Digitization Best Practices in a Financially Challenged University Library: The BSU Experience

Lauren P. Kipaan
Benguet State University Library& Information Services
La Trinidad, Benguet, Philippines

Abstract. The paper describes the strategy for application and implementation of digitization in the financially challenged Benguet State University Library. It discusses best practices that enabled the university library to optimize resource utilization, and deliver high quality, value added services to its users.

Keywords: Digitization, Digital library, Benguet State University Digital Library, Digitization Process, GSDL, Digital Library Best Practices, Digitization strategies

Introduction

Digital libraries have received wide attention in the recent years allowing access to digital information sources in areas of education and other knowledge (Bainbridge, Thompson, & H. Witten 2003 as cited by Ambati, 2006). Libraries around the world aim to provide such services but they are sometimes limited by existing policies, structures, skills, and financial capability.

In 1999-2009, Benguet State University Library and Information Services (BSULIS) was a recipient of Philippine Inter-University Cooperation grants, a collaboration between Saint Louis University and Benguet State University (PIUC-SLU&BSU). The grant was a product of the collaboration of Flemish-Belgian universities, Vlaamse Interuniversitaire Raad (VLIR) that supports developing countries worldwide. Areas of concerns in the collaboration were the development of Information Communication and Technology (ICT), Research, Library development, and Faculty development (Tagarino, 2002). The grant enabled BSU Library to increase its book collections, implement and acquire integrated library systems, send librarians to international trainings, install electronic equipment such as closed-circuit television and alarm systems, and acquire servers and computers for library use(Vanhoute, 2010).These undertakings transformed the traditional library work and compelled BSU librarians to adapt to these technology-driven services.

The role of the library becomes more focused in making the users techno-savvy. Technological developments have affected not only the formats and sources of the information, but also how and where to provide library services. The library and its resources have partially moved to the virtual world of the Internet. As a result, library users can now access the resources remotely (Witten, 2010). In an effort to enable users in accessing the library resources via their computers, though the digitization was not given attention during PIUC-SLU&BSU cooperation, BSULIS smoothly expanded its services to include virtual reference and digital library. Technology now allows users to submit their queries to the library at any time from any place in the world and obtain what they need. However, in some instances, technological services like digitization projects in the library can't be fully implemented due to managerial and financial constraints that confront most state university libraries in the Philippines.

BSULIS occupy a place of pride in BSU community and it has been an essential instrument in achieving the university's outstanding researches and educational missions. It expedites learning that promotes discovery and advancement of knowledge. The mission of the library is to create an ICT model library through acquisition, organization, and dissemination of knowledge in all formats and in facilitating information access to all. Such is made possible by adapting and devising digitization processes and practices that are not only considered to be the best but are suitable to the existing setup and skills of the librarians.

This paper presents these processes and practices with an end view of guiding the management of other financially-challenged libraries towards improving their services via digitization projects. Its intent is to share to interested audiences the tips and coping mechanisms or digitization techniques towards providing digital library services to clientele even with a limited budget.

How it Began

One of the library staff, Mr. Lauren P. Kipaan, attended a series of seminar-workshops related to digitization. In 2004, his first skills training sponsored by Knowledge Spaceway, a commercial company, exposed him to advance equipment used in digitizing documents acquired through a huge budget. To further advance his skills and knowledge, he again attended a relevant seminar-workshop hosted by BSU in 2005which led him to realize that the best solution to provide a perfect Optical Character Recognition (OCR) is through manually encoding or getting the born-digital documents. In January 2009, he was given the opportunity to participate in the international training program on Scientific and Technological Information Management in Universities and Libraries- International University Cooperation (STIMIUC) in Belgium with funding from the Vlaamse Inteuniversitaire Raad-University Development Cooperation (VLIROUS). One of the training courses includes the Greenstone Digital Library (GSDL). His basic knowledge in metadata and manipulation motivated and enabled him to apply the concept to BSULIS where he works. In the early part of 2010, Mr. Paul Garcia, an ICT Division Chief of the same university, expressed interest in GSDL and helped Mr. Kipaan formulate a one-day training on GSDL which they implemented in June 2010 using the ICT Division funding.

