

Digitization of Library Resources: Challenges and Implications

By

Goshie RhodaWusa and Jimada Aisha

Kashim Ibrahim Library A.B.U. Zaria

Rhoda4eva@yahoo.com and shatuindo@gmail.com

08035054080, 08066968840

A Paper presented at the Nigerian Library Association (NLA) Annual General Meeting (AGM)/Conference.

Venue: Kaduna State University (KASU)

Date: 12th December, 2012

Abstract

This paper examines the concept of digitization, its implications, challenges and the digitization software. It also discusses the challenges often encountered during digitization and the implications. Digitization is a process in which materials are converted from the hard copies to electronic copies. The major purpose of digitalization is: to enhance access and improve preservation of library materials. A number of challenges are encountered in the process of digitizing library materials. These challenges include; finance, Bandwidth and copyright, which have implications for library. The implication is that library will now be able to provide access to materials that were indigenous to them. It was concluded that digitization is an essential task in modern day libraries, because of the current Challenges, and the need to go digital, that is, improve the efficiency of information search mechanisms and enhance access to library resources.

Introduction

In recent times, universities in Africa have embarked on integration of technology in their operations. This serves as a tool for improvement and development, especially in a situation where attention is drawn to the academic contents and ways of digitizing and preserving them. University libraries are fully involved in this development as they are the heart of the University, being that they house the intellectual contents of the institutions. This effort to digitize the intellectual property of the institution is what is known as digitization.

Digitization has been possible for nearly every format and medium presently held by libraries, from maps to manuscripts, and moving images to musical recordings. The use of hardware and software for capturing an item and converting it into bits and bytes, matched by a quickly developing set of practices for describing and retrieving digital objects, is giving form to the talk of a "library without walls." But such a virtual library has a very real price. Managers of cultural institutions and those responsible for policy matters related to digitization often find themselves struggling not only to understand the new technologies, but also, and more importantly, to grasp the implications of those technologies and to understand what digitization of their collections means for their institution, its patrons, and the public.

It is often said that digital information is transforming the way we learn, the way we communicate, even the way we think. It is also changing the way that libraries and archives not only work, but, more fundamentally, the very work that they do. It is easy to overstate and underestimate the transformative power of a new technology, especially when we do not yet understand the full implications of its many applications.

A critical mass of experience is accumulating among libraries and archives active in digitizing parts of their collections, ranging in size from the Library of Congress, the National Archives, and major research libraries in the Digital Library Federation, to smaller institutions. Their experiences reveal patterns that can help us assess when the technology is able to meet expectations for improvement of traditional library services, when it cannot, and when it may do so, but not in a cost-effective manner. In a digital library, resources are stored and made available in electronic forms, and the services of the library are also made available electronically. Rosenberg (2005) noted that these services are made available frequently over the

Internet so that users can access them remotely. This refers to e-services which most libraries are trying to embrace in the digital environment. Rosenberg stated that as libraries embrace the digital environment, their most crucial role is not that of providing e-resources, but of establishing services that facilitate access to the information available. According to Fabunmi, Paris and Fabunmi (2006), library digitization has become part of the work of librarians, and most libraries are involved in digitization. Libraries in Nigeria are not left aside in this digitization trend in universities.

Digitization

Digitization is a process in which materials are converted from the hard copies to electronic copies. Digitization can involve simple data conversion from catalog cards or paper to digital form, video and audio migration to digital form.

The term 'digitization' has been variously defined by different authors. Digitizing is the art of converting the contents of a document from hard copy into machine-readable format. Digitization implies conversion of documents and art works into digital images (Fabunmi, Paris, and Fabunmi 2006). Digitization makes materials available electronically and the "conversion of non-digital material to digital form" (Tsebe, 2005). Digitizing or digitization is representing an object, image, document a signal (usually an analog signal) by a discrete set of its points or samples Wikipedia (2011). The result is called "digital representation" or, more specifically, a "digital image", for the object, and "digital form", for the signal. Fabunmi, Paris and Fabunmi (2006) refers to digital images as electronic copies of documents.

