Contents

WDL (World Digital Libraries)
Volume 5, Issue 1 • June 2012

v Editorial

1 An Efficient Approach for E-Content Management and Delivery in Digital Library
Achintya K Mandal and Utpal Roy

11 Anwesan: A Search Engine for Bengali Literary Works
Suprabhat Das, Shibabroto Banerjee and Pabitra Mitra

19 Adding Value to an Abstracting and Indexing System: The Case of MyAIS, Malaysia
Zainab A N

37 Collection Management of E-Books: A Developing Country Perspective
Abdul Waheed

51 Designing and Evaluating an Integrated Digital Library System for the National Oil Corporation in Libya
Ramadan Elaiess

75 Conservation and Preservation of Archives: A Case Study of Punjab Digital Library
Balbir Kaur Pandher

85 News

87 Index
Editorial

MICHAEL SEADLE
Dean, Faculty of Arts and Director, Berlin School of Library and Information Science and Editor, Library Hi-Tech, Germany

E-content management, search engines for literature, abstracting and indexing, collection management, designing and evaluating, and conservation and preservation – the topics in this issue of World Digital Libraries reflects the range of tasks and challenges that libraries face in implementing the modern digital library, and in that sense the journal serves to be an excellent primer for libraries with an earnest interest in digital library development.

This journal operates on the assumption that digital libraries are now an integral feature of the library landscape and should not be conceived and built merely as boutique projects for a very limited and specialized range of materials. The scale of digital projects has been growing, not just because of local digitization projects, but because commercial publishers are providing many more digital journals and monographs. There is good evidence that British scholars rely primarily on electronic sources for their scholarly research (Tenopir, 2011)*. Britain is not alone or unique in this venture. It puts pressure on libraries to expand their digital full-text content and encourages publishers to favour digital formats. Journals have gone further in this direction than monographs, partly because centralized journal ownership of copyrights has made it easier to avoid legal complications when digitizing, while author ownership of the copyrights in monographs has made the permission process administratively harder. Nonetheless, new eReading devices and screen-friendly formats have given an impetus to digital publication of monographs as well. Digital libraries that a decade or so ago had only a few hundred items can now have contents in the same or greater scale than the number of physical items in traditional library buildings.

The technical and administrative complications for management grow with the scale of digital contents. Today, the boundary between traditional publisher content in digital formats and the digitized content that libraries themselves provide is an artificial divide based on licensing and legal issues. Overcoming that divide is not always easy, especially when the publisher content remains on the publisher server. Efforts to do this matter because readers care chiefly about finding and being able to use content that answers their information seeking needs. Discovery tools such as indexing are of key importance here, but the solutions are not purely technical. E-content management involves integrating digital content into the library as a whole and the way in which libraries present

digital content can make a significant difference to its usability.

The digital preservation issues are particularly important and complicated. The issue today is not simply that storage media will fail or that formats will change. Digital content presents questions in areas like integrity and authenticity that are significantly different than the comparable issues in the preservation of traditional physical materials. Digitization itself is also a widely recognized way to preserve fragile physical content, when the digitization of fragile materials becomes an integral part of a collection development plan. Digitization need not mean disposing of physical content. Digitization actually helps to preserve fragile content by providing access while the original object can be protected in closed and environmentally-friendly locations.

Digital libraries are growing in importance. The speed of this growth varies from country to country and depends to some extent on the communications infrastructure, without which digital access is not possible. Digital libraries also require staff with appropriate technical training. The number of possible discussion and research topics is legion and World Digital Libraries will continue to cover them.
An Efficient Approach for E-Content Management and Delivery in Digital Library

Achintya K Mandal
Deputy Registrar, Assam University, Silchar - 788011, INDIA
Email: achintyak99@hotmail.com

Utpal Roy
Department of Computer and System Sciences, Visva-Bharati University, Santiniketan, INDIA,
Email: utpal.roy@visva-bharati.ac.in

Abstract
In a digital library environment, the delivery and management of e-content is still a challenging task. The present study addresses an easy-to-implement technique applicable for the management of e-content and its delivery in digital library. The experiment has been performed upon some paper-based derived digital documents of many variants. The idea behind it is that the non-text region of a book-page, mainly the white margin around four sides and the space between two paragraphs have been discarded during scanning and then the conventional algorithm has been applied for removing the obvious noise incurred during scanning process. In essence, a reasonable amount of storage space is saved when such a document is placed in the secondary memory of the digital library server. Obviously, the web request for such pages is served in a better speed than usual. Moreover, a margin resetting algorithm in client side is applied to return these pages in its original form. Considering, mainly the storage space and the access speed while serving web request, the obtained results have some positive implications towards the problem of e-content management and its delivery in digital library system.

