Digital Curation Projects: A Study of Selected Academic and Research Repositories in the Philippines

Lourdes T. David
Professional Regulatory Board for Librarians, Professional Regulation Commission, Manila, Philippines
lourdesdav@gmail.com

Stephen B. Alayon
Aquaculture Department, Southeast Asian Fisheries Development Center, Iloilo, Philippines
stpnalaun@yahoo.com

ABSTRACT
The paper examines digitization and preservation initiatives in selected academic and research libraries in the Philippines. It identifies the drivers for the projects, practices followed, and issues encountered. Most of the projects were initiated by the librarians in cooperation with information and communication technology personnel. The selection of the materials for digitization was based on demand, the condition of the print format, and importance of the materials for future use. Conversion into microfilm from digital format, use of Internet mirror sites, and off-site repositories of data on hard disks were some of the preservation methods practiced by libraries. Challenges faced by digital curation activities include the need to develop digital curation skills of information professionals and researchers, and to develop the information infrastructure of the institutions and the country.

INTRODUCTION
This paper provides an overview and describes digitization projects and preservation initiatives in selected academic and research libraries in the Philippines. Particularly, it describes a digitization project at the Rizal Library of the Ateneo de Manila University, to digitize its newspapers and special materials, specifically photographs. The digitization project is a case study based on firsthand experience of the first author who at the time was the director of the library. The survey of other digitization initiatives is based on reports and publications, and websites gleaned from the Internet.

THE RIZAL LIBRARY EXPERIENCE
Digitizing Newspapers
The digitization project was conceived in 2001 upon the request of users of the product Newspapers on Microfilm. The users were having problems maintaining the quality of the microfilms because of the expense of storing them in a room with 24 hour air-conditioning. Many of the users found that their microfilms could no longer be used after some time...
because they were stored in rooms where the temperature and humidity were not controlled. In the Philippines, the room temperature varies from 25 to 40 degrees Celsius, depending on the season and the location of the library. Relative humidities were also in the high 70s or 80s. There was also the high cost of microfilm readers and printers. The cost of printing microfilm images was also high compared to photocopying. They suggested digitizing the newspapers since all libraries in the Philippines have computers with DVD drives. The DVD was the preferred format because not all libraries are connected to the Internet.

There were other objectives in digitizing newspapers and photographs. One was that the handling of microfilms was unwieldy compared with the DVD. There was always a need to assist the users in loading the film and in printing the images. A second reason was ease of access. Information in digital format can be indexed and accessed randomly while those on microfilms were accessed sequentially. A third reason was the ease in saving the information to thumbdrives or sending it to one’s email address. A fourth reason was the ease of duplication should the disk be damaged due to heavy use.

The Ateneo de Manila University has been microfilming newspapers since the 1980s to preserve and share its content with users all over the country (David, 2002). In a country such as the Philippines however, temperature and humidity, the natural enemies of microfilm, are difficult to control. Hence when the requests to digitize came in, the Rizal Library in cooperation with the Philippine Library Materials Project, the agency handling the distribution of the microfilms, conducted a survey with current microfilm users and potential DVD users on their preference in acquiring newspapers for research purposes. All respondents preferred to acquire newspapers in digital format.

In 2002, a team was formed to study equipment costs, administrative costs, and costs of the creation of digital objects from print, conversion to microfilm and vice versa, and what and where to acquire the equipment needed. After a thorough study of the above factors and conducting a feasibility and market study, a pilot study using four newspapers was done. Digital objects were created and stored in hard disks and DVDs. The first batches of the DVDs were sent out to users for evaluation and comment. The DVDs were also subjected to testing by users on-site. The results were collated and used to improve the product which included an index for fast retrieval. After a year, another survey was conducted to determine user satisfaction and other problems met by the users of the product. The survey revealed that the users were very satisfied with the product. The production of newspapers on DVD is now a regular function of the Rizal Library photoduplication section (Figure 1). The DVD is still the medium used to this day as not all libraries in the Philippines have access to the Internet. The newspapers can be browsed by date, searched by keywords, cropped, printed, saved or sent by email. The software used in indexing, cropping, printing, saving or emailing was developed by computer programmers at the Rizal Library.