The aim of digitizing library materials is for preservation and easy access by any user or researcher. In the words of Fabunmi, Paris and Fabunmi (2006) Digitization improves access to library resources. By digitizing library collections, information will be accessible to all instead of a group of researchers. It is a project for the digitization of Theses and Dissertations in Nigerian Universities which started with the University of Jos and the Obafemi Awolowo University (OAU), Ile-Ife (Okiy, 2008). Digital projects allow users to search for collections rapidly and comprehensively from anywhere at any time. Digitization makes the invisible to be visible. Several users can access the same material the same time without hindrance. It also removes the problem of distance, as users do not have to travel to libraries that possess the hard copies of

library materials before they can access and use such materials. This entails that the digital library would be open at any time for consultation of materials. Materials uploaded on the website are always there for people to consult, except when erased by the website administrator.

However, there are three major reasons for digitization

- The need to preserve endangered library resources
- Improvement of the efficiency of information search mechanisms, and
- Digitization improves access to library resources.

Most libraries are digitizing materials which might be lost in the future, such as old manuscripts, research projects, photo images, analogue maps, non-live musical recordings, government official gazettes and several other historical records. Digitization is useful in preserving precious materials. Making high-quality digital images available electronically will reduce wear and tear of fragile items. However, digital copy should not be seen as a replacement for the original piece, therefore original document should be cared for even after digitization. Preservation remains a secondary benefit of digital projects. Warraich N.F. and Tahira (2009)

One of the advantages of digitization is the ability to search for an item electronically. Silkroad (2009) noted that rather than scan through table of contents in a book or newspaper, you can do a quick electronic search and find what you are looking for in seconds. It saves the time of researchers, students and corporations. This implies that a large number of users can access a single material at the same time. This also saves time and it goes in line with Ranganathan's fourth law of library science which states: "save the time of the reader". Digitization also helps to reduce handling and use of fragile documents. As these invaluable resources have become old, they need less handling and an effective back-up is established.

Hirtle (2002), cited in Fabunmi, Paris and Fabunmi (2006) argued that the biggest benefit of digitization is the tremendous increase in the use of digitized material. He used the cases of the Cornell University and the University of Michigan as examples. In Cornell, prior to digitization, a few volumes of the hard copies are circulated each year. However, with digitization, the views per month are above 4,000 web pages. Michigan has over 5,000 web page views per day. Michigan started earlier than Cornell.

Though digitization is sometimes loosely referred to as preservation, it is clear that, so far, digital resources are at their best when facilitating access to information and weakest when assigned the traditional library responsibility of preservation. Regrettably, because digitization is a type of reformatting, like microfilming, it is often confused with preservation microfilming and seen as a superior, if as yet more expensive, form of preservation reformatting. Digital imaging is not preservation, however much is gained by digitizing, but permanence and authenticity, at this juncture of technological development, is not among those gains. Eke, Helen Nneka (2009)

The reasons for the weakness of digitization as a preservation treatment are complex. Microfilm, the preservation reformatting medium of choice, is projected to last several centuries when made on silver halide film and kept in a stable environment. It requires only a lens and a light to read, unlike computer files, which require hardware and software, both of which are developed in often proprietary forms that quickly become obsolete, rendering information on them inaccessible

Access to Digital Information

Digital files can provide extraordinary access to information. Digital surrogates can bring together research materials that are widely scattered about the globe, allowing viewers to conflate collections and compare items that can be examined side by side solely by virtue of digital representation. The easy access to reference surrogates images that provide a great deal of the information contained in the original, even if at fairly low resolution is a boon to researchers when developing efficient and effective research strategies through the use of thumbnail images, which do not require high resolution, one can at a minimum acquaint oneself with the source enough to know whether or not one needs to consult the original. Very often one can make do with the digital surrogate because it provides all the information required.

One must think about the nature of the source materials (color, black and white, or shades of gray) and the use of the images (who will be consulting them and for what) when making decisions about the parameters for image capture. The quality and utility of an image depend upon the technology of capture and display, and the usefulness of an image, even if only for reference, can be severely compromised by a low-resolution monitor on which the image will be displayed. While work is ongoing to address the quality control and variability of computer

monitors, as yet the lack of control over display mechanisms constitutes one of the weakest links in the digital chain of transmission.