Keywords: Digital Library, Content Management, Content Delivery, Derived Digital Document and Document Images.
1. Introduction

The digital library today uses wealth of different document formats storing and representing the content. The genesis and development of digital library is a long saga and has presently become a very common source for information seekers. Digital libraries come in many forms. They may contain text of the document instead of only simple metadata or catalogues of bibliographical information. They can also contain images, audio, or multimedia materials. All this information may be available in different formats, created with different softwares. The resources may reside on different servers using no unified thesauri or heterogeneous indexing schemes. All this makes information retrieval a very complex process. Every information system is unique when it comes to retrieval methods, and it is more or less necessary to have a fair idea of the characteristic features of each system to be able to perform a relevant search. This becomes even more complex when some digital libraries allow users to conduct search across a range of distributed services as described by Chowdhury and Chowdhury (2003). With many general issues of digital object storage, moderately large and secured storage system is an important and challenging aspect of modern day digital libraries. Storing of digital objects leads to a large byte-count. An institution’s digital repository (which may have a label such as ‘library’ or ‘archive’) may include terabytes of data spread across thousands of objects in hundreds of formats. It may, for ease of handling and control, transfer the contents to a more readily managed medium. The sheer ease of producing and proliferating electronic files, combined with the perception that storage is so cheap, leads to an uncontrolled explosion in numbers and storage volumes, erratic and unplanned use of storage devices. The ‘Total Cost of Ownership’ (TCO) of storage (which includes the life-long cost of its management) varies from organization to organization, but can only be minimized if the techniques for minimizing the size of the content are developed. Content management in digital library has been discussed by many researchers. Wu and Liu (2001) discussed Internet-based e-content management. Specifically, they reviewed the technologies, the criteria, the issues and concerns in content repository, content contribution, workflow, automation services, and lifecycle of automation services for controlling managing content and processes. Warren and Alsmeyer (2005) describe the application of semantic knowledge technology to a case study in intelligent content management, specifically the BT digital library.

In this context, the Trove (launched in 2009) is an important discovery by Australians for their libraries. It harvests metadata from over 1,000 Australian libraries and other cultural heritage organizations, allowing free public access to over 100 million items. The guiding principle of Trove is ‘Find and Get’ (Holley, R, 2010). The details of its application and development has been described by Holley, Rose (Holley, Rose 2011) Most of the libraries have a good collection of old and rare books, as there is no electronic equivalent of those collections. The only way to integrate them into digital library is by scanning the individual pages (document images) with a good quality scanner preferably by a planetary scanner. The file size of the document images is very large in comparison to the electronic documents, though it depends on the resolution of the scanning and the image file format chosen for saving. In normal practice, the document images are not properly processed before sending for final use of the service. Optical Character Recognition (OCR) software plays an important role here to convert the document images into its electronic equivalent. As a result the file size reduces to a great extent and all the words of the document become electronically searchable. Several OCR software for European languages particularly in English Nagy (2000), is available, that can also handle multi-columns, images, tables, and also preserve the document layout analysis. On the other hand, they work
on document analysis and recognition in other languages particularly Indian languages are still in a nascent form. However, a few basic OCR modules are available. Most of them, from north Indian Brahmi based scripts – such as Devnagari (Hindi), Bengali (Bangla) as discussed by Chaudhuri, Garain and Mitra. Ray Chaudhuri, Mandal and Chaudhuri, (2002) have developed a page layout analyzer that can locate the textual zones in multi-column and multi-font Indian documents in Devnagari and Bengali scripts, but are not friendly with images and tables. On the other hand, in real time work (especially for the old and rare books), we noticed that the output of the OCR is also not very surprising and highly depends on the quality of the documents and is not robust in font variations. In such circumstances, the only way to give effective service to the requesting user is to send the scanned document images directly from the server via intranet and/or Internet.

