Systems function 4 Telecommunication facility 5 Human resources

# Planning for Digital Library

A digital library committee should be formed to plan for its creation and maintenance. The members must be from various library departments, and, if necessary, consultants can be hired. There are at least two ways of developing a digital library: converting a traditional library into a digital library, and direct development of a digital library.

# Planning includes:

1 IT Infrastructure 2 Digitization 3 Access 4 Staffing 5 Furniture, equipment, and space 6 Services 7 Funding

#### PLANNING

Planning mainly involves identifying various tasks related to creating a digital library collection, developing strategies for handling these tasks, identifying required resources and formulating a timeline for accomplishing these tasks. If there is a need to have a large digital project, you may consider conducting a feasibility study to assess the viability of the project before detailed planning. The outcome of the feasibility study could be a formal proposal for obtaining management approval or grant for the project.

a. The first step in planning a digital library collection development project is to specify the need for creating the digital library collection, its purpose and target user community. You should indicate if management, the users or others have expressed this need and defined what this need is. The purpose could be improving preservation of some rare or delicate materials, improving access to and the visibility of certain material or facilitating re-use of documents. It is important to identify the target user community for a digital library collection and their profile.

b. There is the need to define the source material that constitutes the digital library collections and the key attributes of this source material. Examples of source material include project reports, staff publications, working papers, theses, dissertation, audio and video lectures, songs and musical scores etc. There is also the need to specify what portion of the material is to be digitized and if all the material or only a sub-set will be covered in the digital collection. Remember to assess copyright restrictions.

c. Define the key features of the digital library collection you plan to build. Identify the nature of the collection e.g. static or dynamic. Indicate the type of usages you would allow the users to adhere to and the kind of service delivery they should expect from you e.g. CDROM or on-line or both. Define metadata, search and retrieval requirements.

d. The important task in creating a digital library collection is the conversion of the source materials available in hardcopy into a digital format. There should be a clear cut statement about the related requirements and their processes, namely:

i. How to convert the source material into required digital format.

ii. What are the digitization requirements?

iii. The workflow involved in digitizing the source material.

e. Identify the resources and money required for creating and maintaining digital collections. There is a need to identify:

i. What type of information technology (IT) infrastructure is required for establishing and maintaining the digital collections?

ii. What are the personnel requirements and

iii. What are the financial requirements involve for setting up and maintaining the collection.

f. Finally, there is the need to define how the project is going to be implemented and what the major milestones and time requirements are?

#### IMPLEMENTATION

Planning is followed by implementation. That is getting down to the actual steps required to set up the collection. This means that there must be a need to obtain the management approval for the plan and the required resources before proceeding with the implementation. There is a need to identify and designate a project manager to lead the Implementation of the digital project. For large digital library projects, it is Essential to have a full time project manager for the project period. The Implementation of a digital library project involves the following Activities.

i. Establish the project team

ii. Set up the Information Technology (IT) infrastructure

iii. Procure and install digital library software

iv. Finalize policies and specifications

v. Complete arrangement of workflow for digitization

vi. Set up the digital library collection site in case of Internet Distribution

vii. Obtain copyright permissions and

viii. Release the digital library collection for use.

## **PROMOTION AND PROVISION OF SERVICES**

The digital library collection created should be visible, and it should provide an easy access for users. Oneway of achieving this is to include links to the collection site in the appropriate pages of the library website and other related on-line services in the organization.

In addition to, or in the absence of remote on-line access to the digital collection, there is the need to explore other modes of providing access to the digital collection. These may include:

i. Setting up local public access computers on the library Local Area Network.

ii. Provision of e-mail based services and

iii. CD-ROM based distribution of the collection.

## DIFFERENT STAGES IN DIGITIZING DOCUMENTS

Six stages in digitizing documents for a digital library: Registering, Scanning, Optical Character Recognition, Proofreading and formatting and producing the Final Version.

#### i. Registering

Before scanning large number of documents, there is the need to first register them and use a filing system to keep their track. If not, you risk misplacing hardcopies, losing files, skipping steps in the process or duplicating work, perhaps without realizing it. There is also the risk of losing electronic versions of files because they have been misnamed or saved in the wrong subdirectory. Moreover, a good filing system is vital, so everyone in the digitizing team knows what he is supposed to do, and he can fill in for another person in case of absence.