Digitization in the Networked Environment

Making information available on the Internet removes the very barriers from use that we take for granted in physical collections. No one has to travel to a library, nor do they have to present proof of their serious research interest in order to gain access to complex, disturbing, and uninterpreted material. There are ways to build in electronic barriers to access for all or portions of a site, using much the same technology that commercial entities use in granting fee-based access. However, constructing these barriers adds a layer of administrative complexity to managing the site that libraries and archives may not be prepared to take on, even if the technology does exist, only when digitization is viewed specifically as a form of publishing, and not simply as another way to make resources available to researchers, are the thornier issues of selection for conversion put into an editorial context that provides a strong intellectual and ethical basis for imaginative selection of complex materials.

Process of Digitization

In contemplating a digital conversion project, an institution must ask itself what can be gained from digitization, and whether the value added is worth the price. Many libraries have begun the difficult task of developing criteria for selecting for digitization and have published their criteria on the Internet. Columbia University, for example, was among the first to post guidelines for selection of materials for digital conversion, which include the criterion of added value. They define the added value of digital capture as enhanced intellectual control through creation of new finding aids, links to bibliographic records, and development of indices and other tools; increased and enriched use through the ability to search widely, manipulating images and text, and to study disparate images in new contexts encouragement of new scholarly use through the provision of enhanced resources in the form of widespread dissemination of local or unique collections; enhanced use through improved quality of image, for example, improved legibility of faded or stained documents; and creation of a "virtual collection" through the flexible integration and synthesis of a variety of formats, or of related materials scattered among many locations.

The following are therefore the process of Digitization in Kashim Ibrahim Library of the Ahmadu Bello University, Zaria.

- Acquisition of thesis and Dissertations
- Pulling of the thesis and dissertations into loose sheets
- Scanning of the documents
- Launching the Abby Fine Reader, using the optical character recognition (OCR) software
- Loading of loose sheets into document feeder on the scanner
- Clicking on scan to start scanning
- Capturing of pages as text or picture. Note: for thesis and dissertations you only capture the title page and the abstract as text, while other pages are captured as pictures.
- Read the captured images
- Check spell text pages only
- Save batch as a portable document file (PDF). Note: save the batch with the title of the work
- Uploading into a digitization software

Digitization softwares

The digitization softwares are either open source software or proprietary software. The open source softwares includes; Dspace, greenstone, Open cms, Ametys cms, content dm, Fedora Dot cms, NuxeoEP, Magnolia, Open Km, engage digitizer, Yanel, and Liferay. The proprietary softwares includes; Episerver Cms, Adobe Cq5, Cascade server, Alfresco, open text web and AteX software.

Greenstone Digital Library Software

Greenstone is a suite of software for building and distributing digital library collections. It is not a digital library but a tool for building digital libraries. It provides a new way of organizing information and publishing it on the internet in the form of a fully- searchable. Green stone runs on all versions of windows, and Unix/Linux, and Mac OS-X. It is very easy to install.

Content dms

OCLC Content dm is software that handles the storage, management, and delivery of digital collections across the web. CONTENTDM stores images, newspapers, books, maps, slide libraries, or audio/video. All types of cultural heritage institutions (archives, special collections, libraries, historical societies) and digital collaboratives. CONTENTdm can also be used as a digital repository for an institution's Digital Repository Software.

DSpace

Originally developed at MIT, DSpace is open source software used for building digital repositories. DSpace preserves and provides access to various types of digital content including text, images, moving images, mpegs, and data sets. Academic, non-profit, and commercial organizations are using DSpace. Visit the DSpace website to browse the list of users.

Greenstone

Greenstone is an open source option for building and providing access to digital library collections. Greenstone is produced by the New Zealand Digital Library Project at the University of Waikato. Universities, libraries, and other public service institutions use Greenstone. If you're looking for specific examples of Greenstone implementations, visit the Greenstone website.

Fedora Commons Repository Software

Originally developed at Cornell University, Fedora Commons is another open source software option available for building digital repositories. Consortia, museums, university libraries and archives, government agencies, national and public libraries and archives use Fedora. View the Fedora Commons registry for more information.

Enguage Digitizer

This open source, digitizing software converts an image file showing a graph or map, into numbers. The image file can come from a scanner, digital camera or screenshot. The numbers

can be read on the screen, and written or copied to a spreadsheet. The process starts with an image file containing a graph or map. The final result is digitized data that can be used by other tools such as Microsoft Excel and Gnumeric. The term "Engauge" in Engauge Digitizer was invented for this project, since there seems to be no similar term in common use.