The article addresses a useful technique that has much impact on electronic content management and content delivery system, an important issue of modern day digital library system of any kind. The technique presented here is easy to implement. The main idea is that, the non-text portion, say, of a book-page, such as four-side margins, space between two paragraphs and alike blank/white regions have been properly marked and discarded during the preparation phase of a derived digital document. Thus, it saves a reasonable amount of storage space in any means. A simple as well as conventional noise cleaning algorithm is used to make the ultimate image noise free. Further, when the ultimate document images are noise free it is prominent and easy to read. In essence, the technique saves a reasonable amount of storage space when in the server side of the digital library. The digital obsolesce and the management of storage space for digital library has become a burning issue as the electronic documents are going to be doubled in every six months. The detailed discussion about it is beyond the scope of this study. Obviously, the web request for such pages is served in better speed than usual process. Moreover, a margin resetting algorithm at client side has been applied to return those pages in its original form. This helps the reader to read the web document having usual margin. Sometimes, if required, the reader may set the margin according to his convenience. This option is available in the margin setting algorithm. Considering, mainly the storage space and the access speed while serving web request the obtained results have some positive implications. This indicates that this technique is much viable when a derived digital document is served through web and provides a simple probable solution towards the problem of e-content management and its delivery in digital library system. This experiment has been performed with many variants of derived digital documents and the results obtained are very much impressive. The results, point out that the conventional practice for scanning the paper document and the practice of storing the raw document pages is a pedestrian approach though the practice is still going on.

This paper is organized in following five sections. Section 2 describes the image acquisition and noise cleaning techniques. The storage optimization technique to save space in the server is presented in section 3. Some experimental results and discussion with various types of input images are shown in section 4. Finally section 5 embodies conclusion of the study.

2. Image Acquisition and Noise Cleaning

The first step of a digital library is the content collection/creation. Document image collection includes image acquisition that is, to scan a document with appropriate scanning instrument. The purpose requires an imaging sensor and the capability to digitize the signal produced by the sensor. Digitization can be done either by a flatbed scanner or a hand-held scanner. The resolution of a flatbed scanner varies from 100
to 900 dots per inch (dpi) where as the hand-held scanners typically have a lower resolution range. For the present study, the flatbed scanners are appropriate so that a complete document page can be digitized at a time, though flatbed scanners face some difficulties. Flatbed scanners often come in contact with at least that part of the object that has to be scanned. They also require books to be fully opened most of the time. Both practices can damage rare books; for example, opening a book at 180 degrees can damage its spine. A planetary scanner (also called an orbital scanner) is a type of image scanner used for making scans of rare books and other easily damaged documents. In essence, such a scanner has a mounted camera which takes photos of a well-lit environment. Originally, such scanners were expensive and could only be found in archives and museums. This system is equally applicable for the scanned document images from the planetary scanner also. All inputs of the system are as uncompressed gray scale TIFF (Tagged Image File Format) file format and scanned with 300 dpi. Figure 1 shows the paper-based text and digitized text shown in XV image viewer software.

In the TIFF header, number of rows are available in the value field of the directory entry of Tag Value 257; number of columns are available in the value field of the directory entry of Tag Value 256, and Strip Offset from the directory entry of Tag Value 273. We use Tag Values 257 and 256 for number of rows and for number of columns but for Strip Offset = file size – (rows * columns). From Strip Offset position, every character in a 2D matrix, it is the gray-label matrix of that image.

2.1 Noise Cleaning

Binarization is a process for converting a gray-tone image to a two-tone image or binary image. In binary image, the pixels are either black (usually having value 1) or white (usually having value 0). Alternatively, the pixel or point \((r,c)\) with value 1 is called object pixel and the pixel with value 0 is called background pixel.

Thresholding is the easiest method to binarize an image. Check every pixel if it is greater than or equal to a particular value \(T\), then set that pixel with 0 else with 1. The value \(T\) is called a threshold value. \(B(x,y) = 0\) if \(G(x,y) > T\) and 1 if \(G(x,y) < T\). Where \(B(x,y)\) is the pixel in the binary image and \(G(x,y)\) is the pixel in the gray-label image. Threshold selection is an important task for binarization. One popular method is using histogram. In bimodal histogram, select the bottom of the valley between peaks of the histogram. The histogram with two peaks is termed as bimodal. Here, we assume that the images are bimodal, we use threshold value 128 in our binarization module. Though, there is a scope to change the threshold value. Figure 2 shows the binarized image of the image shown in Figure 1(b). In this two-tone image, the noises due to shadow is automatically removed as it has a gray value greater than the threshold value.