#### ii. Scanning documents

It is necessary to clean and dust off the documents to be scanned; make sure that all the pages are present and in the right order. If the document is in poor condition, try to find a fresh copy. If it is a sheet fed scanner, cut the document open to get individual sheets to feed through the scanner. If necessary, you can rebind the documents later. If you do not want to damage the documents, you can photocopy each page and feed in the photocopy through the scanner, though this uses a lot of paper and reduces the quality of the scan.

To scan a document on a flatbed scanner, place it face down on the scanner platen or put the pages into the sheet feeder. Then, in the software, choose a setting, resolution and colour and scan each page of the document at the settings you have chosen.

#### iii. Optical Character Recognition (OCR)

Optical Character Recognition (OCR) software converts a scanned image into a text file that a word processor can read. To do this, it must first recognize where the text is on the page. The software breaks the text blocks down into lines or into an individual character. It tries to match the image of each letter against patterns it recognizes as an "a", "b", etc. There is a problem to encounter with languages that use Latin scripts with accented characters. As a solution, you should use the OCR software that is specific for language.

#### iv. Proof reading.

This is the act of making corrections to the document text and layout. This is done in two ways:

a. Comparing the scanned text on the screen with the hardcopy and entering the corrections directly into the computer. The word processor's spellchecker will help in spelling errors quickly.

b. Printing out the scanned text and comparing it with the original copy. Mark any corrections on the printout, and then enter them into the Computer. This is a slower method, but may be the best option if there are no enough computers for each proofreader.

#### v. Reformatting

The Optical Character Recognition (OCR) software may produce a document that consists of straight text, no columns, no headers and footers. There is the need to reinsert these by hand or correct where they appear on the page. There may be also need to change the typeface, heading styles and so on, to make the document more attractive and readable. Alternatively, you may be able to adjust the settings of your OCR Program to preserve the layout of the page.

## vi. Final Version

For many documents, there is a need to add some information to the text so that readers can identify it easily. As for a book you must make sure that the book title, the author or the editor, the publisher and the publication date are all included. As for chapter in a book, you should include the title and the author of that chapter and the original page numbers in the printed version of the book. As for the journal articles you should include the journal title, the date, the volume and the issue number, the article title and the authors and the page numbers in the original printed journal. In other words there is the need to add Metadata to describe each document.

## Technology Infrastructure and Personnel

Several resources are required for the creation of digital library collections, their maintenance and provision of services. The two major resources needed are technology infrastructure and personnel.

## Infrastructure

Access to a digital library collection can be provided on-line or off-line. The On-line access today typically means that the client uses a web browser on a desktop computer or laptop and access the collection by connecting to the digital library website over the Internet. The On-line access requires a connection to the Internet or to an internal network (Intranet). In Off-line access, the digital library is not accessible over a network. One way of providing an Off-line access to a digital library collection is to receive and respond to the user queries over e-mail. Another way is to distribute the digital library collection on a CD-ROM.

A digital library project would typically require the following equipment: Server computer, Desktop computers, Digitization equipment, Network connectivity and other equipment. Another aspect is the software to be used in digital library. The Digital library software works with the web server in providing various digital library functionalities including creation, organization, maintenance, indexing, search and retrieval. In choosing the software, some features should be taken into consideration. These include: Support for different document types, Support for customized metadata, Collection administration, Support for standards like Dublin core metadata standard, Search and retrieval and Multi-lingual support.

Several free digital library software packages are now available which could facilitate the easy creation and sharing of information through digital library collections. Examples of open source free digital library software include: Greenstone Digital Library software by New Zealand Digital Library; Academic Research in the Netherlands On-line (ARND); Tilburg University, The Netherlands; CDSware; CERN Document server software, Geneva, Switzerland; D-space; MIT Libraries, Cambridge, MA USA. etc.

#### Personnel

Personnel are most important digital library's resource, not only during its initial creation and set up, but also for its operation, maintenance and provision of services. Since the access to the digital library is easy, compared to a physical library, more users are likely to access it. If the digital library does not meet the expectations of the users in terms of currency and quality of content, they will lose confidence, and it is likely for them not to visit the digital library again.

It is therefore important to assign the personnel with the right skills and attitude to handle the various tasks associated with the digital library project.

# Broadly speaking, the personnel will be required for the following tasks:

a. Project management.

ii. Selection and preparation of source material

iii. Digitization and conversion

iv. Cataloguing and metadata assignment

- v. Quality assessment
- vi. System administration and maintenance of digital library server and website.
- vii. System analysis/programming for digital library application/interface development

viii. Promotion and provisions of services.