The Problem

The manual system of searching for information and materials in the traditional library does not permit multiple use of the same material by different library users unlike the Online library services. It is inefficient and time consuming it takes 30 to 35 minutes to digitize a thesis, hence the need to exploit the advantages of the digital library which enables provision of online library services. However, there are a lot of challenges facing the setting of a digital library or conversion to digital status. Digitization is time consuming it takes 30 to 35 minutes to digitize a thesis material depending on the volume of the thesis. It is also a very expensive endeavour. The cost is usually between 3 to 5 million Naira depending on the size of the library and the volumes of the library collections.

The space required to store paper documents can be a problem. Digitizing your documents renders them exquisitely portable you can store an entire library on your e-book reader with ease. And because paper documents can be turned into editable computer documents, they become searchable. Compare typing "Roosevelt" in a search field with spending all day scanning microfiche and old newspapers by eye to research the Square Deal or the New Deal. The digital document is a boon to researchers the world over. However, there are a lot of challenges. There is a need to undertake a psychological preparation of the employees, so that they will not resist digitization of the library resources. There is also a need to retrain the workers. Creating a digital library is a very expensive venture which requires adequate planning and monitoring. The major problem is lack of knowledge and skills required for digitizing, hence most digitization projects often run into problems. There is a need to design flexible and compatible programs. In addition, the interface should be user- friendly, so that users can search for information with ease. It is not all electronic copies of documents that will suit the

application format, hence the need to take this into consideration during conversion or digitization exercises.

Implications for Library

This shift to digitization in the contemporary media world has created implications for traditional mass media products, however these "limitations are still very unclear" McQuail, (2000). The more technology advances, the more converged the realm of mass media will become with less need for traditional communication technologies. For example, the Internet has transformed many communication norms, creating more efficiency for not only individuals, but also for businesses. However, McQuail suggests traditional media have also benefited greatly from new media, allowing more effective and efficient resources available.

As in other forms of reformatting, digital scanning has implications for the original item and its physical integrity. Depending on the policy of a library or archival institution, the original of a scanned item may or may not be retained after reformatting to the extent that a reader can make do without handling the original, the digital preservation surrogate can serve to protect it from wear and tear. If there is concern that the scanning process could damage materials, one would choose to scan a film version of the original.

Library will now be able to provide access to materials that are indigenous to them and the library have to acquire the necessary technology for digitization.

Challenges of Digitization

Most universities involved in digitization of their resources face similar challenges. Generally, the following factors pose challenges to the digitization initiative in most universities.

Legal Aspects

This is related to intellectual property rights. A major challenge for digital libraries is complying with copyright, intellectual property rights and related issues like plagiarism Warraich (2008). This is an aspect where librarians and researchers need to take precaution. There is an increasing unease among members of the library community that copyright changes will adversely affect the ability of libraries to provide digital collections and services.

Finance

A lot of money is required for the purchase of diesel, for the generator; repairs of systems, generators and other equipments demand a lot of attention financially. Sometimes, the work ceases if there is no release of money. Another area where money is required is human resources development. Training needs money and some universities find it challenging to issue out money for such purposes, especially during economic crises as of present.

Technical Support and Security

For librarians, security for digital information is a great challenge that demands attention. Piracy has been a great problem that needs solutions. Virus attack on the files and systems is also a big threat. Such incident occurred sometime mid 2008, when a particular virus named "RailaOdinga" attacked all the scanned files and it went ahead infecting all the computers used for the scanning and uploading of files. Eventually all the files were corrupted and lost, and some of the uploaded files had virus on them, and so could not be properly downloaded by users. To endure, digital libraries need the latest technology and more technical staff or experts to guard against such mishaps.

Technophobia

Many staff involved in the digitization exercise were subjected to series of training and re-training before they were able to pick up on the technology involved.

Bandwidths

Developing countries may have limited bandwidth available. This is a problem that is being experienced in most university libraries. Poor connectivity has always affected the rate at which files are uploaded.

Difficulty in digitizing some materials

Most academic staff supply their thesis and dissertations and scholarly publications in a CD. At times, the CD-ROM drives of the computers used for the scanning are faulty. Some cases have occurred where the drives could not open, nor read the CDs provided.

Editing of works digitized

The scanning requires considerable editing to conform to the standard set for such materials. The library professionals involved in appending digital signatures on the scanned documents must have to edit the work by first checking the spellings, looking out for the bookmarks to ensure it is properly done.