The first digitization-related training was in BSU and was attended by different research units and colleges. This training was a jump start for the BSULIS. The ICT division, using an existing server, established the digital library site (digilib.bsu.edu.ph) after a month from the training. It is online with more than 100 titles uploaded in PDF and HTML formats. In December 13-17, 2010, Mr. Garcia attended the international forum on GSDL in Indian Institute of Management at Kozhikode, Kerala, India using the ICT Division funding. Mr. Garcia's attendance to the forum enabled him in customizing the GSDL.

Basically, the ICT division was fully in-charge of the technical aspects while the BSULIS is in-charge of the contents, including the processing into PDF/A or HTML format. As of August 2012, the Digital Library has a total collection of 713 titles in different sub-databases. Further, the Cariana, a sub-database on indigenous knowledge of the customized GSDL, was updated in contents through the collaboration of the Ifugao State University, University of the Cordilleras, University of Baguio, Benguet State University, ResearchMate Inc., and Tebtebba Foundation.

Objective and Purpose

According to Fabunmi (2006), digitization is an essential task in modern day libraries. If a library is to live up to the current challenges, it has to go digital, that is, to provide online services. A print collection isn't only threatened by human 'outsiders', sometimes, the problem is within: molds, moist, dust, insects, etc. (Vanhoute, 2010). Digitization preserves and provides easy access to information, and improves the efficiency for search mechanisms. BSULIS engages in digitization to make and give opportunities to born-digital and archival print documents of the university to be preserved and at the same time to provide online access.

BSU Experience and Challenge

According to Liu (2004), most of the academic libraries in the United States that implemented digitization are financially capable. A digitization initiative basically requires skills, planning, and budget among other needs. In the case of BSU, the key is the basic skills in digitization, existence of low-end computer units, good inter-unit relationship, and the commitment to put up a digital library. Before the implementation of Digital Library, Mr. Kipaan created and experimented on Wikis (carian.wikispaces.com) and google sites (site.google.com/site/bsulisarchive/home) as a way to provide users' access to thesis and dissertations bibliographic and abstract information.

The team. The Digital Library (DL) or Digitization team is composed of four (4) staff: the ICT Chief, a web developer from the ICT division, the library director and a computer technician assigned in the library. Occasionally, library science graduates volunteer to help in the scanning, encoding, and conversion of documents. Every second semester or summer break, BSU Information Technology students conducted their on-the-job training in the library to help in the digitization process. In most instances, library student assistants were tapped to encode archival documents. The quality of the finish product is ensured by close-monitoring, cautioning the encoders to go slow but sure and finally reviewing the output by the DL team.

The problem and solution. The primary problem of the DL team was on how to create and digitize content in an inexpensive process with minimal people involved. Digitization projects normally involve huge budget and significant number of human resources as the key factors in accomplishing them. However, in the BSU library,

its library director overcame those factors. Financial and other logistics are coordinated with the different departments. The ICT division sponsored the training and offered their extra server for digital library use, and the launching of Cariana, a sub-database produced through inter-university efforts, was sponsored by the University of Baguio Library.

DL team also optimized the use of the library's existing desktop computers and ordinary flatbed scanner for the processing of graphical images while texts are manually encoded for the print archival documents. The people involved do digitization works in addition to their main job assignments.

The BSU DL team did not undergo formal meetings. Constant casual conversations facilitated conceptualization and implementation of strategies and the installation of the server. Several trials were also done that indicated the next step and corrective measures to undertake until the system was made fully operational.

Submission policy. To save money and minimize personnel requirement in digitization, the submission of "born-digital" for theses, dissertations, and faculty manuals was proposed by the library unit in the Office of Vice Presidents for Academic Affairs (OVPAA) monthly meeting. The colleges approved the proposal and the submission practice is now institutionalized through advisory. The issuance of the advisory produced a significant impact in the collection and digitization process.