Moreover, the rapid changes in the digital library technologies require constant re-training and repositioning of staff for an effective practice in technological application.

# Greenstone Digital Library Software

Greenstone is a freely available suite of software for building and distributing digital library collections. It provides a new way of organizing information and publishing it on the Internet or on the CD-ROM. The Greenstone is open source software, issued under the terms of the GNU General Public License. The aim of the software is to empower the users, particularly in the Universities, Libraries and other public service institutions, to build digital libraries. The software has the following features such as multi-platform availability for windows, linux, access and distributed through the Internet, Intranet and CD-ROM, powerful indexing from full-text and creation of indexes for various metadata, powerful search and browse, support different file formats (html, pdf, doc rtf, ppt etc), extensibility by allowing customization and configuration. Greenstone also allows the building of non-textual multimedia such as audio, video and pictures accompanied by textual description to allow for searching and browsing.

# **Creation of Digital Resources**

1 Database of digital material that is open to all users over the campus-wide LAN.

- 2 High bandwidth Internet connectivity
- 3 Focus selectively on acquiring digital resources
- 4 Electronic journals, and gradual elimination of print subscriptions
- 5 Licensed databases
- 6 Creation of local digital content available within the university

# Function of Digital Library

- 1 Access to large amounts of information to users wherever they are and whenever they need it.
- 2 Access to primary information sources.
- 3 Support multimedia content along with text
- 4 Network accessibility on Intranet and Internet
- 5 User-friendly interface
- 6 Hypertext links for navigation
- 7 Client-server architecture
- 8 Advanced search and retrieval.

9 Integration with other digital libraries.

# Objective of Digital Library

1 Expedite the systematic development of procedures to collect, store, and organize, information in digital form.

2 Promote efficient delivery of information economically to all users.

3 Encourage co-operative efforts in research resource, computing, and communication networks.

4 Strengthen communication and collaboration between and among educational institutions.

5 Take leadership role in the generation and dissemination of knowledge

#### **Advantages**

#### The advantages of digital libraries

No physical boundary. The user of a digital library need not to go to the library physically; people from all over the world can gain access to the same information, as long as an Internet connection is available.

Round the clock availability A major advantage of digital libraries is that people can gain access 24/7 to the information.

Multiple accesses. The same resources can be used simultaneously by a number of institutions and patrons. This may not be the case for copyrighted material: a library may have a license for "lending out" only one copy at a time; this is achieved with a system of digital rights management where a resource can become inaccessible after expiration of the lending period or after the lender chooses to make it inaccessible (equivalent to returning the resource).

Information retrieval. The user is able to use any search term (word, phrase, title, name, subject) to search the entire collection. Digital libraries can provide very user-friendly interfaces, giving click able access to its resources.

Preservation and conservation. Digitization is not a long-term preservation solution for physical collections, but does succeed in providing access copies for materials that would otherwise fall to degradation from repeated use. Digitized collections and born-digital objects pose many preservation and conservation concerns that analog materials do not. Please see the following "Problems" section of this page for examples.

Space. Whereas traditional libraries are limited by storage space, digital libraries have the potential to store much more information; simply because digital information requires very little physical space to contain them and media storage technologies are more affordable than ever before.

Added value. Certain characteristics of objects, primarily the quality of images, may be improved. Digitization can enhance legibility and remove visible flaws such as stains and discoloration.

# Easily accessible.

#### Limitations

1 Lack of screening or validation

2 Lack of preservation of a fixed copy (for the record and for duplicating scientific research)

3 Lack of preservation of "best in class"

4 Difficulty in knowing and locating everything that is available, and differentiating valuable from useless information.

5 Job loss for traditional publishers and librarians

6 Costs are spread and many become hidden.

#### CONCLUSION

Digitization has opened up new audiences and services for libraries, and it needs to be integrated into the plans and policies of any institution to maximize its effectiveness. Digitization is a complex process with many crucial dependencies between different stages over time. Utilizing a holistic Life-cycle approach for digitization initiatives will help develop sustainable and successful project.

It is hoped that the approach of the issues outlined, the software mentioned in this paper and the references to more detailed source and past project will contribute to the future success of initiating digitization of library Resources.



Miss. Avinashi M. Mahant Ph. D Research Scholar, Rai University, Ahmedabad.

Available online at www.lbp.world