![Figure 1](image1.png) (a) Paper based text (b) Digitized text shown in XV

![Figure 2](image2.png) Binarized image
3. Storage Optimization

A close look at any book or a printed document will reveal that almost 30%-50% of the total page size of any book are blank mainly for the margins setup on the four sides of a page of the book and more accurately some half and blank pages also increase the percentage of the blank space available normally in all books. In other words, on an average, 50%-70% part of the page area contains valid text. Our objective is to deal with the textual portion without changing the original arrangement of the page setup and store it into the server so that the text portion can be retrieved as and when required. It also reduces the document image size as it excludes the four margins. As a case study, we considered a rarely available book, ‘Rabindranath Tagore His Mind and Art’ authored by B S Chakraborty, Young India Publications, New Delhi-48. The book contains 303 pages. Arbitrarily, a normal page was chosen from the book and scanned with 300dpi in a flatbed scanner. Figure 3a depicts the document image of the page with full margin whereas Figure 3b presents the document image of the same page with the text part only, without four margins. The file sizes of the two document images with (Figure 3a) and without margin (Figure 3b) are respectively 3894 KB and 2636 KB and eventually the size difference is 1258 KB. So, for every single full page one can save 1258 KB memory space. Similarly, to store 303 pages of the entire book one can save 381 MB storage space even less as normally all the books have some half and quarter pages, as described above.

But in normal convention readers are not habituated to read a book, either in the hard or soft form, without margins. Our system has been designed in such a way that the margins of the pages can be computed according to the choice.
of the reader and the reader can easily put the margin size to view the actual text according to one's need. The document is re-written in a new form after noise removal as discussed in section 3. The noise removal algorithm has been applied in the document images of Figure 3b and the outputs have been shown in Figures 4a and 4b with margins 1 inch and 1.5 inches respectively. The retrieval algorithm also works a bit faster than usual as its performance depends mainly on the file size. As a result, the electronic document serves the requests through the network in a better speed and in a more readable form. The program for re-writing the images with the desired margin may reside at server-side or in the client-side. Though, for both the cases there are some advantages and disadvantages and our algorithm is equally applicable for both the cases and the choice between the two is a web-designing issue only.

4. Experimental Results
An efficient as well as easy-to-implement technique has been proposed in this study for the management and delivery of e-content in the context of digital library. The proposed method has been applied with many variants of document images having different languages, font size and documents having possibility of containing considerable amount of noise due to old age and careless holding of the original documents. Sample of original scanned images and the obtained results have been presented in the Figures 5a, 5b, Figures 6a, 6b and Figures 7a, 7b respectively and the detailed analysis of

Figure 4a The output image with 1” margin. Figure4b The output image with 1.5” margin.
the results are presented in Table 1. Figure 5a is the scanned image of the original documents written in Bengali script and Figure 5b is the noise-free output image obtained with the application of the proposed method. The four side margins of the page have been set to 1 inch, here. Qualitatively, with the normal vision it is seen that the original scanned image (Figure 5a) has some obvious noise, which may have occurred due to back page shadow or something else. From Table 1 it is seen that, the input image shown in Fig. 5a has the file size of 3754 KB where as processed image (Fig. 5.b) is of 5040 KB. The four-sided white margins have been set with 1-inch width. The percentage of reduction of space here is 34. It should be noted that, we are accessing a page of size 5040 KB for which the original file size is 3754 KB stored in the server taking only the textual regions. Similarly Fig. 6.a and Fig. 6.b are examples of another document image of different type but printed in Bengali. In this case the margin set up is of 1.5 inch and the percentage of space saved is 61. Fig. 7a is a sample document written in Devnagari script. Here, the reduction of space has also occurred. It is important to note that the proposed method is script independent. Considering the amount of the total size of a document image of all pages of a book, it can be concluded that the saved space

Figure 5a The original scanned page.

Figure 5b The output image with 1” margin.
Figure 6a The original scanned page with font variation.