**Photographs Digital Archives**

The success of the *Newspapers on DVD* project led to another project, the digitization of photographs. The Rizal Library has approximately 250,000 photographs in its collection. This includes 200,000 black and white photographs from the photo morgue of the magazine *Mr. and Ms.* The photographs date from 1945 to 1972, which was the active life of the magazine. The photographs have been used by researchers on Philippine history and culture.

The American collection, a special collection covering the American colonial years in the country has 17, 408 photographs, 9,240 of which are negatives. Digitization of the photographs in the American Historical Collection started on April 17, 2002, and was
completed on May 20, 2003, while the digitization of photographs in the Mr. and Ms. photo morgue is ongoing.

Other special collections in the library contain rare photographs of Dr. Jose Rizal and other national heroes. These will eventually be digitized. The software that will ensure data security when they are uploaded to the Internet is still under development.

**Issues and Concerns**

The factors to consider when embarking on a digitization project are shown in Figure 2. This was the procedure followed in the digitization of the Mr. and Ms. collection. In addition, some issues and concerns were raised before the commencement of the project. These were:

1. **Technology.** These included continuing technological developments, image formats, compression schemes, network transmission, monitor and print design, computing capacity, image processing capabilities, particularly for the automatic extraction of metadata (OCR) and visual searching (QBIC) (Chapman & Kenney, 1996). Questions such as: will the media last; will the media's hardware (reading device) and its controlling software remain available; does the compression technique involve risks; will the logical formatting of the data be clear to future generations; is the technology available and affordable? The scanner output setting used is shown in Table 1.

2. **Retrieval software and indexing software.** The ICT team of the library wrote a program for the indexing and retrieval of information.
3. **Policy and legal issues (copyright, fair use, etc.).** The publishers of the newspapers and the photographers were requested for permission to digitize their publication. They graciously granted the library permission to digitize and distribute the products in digital format.

4. **Criteria for the selection of materials.** The following criteria were observed in the selection of materials to be digitized: Intellectual content, heritage value, research value, user demand (market value), and quality of the original material.

5. **Start-up costs, cost recovery/sustainability.** A market survey was conducted to determine the demand and matched with start-up cost and return-on-investment.

6. **Expertise/staff training required.** Training was one of the conditions stipulated in the purchase of the equipment. Three pieces of equipment were purchased. The first is for digitizing from print, the second was for digitizing from microfilm, while the third was for converting from digital format to microfilm.

   To date both projects have been very successful in providing access to information and in preserving content. The original photographs in the American Historical Collection and the Mr. and Ms. photo morgue are no longer made available to the public, unless the researcher wants a higher resolution for purposes of publication. Other digitization projects have followed, and access is shifting from onsite access to remote access. The theses digital
OTHER DIGITIZATION INITIATIVES IN THE PHILIPPINES

According to Batiancila (2010), early digitization efforts in the country began in the late 1990’s. Selected Filipiniana materials were digitized through a project funded by the United Board for Higher Education and continued by the De la Salle University (Versoza, 2005). The pioneers of digitization in the country are the National Archives, The National Library of the Philippines, and the libraries of De la Salle University, University of the Philippines, Ateneo de Manila University and the Supreme Court.

1. The National Archives of the Philippines started its digitization projects of historical documents in 1998 (Darlucio, 2009). More than 13 million documents have been digitized but have not been made accessible on the Web. Moreover, the digitization project is done only to improve access to the documents but microfilming is still the focus of preservation efforts. The vision is for the National Archives website to be the focal point of access to all government records. The digitization project of the National Archives is done in partnership with the Integrated Government Philippines Project through the National Archives and Records Management Information Systems Program.