Copyright. The library supports the open access movement and the utilization of the open source software. It also aims to protect and uphold the rights of the authors or creators over their work. Hence, a copyright waiver through which the author categorically authorized the BSULIS to archive and provide public access to the document and information contributed was conceptualized and implemented. It has two options: first, the Close Community which allows the title, abstract, introduction, recommendation, conclusion and references only, and second, the Open Community which allows the whole contributed document or information for digitization and uploading online.

Lay out. Born-digital documents are arranged according to the approved sequencing of text as agreed verbally by the DL team. Before the document goes to the process of digitization, DL team will identify Close or Open Community waivers. For documents submitted for Open Community, all contents are processed following the sequence starting from Bibliography, Abstract, and Introduction down to the References. Using the MS Word as text editor, the footnoting, page numbering, and creating the table of contents are first part of laid-out. Then Acrobat software will be used to insert watermarks and to create bookmarks. The last process will be the reduction or compression of the files before saving as PDF/A. PDF/A is the archival version of PDF file format, an open source tool managed by International Standards Organization (ISO) that enables long-term digital preservation of electronic documents (Noonan, 2010).

The documents for Close Community mode are processed in the same manner as that of Open Community. In practice, any document exceeding 100 pages with numerous images are partitioned to make it easier for the researcher to download and open it. A suffix in the title indicating the part 1, part 2 and even part 3 of the document is provided.

Ownership and packaging details. The Submitted "born-digital document" is arranged first in MS Word format, checked and prepared based on the prescribed layout, then processed further in Adobe Acrobat. The final standard output will be that of a PDF or PDF/A. Metadata encoding and uploading to the GSDL server follows. The university library logo was used as the image water mark. This will remind the users that the work is owned or processed by BSU and that proper citation should be attributed correspondingly to BSU. In the case of scanned documents, it is processed immediately using the Adobe Acrobat software. The arrangement, optical character recognition, compression, and saving as PDF or PDF/A are done also in the Adobe Acrobat. Most of the scanned documents, particularly those predominantly of graphical images are not subjected for OCR. Unwanted marks and flaws in the scanned documents will serve as an authenticity indicator and remain due to the unavailability of software and equipment to delete them. Audio-Visual multimedia are encoded directly and uploaded in the GSDL server.

Workflow process. Figure 1 shows that in the BSULIS, submission of born-digital and printed archival documents are the top priorities for the digitization packaging. The copyright waiver is the pre-requisite before undergoing with the next process of digitization. The present Library Director, who is the initiator of the project, is in-charge of the selection and prioritization while the scanning, packaging, conversion, and compression are done by the Library Assistants, Volunteers and OJT's. The in-charge is assisted by the computer technician assigned in the Virtual Library section. Cautioning the people involved to "go slow but sure" and the quality