Figure 6b The output image with 1.5” margin.

Figure 7a The original scanned page from Hindi Page.

Figure 7b The output image with 1.5” margin.
Table 1 Analysis of the obtained results from the proposed technique with few sample document images.

<table>
<thead>
<tr>
<th>Images</th>
<th>File Size (KB) [Textual zone]</th>
<th>File Size (KB) [Textual zone]</th>
<th>Save in file size (KB)</th>
<th>Margin size (inch)</th>
<th>% of the save in file size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 5</td>
<td>3754</td>
<td>5040</td>
<td>1286</td>
<td>1</td>
<td>34</td>
</tr>
<tr>
<td>Figure 6</td>
<td>2889</td>
<td>4655</td>
<td>1766</td>
<td>1.5</td>
<td>61</td>
</tr>
<tr>
<td>Figure 7</td>
<td>2627</td>
<td>4335</td>
<td>1708</td>
<td>1.5</td>
<td>65</td>
</tr>
</tbody>
</table>

is of considerable amount, more importantly the output image is noise free and can be margined accordingly as well at the user side.

5. Conclusions
Digital libraries are emerging technologies for document management. It has many branches to be taken care of with equal respect for the overall development of the digital library. Proper and secured content management and delivery is one of the key issues of the digital library. More and more digital material is added every day to the digital library server. Content management technologies will be a big thing in the future. Moreover, the everyday problems of digital library help in the long term preservation of digital objects; copyright of digital material; good solutions for micro charging, and pay per view etc. All these are related more or less with the proper storage of content and planning of storage device management. Due to these problems the present status of the digital library in the third world countries is in a nascent stage.

In the present study, a simple as well as easy-to-implement approach has been presented for the storage management and delivery of e-content materials related to digital library. One of the attractive features of this approach is that it can reduce the space of any scanned document image to a considerable amount and can also in principle be applied to any digital document. Further more the noise reduction algorithm becomes very effective for this type of electronic document delivery. One may say that there are many other noise cleaning algorithms available in literature, ‘Gonzalez and Woods’ (2004), ‘Weeks (2007), but for the present purpose the effectiveness of our noise-cleaning algorithm is sufficient enough. This is in sharp contrast to the noise-removing algorithm used for the present purpose so far. We do not feel the need of complexity analysis of the present algorithm as it can easily run on a desktop or laptop of normal configurations.

The present day digital libraries have to use all modern network and server technologies in order to supply services of a high quality. For faster access and retrieval of a data or metadata smaller file size is one of the issues. The present technique can effectively be applied for the files supposed to serve the user requests.

References


Anwesan: A Search Engine for Bengali Literary Works

Suprabhat Das, Shibabroto Banerjee and Pabitra Mitra
Department of Computer Science and Engineering
Indian Institute of Technology Kharagpur
West Bengal 721302, India
Email: suprabhat4u@gmail.com, shibabroto@gmail.com, pabitra@gmail.com

Abstract
Most of India’s literature was written in Bengali since the beginning of the 19th century. Hundreds of authors have contributed to the enrichment of Bengali literature for years. Besides that, nearly 300 million people around the world speak in Bengali. The language, having a rich traditional background and popularity throughout the world, must be taken care for the web users in the present era of World Wide Web (WWW). The digitization of Bengali literary works and the development of the search engine is very important for the benefit of the Bengali language users all over the world. The paper describes Anwesan, a search engine for Bengali literature. Currently the entire work of Rabindranath Tagore and a part of Bankim Chandra Chattopadhyay’s work is searchable through Anwesan. Several advanced search features necessary for simple and expert users are supported. It also serves as a digital library with various metadata information. The engine is implemented by customizing DSpace in Bengali language and is perhaps the most exhaustive exercise in this direction. This search system was primarily open for the public in Kolkata Book Fair 2010 only with Rabindra Rachanabali collection. Since then, its been in high use.

Keywords: Digitization, Anwesan, Rabindranath Tagore, Information retrieval, Metadata search
1. Introduction

Bengali literary works have a long and rich tradition of over hundreds years. It became rich with its variations over a period of time as it started to spread its different genres, like poetry, story, novel, essay, drama etc. From the very beginning of the history of Bengali literary works, Ishwar Chandra Bandyopadhyay, Michael Madhusudan Dutt, Bankim Chandra Chattopadhyay, Rabindranath Tagore, Kazi Nazrul Islam, Sarat Chandra Chattopadhyay, Sukanta Bhattacharya and many others contributed to enrich the literature. This panorama of literature is now extended by Sunil Gangopadhyay, Buddhadev Guha and others.