2. The Philippine eLibrary project is a collaborative project of the National Library of the Philippines, University of the Philippines, Department of Science and Technology, Department of Agriculture and the Commission on Higher Education (Philippine eLib, n.d.). The project sought to make available in digital format and through the World Wide Web, the holdings of the participating agencies which is approximately 24 million images. Digitization was outsourced and librarians checked each image for quality control. The librarians found many unreadable pages, too large output files, and damage to the original documents. They also found that the time spent on checking for quality was longer than rescanning in-house to ensure quality (Quiros, 2008).

3. The National Library of the Philippines digitized its rare books, maps, photographs; Presidential papers and speeches; and other materials that are of cultural and historical value as part of the eLib project (National Library of the Philippines, 2007). Its aim in digitization was two-fold: to increase access to Filipiniana materials and to preserve the original material (Quiros, 2009).

4. The Supreme Court developed an e-library which is a compilation of the sources of laws, statues and jurisprudences in the country (Ong, 2014). It aims to be the primary research tool in the speedy delivery of justice, serve the entire judiciary and support legal education. The Supreme Court e-library is accessible at http://elibrary.judiciary.gov.ph/.

5. The Department of Science and Technology Science and Technology Information Institute (DOST-STII) developed STARBOOKS, the first Philippine science digital library which contains thousands of science and technology resources stored in “pods” with a user-friendly interface. The project aimed to make available to the general public the more than 150,000 volumes of information materials held at the DOST-STII without the need for an Internet connection. The projects received a special citation from the...

6. *The University of Santo Tomas Miguel de Benavides Library*, being the library of the oldest university and the only pontifical library, has a collection of rare materials dating back to the Spanish era. With a grant from a bank in 2011, it was able to embark on an ambitious three-phase program, the conservation, digitization and publications of its collections. The project, Lumina Pandit, aimed to conserve the rich heritage collection in the library and archives, publish its historical collection catalogues, digitize its resources and upload them on the World Wide Web (University of Santo Tomas Miguel de Benavides Library, n.d.).

7. *Filipiniana.net* was conceived in 2005 by Gaspar A. Vibal and Jaime Marco. It went online in September 2006. It has a growing collection of textual and non-textual materials. It allows the user free online access to the entire book which is fully indexed by keyword in English, Spanish and Filipino languages. It also established the wikipilipinas.org knowledge portal (Wikipilipinas, n.d).

8. *The Philippine e-journals* is an online product produced by C&E Publishing in partnership with educational institutions in the country. It contains academic journals published by the different academic institutions in the country in compliance with a requirement of the Commission on Higher Education for faculty and academic institutions to publish (Philippine E-Journals, n.d).

9. The digitization project of SEAFDEC/AQD was carried out to develop an institutional repository where SEAFDEC/AQD researchers can store, preserve and share their research publications to enable access and increase visibility. Alayon, Nemiz, Superio, de la Peña, and Pacino (2013) described their experiences in digitizing institutional publications and developing an institutional repository.

10. Other libraries and research centers with institutional repositories include Saint Louis University Library, the International Rice Research Institute Library, World Health Organization, Asian Development Bank. Saint Louis University Library pioneered the use of DSpace in the Philippines. Perez, Angeles and Alayon (2013) presented various digitization initiatives and digital repositories of selected libraries, archives and museums. These are the Ateneo Professional Schools Library, House of Representatives Archives and Museum, Lopez Memorial Library and Museum, Lyceum of the Philippines University-Jose P. Laurel Digital Collection, Manila Observatory Library and Archives, Official Gazette Digitization Project, Southeast Asian Fisheries Development Center/Aquaculture Department Library, University of the East Ramon Magsaysay Memorial Medical Center, Inc. Library, University of the Philippines-Center for Ethnomusicology, and the University of the Philippines-University Archives and Records Depository. The types of materials digitized by these institutions are presented in Table 2.
Table 2. Types of materials digitized in selected libraries, archives and museums in the Philippines