review are parts of the process to ensure accuracy; the latter is done before a backup is created. Dublin core sets of metadata were used to describe the digitized document and clustered according to courses and type of document. Before uploading to the server, a test is done to check the functionality. An additional back-up is created in an optical disc, together with the GSDL software so that it can run as stand-alone in all computer units. Uploading of data to the server is done manually with ease.
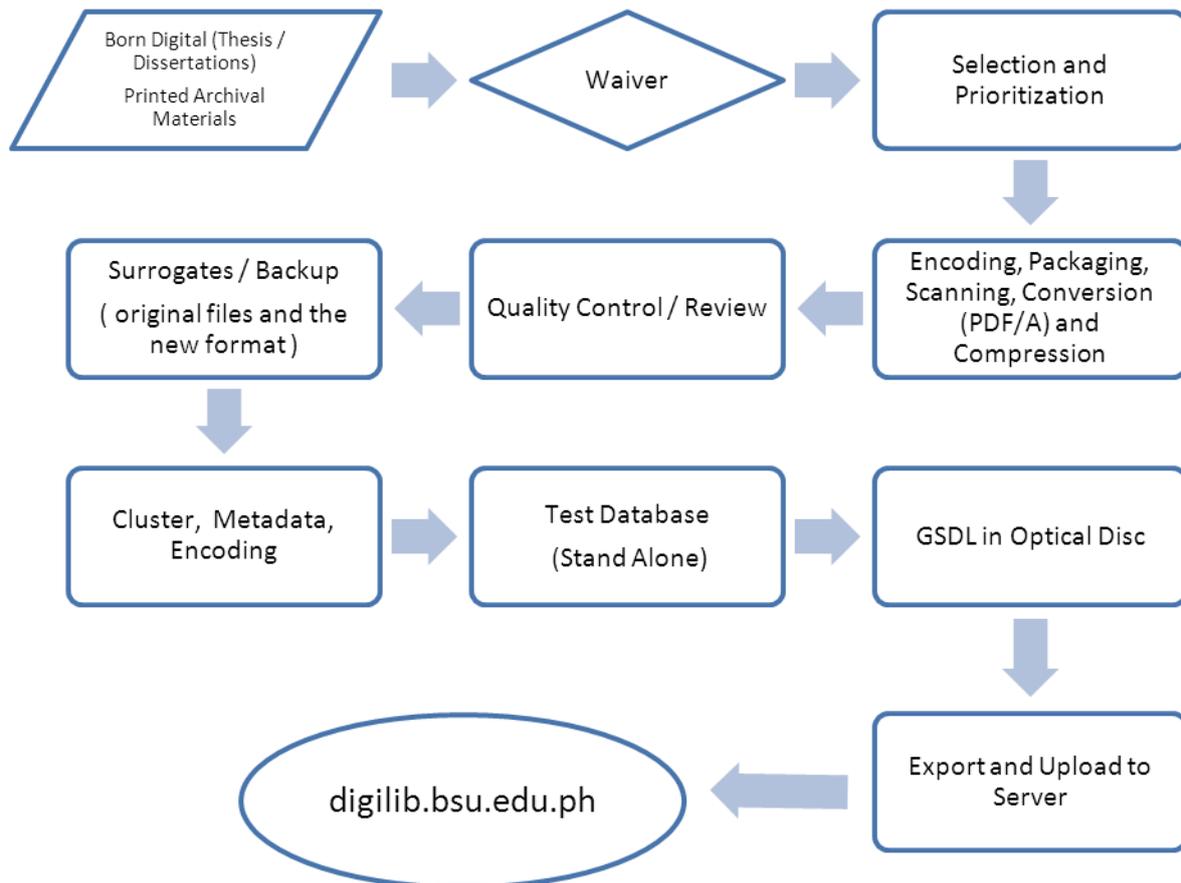

Figure 1. Digitization process in BSULIS

Platform. BSU uses the Greenstone. It is a 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 CD-ROM. Greenstone is produced by the New Zealand Digital Library Project at the University of Waikato, and developed and distributed in cooperation with UNESCO and the Human Info NGO. It is an open-source and multilingual software issued under the terms of the GNU General Public License. The aim of the Greenstone software is to empower users, particularly in universities, libraries, and other public service institutions, to build their own digital libraries. The complete Greenstone interface, and all documentation, is available in English, French, Spanish, Russian and Kazakh. Greenstone also has interfaces in many other languages (Greenstone Digital Library Software).

Greenstone runs on all versions of Windows, and Unix, and Mac OSX. It is very easy to install. For the default Windows installation, absolutely no configuration is necessary, and end-users can routinely install Greenstone in their personal laptops or workstations. Institutional users run it on their main web server where it inter-operates with standard web server software (e.g. Apache).

BSU used the 2.85 version of GSDL and experimenting on the Ubuntu 10.10 version as an operating system for the GSDL server. The server specifications are; HP Proliant ML150, Processor (Intel Xeon (R) CPU E5410 2.33Ghz, Memory (2.00 GB), Storage (80GB), and Graphics (Matrox G200e (Server Engines). It has only 6mbps bandwidth for the whole university though its internet connection is through leased line.