Unavailability of needed Materials

Sometimes most materials that users consult in the library, especially theses and dissertations are shifted for digitization. Some cases arise whereby a material is not yet uploaded and at the same time not available in the library. This keeps the library users stranded if they are in desperate need of such materials.

Recommendation

The issue of copyright should be properly addressed. One good way of handling the issue of copyright is to obtain copyright permissions. In other words, copyright permissions have to be obtained to enable digitization of materials.

- User orientation
- On general occasions, staff, students, and researchers have come asking questions on how to access the uploaded materials. This is almost becoming a read-reference question to be developed by the librarians to whom these questions are posed. Therefore, it is crucial that an orientation be given to such users on how to locate the materials online
- The federal government should put more effort with regard to funding of tertiary institutions in Nigerians.
- Staff training;
- Constant training of staff is necessary
- As new technologies emerge, librarians and other staff involved in the digitization project should be trained. A train-the-trainer exercise should also be encouraged so that when some group of staff retire, other ones can pick up from where they dropped. This is also a way of addressing technophobia. Workshops should be organized at local and national

level by associations and interested bodies for the training of library personnel on digitization issues.

- Back-up sources should be provided
- In a case where a virus attacks the files and data is lost, there should be a back-up for continuity before then. An external hard drive of large storage capacity could be purchased for this purpose. The stated solutions will go a long way in strengthening the digitization exercise in the University libraries.

Conclusion

This paper has established that digitization is an essential task in modern day libraries. If a library is to live up to current challenges, it has to go digital, that is, provide access to digital materials. This will enable it to preserve endangered library resources, improve the efficiency of information search mechanisms and enhance access to library resources. It is essential for the library management to provide policy guidelines and articulate plans for the exercise. Digital library, otherwise known as virtual library, has grown to a special field of study. Courses of instruction and research opportunities are now made available in this area of specialization by some university.

References

- Bekele, S. (2002) The role and impact of the digital library on capacity building in the developing world- a case study of the OSSREA Digital Library, *International Information and Library Review*, 34, 129-137. Retrieved December 12, 2006 from <http://www.idealibrary.com>
- Broncolini, K. (2000) Selecting research collections for digitization: applying the Harvard model, *Library Trends*, 48, 783-98.34 *International Journal of African & African American Studies*.
- Fabunmi, B.A., Paris, M., & Fabunmi, M. (2006). Digitization of library resources: Challenges and implications for policy and planning. *International Journal of African & African American Studies*, 5(2), 23-36. V, No. 2, Jul 2006
- Graham, A. & Wroth, C. (2000) The www.virtual library: library resources. Retrieved Dec 9, 2006 from the Indiana University website: www.indiana.edu/~vlib/Digital_Libraries/
- Hirtle, P. B. (2002) The impact of digitization on special collections in libraries", *Libraries and Culture*, 37, 42-52
- Eke, Helen Nneka (2011). "Digitizing resources for University of Nigeria repository: Process and challenges." *Webology*, 8(1), Article 85. Available at: <http://www.webology.org/2011/v8n1/a85.html>
- McQuail, D (2000) *McQuail's Mass Communication Theory* (4th edition), Sage, London, pp. 16-34
- Pinnell-Stephens, J. (2005) "Alaska libraries rescue two native oral history resources", in *Saving the Stories*, *American Libraries*, pp.46-47.

Rosenberg, D. (2005). Towards the digital library: Findings of an investigation to establish the current status of university libraries in Africa. UK: INASP. Retrieved April 29, 2009, from <http://www.inasp.info/uploaded/documents/digital-libr-final-format-web.pdf>

Stefano, P. (2001) Selection for digital conversion in academic libraries, *College and Research Libraries*, 62, 58-69

Silkroad, I.T. (2009). Digitization service. Retrieved June 28, 2009, from [s
http://digitizationservice.com/?p=3](http://digitizationservice.com/?p=3)

Tsebe, J. (2005). Networking cultural heritage: Africa. World Library and Information Congress: 71th IFLA General Conference and Council, "Libraries - A voyage of discovery", August 14th-18th 2005, Oslo, Norway. Retrieved April 2, 2009, from <http://archive.ifla.org/IV/ifla71/papers/157c-Tsebe.pdf>