<table>
<thead>
<tr>
<th>Institution</th>
<th>Books/ Monographs</th>
<th>Manuscripts</th>
<th>Company Records</th>
<th>Audio- Visual</th>
<th>Theses/ Dissertations</th>
<th>Published works</th>
<th>Correspondences</th>
<th>Unpublished works</th>
<th>Photos</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ateneo Professional Schools Library</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>House of Representatives Archives and Museum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lopez Memorial Library and Museum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lyceum of the Philippines University-Jose P. Laurel Digital Collection</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manila Observatory Library and Archives</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Official Gazette Digitization Project</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southeast Asian Fisheries Development Center/Aquaculture Department Library</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of the East Ramon Magsaysay Memorial Medical Center</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of the Philippines-Center for Ethnomusicology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of the Philippines-University Archives and Records Depository</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Perez et al (2013)
PROBLEMS AND BEST PRACTICES

The common difficulties encountered by the institutions that have embarked on digitization are:

1. **Employing a commercial service in the digitization process** raised a major problem in the Philippine eLib project and projects of other government agencies. The images had to be reviewed and returned for retake because there were many blank, incomplete or omitted pages. It was found that the editing or quality assurance procedure took longer than if library staff had shot the images themselves. For the eLib project, this was the major reason why outsourcing the project was stopped in 2007. The best practice in this case is to “Do it right the first time.” Quality assurance should be ensured in outsourcing digitization.

2. **Return on investment (ROI) or sustainability.** Many of the projects did not include a market study and return on investment study. The market study will identify potential users and long term sustainability of the project. The Ateneo de Manila University did not find this a problem because prior to the commencement of the project, a feasibility and market study was conducted to determine start-up costs, market demand and return on investment. In case of digital contents in open access repositories, impact could be measured through increased visibility and citations of institutional publications. To ensure sustainability of the project, an inventory of all potential publications for digitization and uploading in the repository is necessary. Digitization is time consuming and imposes demands on staff time and workload. Identifying time frame, targets, workflow is necessary. Strong collaboration and support from the management should be established.

3. **Start-up costs which included major investments in technology and staff training** were a problem for most of the projects. Luckily, most of the projects were funded by external agencies through loans or grants. The Spanish government was a major supporter of the digitization projects since many of the materials to be digitized were written in the Spanish language. The best practice for this problem is to conduct a feasibility study and market study to ensure a return on investment in a short period of time. Submit a project proposal to an agency that might potentially be interested due to the relevance of the content of the materials to the agency. The results of the market survey and feasibility study providing evidence of sustainability will help in getting a positive response from the funding agency.

4. **Lack of trained staff.** The training of staff should be part of the negotiations with the supplier of the equipment. Such training should be “echoed” to other members of the team, and should continue during and after the lifecycle of the project when it becomes a regular function of the library.

5. **Space.** The project team and the equipment should be given adequate space, which may be their permanent office space when the project is successfully implemented and becomes a regular function of the library. The space should have the necessary equipment and facilities for long term use.

6. **Compliance with the Open Archives Initiative (OAI).** According to Alayon et al. (2013), some of the digital libraries and institutional repositories in the country were not OAI compliant. Training of personnel in this regard should be addressed.
CONCLUSION

The study found that most of the digitalization projects were initiated by the librarians in cooperation with ICT personnel and funding agencies. The selection of the materials for digitization was based on demand, the condition of the print format, and importance of the materials for future use. The results indicate that digitization was practiced primarily for better access to content and preservation of the original print format. Conversion into microfilm from digital format, use of Internet mirror sites, and off-site repositories of data on hard disks were some of the preservation methods practiced by libraries. Different platforms were used by the different libraries for various reasons. Some are not OAI compliant.

Challenges faced by digital curation activities include the need to develop the digital curation skills of information professionals and researchers, and to develop the information infrastructure of the institutions and the country.

The study provides a profile of digitization initiatives in the Philippines. It has identified best practices and problems encountered in undertaking the projects, which can provide a basis for making decisions for future digitization projects of other libraries.

REFERENCES