Database structure. BSU digital library follows the basic or common database design for GSDL in which one collection creates one database file. It has seven databases interconnected with the "cross-collection search" feature of GSDL. From the main page of the site, logos of each collection are displayed and linked to its contents. Digital library also served as a Document Management System (DMS) for some of the university issuances, though it is not visible from the main page but accessible from the university portal site. Table 1 shows the structure of visible databases classified either on subject field or type of document.

Item	Name of databases	No. of titles
1	Undergraduate Theses	579
2	Master Theses	69
3	Dissertations	7
4	Cariana	14
5	Faculty publications or manuals	2
6	BSU Research Journals	3
7	Records Archives	39
Total		713

Table 1. Structure of Databases

Keywords and metadata. The keywords assigned are supplied by the author or are taken from the content of the documents provided. Some are supplied from the subject headings available. The library encourages free language especially for the grassroots researches and works in which indigenous or native dialects are used. Dublin core set of metadata is automatically used since it is the most common standard for archiving platforms.

Promotion and feedback. Though technologies proliferated tremendously, it was noted that some promotional activities are still traditional due to digital divide and ease of reach.

The promotional initiatives or activities are done both in traditional and web-based by BSULIS. It is through the university newsletter and the social media like Facebook page, (groups.yahoo.com/group/bsulibrary), twitter (bengsulibrary), and the blog site (bsulis.wordpress.com) of the library. Desktop screens of all computer units in the internet laboratory or virtual library are customized providing noticeable icons for easy access or link to the digital library site. Colleges were informed of the ongoing digitization activity and all students, faculty and staff are invited to submit their research results for archiving and uploading online through a series of advisories. The library's Light-Emitting Diodes (LED) signage and outdoor banners were also utilized for announcing library activities including the promotion of BSULIS digitization initiatives.

Feedbacks are encouraged and BSULIS received them through social media, landline telephone, short messaging services (SMS), drop box, and as direct personal comments. Feedback through SMS is hosted by GLOBE, one of the country's leading mobile phone companies, and another dedicated two personal numbers that serve as real-time reference services of the BSULIS.

Current Status

Digitization initiative in the BSULIS is managed and continually updated for users free of charge. The task is headed by the library director assisted by a computer technician assigned in the printing and Virtual library or internet section, in addition to his frontline duties. Due to the in-availability of full-time staff to do the task, selection of archival materials for digitization depends solely from the decision of the library director and suggestions of the library staffs.

Figure 2 shows the statistical usage by country of BSU Digital Library based on Google Analytics report (March 1, 2012 to August 31, 2012). BSU Digital Library had a total of 6693 visits, 620 unique visitors (based on IP address), 2.94 pages/visit (Percentage of average pages visited), 00:01:05 average visit duration, 48.17% bounce rate, and 20.20 % new visits (Percentage of new visits based on IP address within the span of report generation). The top nine visiting countries are Philippines, USA, China, France, Canada, Hong Kong, Ukraine, India, and Japan. The Not set are the visits of other countries not specified.
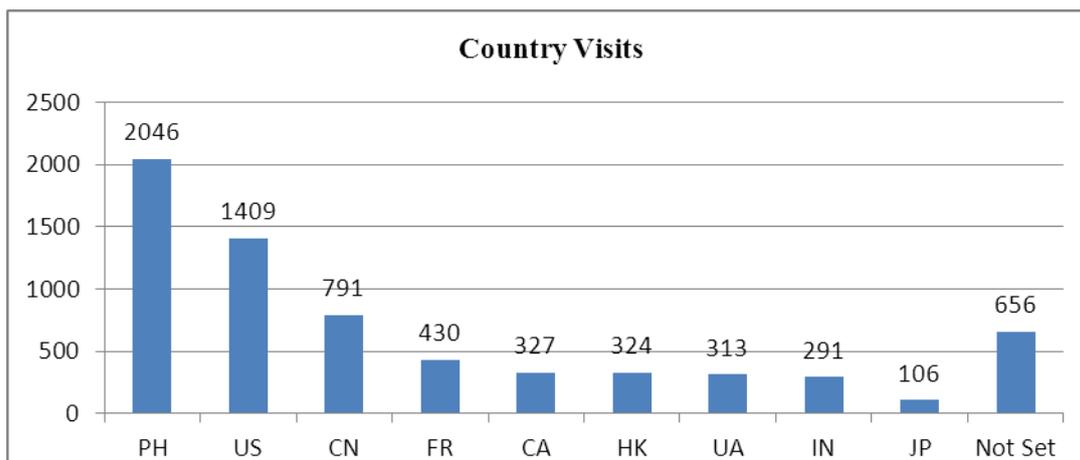
Figure 2.Statistical usage of BSU Digital Library.

Forthcoming Ideas

Review and strengthening of the DL team set-up for enhancing the processes and sourcing for funding are the main challenges for BSULIS. A university policy setting standard format and requiring the submission to the BSULIS the results of theses, dissertation, and other research works in the university, and giving a corresponding credit in the National Budget Circular (NBC 461- scheme for promotion), for faculty who submit his research results for uploading in the Digital library is proposed.

Migrating to a more superior and interactive platform is another consideration to ensure quality service and excellent user satisfaction. In order to comply with correct syntax, the existing web address is being revised in order to gain and generate more statistics or hits. The setting up of the realistic book, another presentation of book type submitted by faculty, is being considered to add as enticing feature in the interface. The idea as to the main feature of this presentation type is the opening and closing of book page by just clicking the mouse.
Tracking number of the document is also conceptualized to easily locate and track documents. In such case, if clients will demand for the original document, particularly those of the printed archival materials, this tracking number system will be used. A dedicated task force will be formed to formulate a procedural manual, fix lapses, improve and maintain the BSULIS digital library.

BSU DL Team is presently working for its registration to different directories such as the Directory of Open Access Repositories (DOAR), and to make it accessible through Google scholar search engine including its Open Archives Initiative Protocol for Metadata harvesting (OAI-PMH) compliance. BSULIS is currently working for the creation, through the ICT Division, of a dedicated bandwidth for all BSULIS online services to provide a faster retrieval access.

According to Alayon (2012), there are only three registered digital or repository site in the DOAR for Philippines. There are other digital libraries or repositories operating like BSU, Rizal Library Digital Repository, and University of the Philippines (UP)-University Archives and Records Depository, but are not registered. Table 2 shows the status of digital libraries or Institutional Repositories (IR) in the Philippines. It appears that BSULIS will be the first user of GSDL software in the Philippines if registered in the DOAR (University of Nottingham, n.d.).

Repository name	Country	Num. Recs.	Pubs	Confs	Theses	Unpub	Other	Base URL	Software
DSpace at IRRI	Philippines	1345			+		+		DSpace
Southeast Asian Fisheries Development Center, Aquaculture Department Institutional Repository	Philippines	970	+	+		+	+	OAI	DSpace
DSpace at Saint Louis University	Philippines	341			+		+		DSpace

Table 2. Number of Open Access IR and Software Used in the Philippines.

Conclusion

Collaboration and good public relation are essential in implementing digitization initiatives. Simplifying the application in creating processes to save time and financial resources, and resourceful staff committed for better service are the key elements. These elements play a dynamic part in the process of digitization and the development of digital libraries especially in a financially challenged library. Riding with the provision of free cloud applications and open source technologies enabled BSU Library to cope with the emerging trends in library services at a low and affordable cost.

Technology-based library services will become more common and sophisticated as the web becomes a public place throughout the world. As we actively enhance the library services, our main goal remains the same, to serve and teach users to find, evaluate, and use information effectively. The librarians should go on the road to prepare themselves go with the new communication models, a change from face to face interaction to human machine connection, from paper to digital delivery, from text-centered mode to multimedia and from physical presence to virtual presence.

Recommendation and Challenge

Libraries can still digitize by manually encoding print resources and utilizing free cloud applications on the web for access. For libraries without internet connection, standalone software can be installed with an ordinary scanner for digitizing rare images for its clients. As the time of librarians are consumed in the frontline services, digitization can be achieved by setting up support groups like library club, inviting volunteers, and/or on-the-job trainees, and by requiring researchers to submit born-digital and formatted documents.

References

Alayon, S. B. (May 28-31, 2012). The Development of an Institutional Repository at the Aquaculture Department of the Southeast Asian Fisheries Development Center, Philippines. *15th Congress of Southeast Asian Librarians (CONSAL) Meeting and General Conference.* Bali, Indonesia.

Ambati,Vamshi, N Balakrishnan, Raj Reddy, Lakshmi Pratha, and CV Jawahar (2006). The digital library of India project: process, policies, and Architecture. Retrieved August 26, 2012, from http://www.serc.iisc.ernet.in/~balki/papers/ICDL2006-Delhi.pdf

Fabunmi, B. A. (2006, July). Digitization of Library Resources:Challenges and Implications for Policy and Planning. *International Journal of African & African American Studies, V*(2). Retrieved August 29,2012, from https://ojcs.siue.edu/ojs/index.php/ijaaas/article/view/80

Liu, Y. Q. (2004). Best practices, standards and techniques for digitizing library materials:a snapshot of library digitization practices in the USA. *Online Information Review, 28* (5), 338-345. Retrieved August 26, 2012, from http://www.emeraldinsight.com/journals.htm?articleid=862266

Noonan, D. W. (2010). PDF/A: A Viable Addition to the Preservation Toolkit. *D-Lib Magazine, 16* (11/12). Retrieved August 25, 2012, from http://www.dlib.org/dlib/november10/noonan/11noonan.html

Tagarino, D. (2002, January-February). All About the PIUC. *PIUC Newsbytes* , 1(1), 2.

University of Nottingham.(n.d.).OpenDoar-The Directory of Open Access Repositories. Retrieved August 26, 2012, from http://www.opendoar.org

University of Waikato – New Zealand Digital Library Project.(n.d.).Greenstone Digital Library Software. Retrieved August 26, 2012, from http://www.greenstone.org

Vanhoute, W. (2010). Ten years of joint and parllel library development in BSUand SLU: an overview. *Proceedings of the International Conference on Best Practices in University Development through International Cooperation* (pp. 48-50). Baguio City

Witten, I. H. (2010). *How to build a digital library* (2nd ed.). Boston: Morgan Kaufmann Publishers.

Supplementary readings

Angeles, C. N. (n.d.). Standards and Procedure in Digitization and Digital Preservation [pdf]. Retrieved August 29, 2012, from http://paarl.wikispaces.com

Bhatnagar, A. (2005). Web-based Library Services. *3rd Convention PLANNER -2005.* Ahmedabad. Retrieved August 30, 2012 from http://ir.inflibnet.ac.in/dxml/bitstream/handle/1944/1418/58.pdf?sequence=1

Bist, R. S. (n.d.). Digitization progress and Challenges at Gandhi Smriti library of LBSNAA [ppt]. Retrieved August 26, 2012, from www.naclin.org/RajenderBistREVISED.ppt

Fleischhauer, C. (1996, January). *Steps in the Digitization Process*. Retrieved September 1, 2012, from Library of Congress: http://lcweb2.loc.goc/ammem/award/docs/stepsdig.html

Ramos, M. S. (April 18-20, 2012). Digitization for Access and Preservation:Role of Academic and Research Librarians[pdf]. *PAARL National Summer Conference.* Coron, Palawan, Philippines. Retrieved August 29, 2012, from http://paarl.wikispaces.com

Samaniego, C. C. (April 18-20, 2012). Digitization Projects for Libraries:Development and Management[pdf]. *PAARL National Summer Conference.* Coron, Palawan, Philippines. Retrieved August 29, 2012,from http://paarl.wikispaces.com

Verzosa, F. A. (April18-20, 2012). Copyright issues in a Digital Library Environment[pdf]. Coron, Palawan, Philippines: PAARL. Retrieved August 25, 2012,   from http://paarl.wikispaces.com

Acknowledgements