In the present era of World Wide Web, the digitization of the huge Bengali literary work is extremely essential. Digital library framework is a convenient mean to manage this huge collection of literary works. However, only storing the documents in digital forms is not the complete task. Finding out the desired document fast from this digital library is also important. Hence, efficient searching technique along with digital library makes the system more user friendly. Apart from that, storing additional information regarding the documents, i.e. metadata, and applying searching procedure on the metadata also, makes the system more informative and robust.

This paper presents Anwesan, a digital library as well as search engine for Bengali literary works. Now the digital library has the complete Rabindra Rachanabali collection and part of Bankim Chandra Chattopadhyay’s work in the database. This is perhaps the first effort of its kind in any Indian language. The rest of this paper is organized as follows. In the second section, details of the Bengali literary collection with metadata are given. The third section gives a brief overview of the architecture of Anwesan. The implementation procedure using DSpace is given in the next section. The fifth section describes the advanced features for Anwesan. The Deployment details are given in section 6 and some future works are described in section 7.

2. The Bengali Literary Collection with Metadata

2.1 Rabindra Rachanabali

The most prolific writer in Bengali literature is Nobel laureate Rabindranath Tagore. He wrote successfully in all literary genres in his lifetime (1861-1941). Although known mostly for his poetry, he also wrote novels, stories, plays and thousands of songs. Besides these, he wrote musical dramas, dance dramas, essays of all types, travel diaries etc. The complete collection of Rabindra Rachanabali is stored in the database of Anwesan along with information about the documents like date of writing, date of publication, name of the main characters etc. The documents are stored in database after classifying according to genres. There are many sub-categories in নাটক (drama), like গীতিনাট্য, নৃত্যনাট্য, ব্যঙ্গকৌতুক, হাস্যকৌতুক, প্রহসন. All these sub-categories are stored under the main genre. There are more than five thousand documents from Rabindra Rachanabali in the database of Anwesan. The number of documents in Rabindra Rachanabali collection according to genres is given in Table 1.

<table>
<thead>
<tr>
<th>Genres</th>
<th>No. of documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>উপন্যাস</td>
<td>13</td>
</tr>
<tr>
<td>নাটক</td>
<td>64</td>
</tr>
<tr>
<td>গল্প</td>
<td>162</td>
</tr>
<tr>
<td>গান</td>
<td>1815</td>
</tr>
<tr>
<td>কবিতা</td>
<td>2475</td>
</tr>
<tr>
<td>প্রবন্ধ</td>
<td>719</td>
</tr>
</tbody>
</table>
3. Bankim Chandra Chattopadhyay’s Writings

Bankim Chandra Chattopadhyay (1838-1894) was the first novelist in Bengali literature. He is regarded as one of the founders of modern Bengali literature and also a key figure in literary renaissance of Bengal as well as India. He is mostly known for his astonishing variety of novels. Vande Mataram (বব্দে মাতরম্) is an eternal poem from his novel Anandamatha, which came to be considered as the national song of India and a part of it was played during the Indian independence movement. The number of documents in Bankim Chandra Chattopadhyay’s collection according to the genres is given in Table 2.

We are thankful to the Society for Natural Language Technology Research\(^1\) for providing us the collection of Rabindranath Tagore and Bankim Chandra Chattopadhyay with relevant information like date and place of writing, date of publication, name of publisher, name of the main characters and many others.

The structure of the database has three tiers. The topmost tier is a community, and under it, there are many collections. In the database of Anwesan, there are two main communities, named Rabindranath Tagore (রবীন্দ্রনাথ ঠাকুর) and Bankim Chandra Chattopadhyay (বঙ্কিমচন্দ্র চট্টোপাধ্যায়). There are many collections according to the genres inside the communities. The collections are novel (উপন্যাস), drama (নাটক), story (গল্প), song (গান), poem (কবিতা) and essay (প্রবন্ধ) for Rabindranath Tagore and only novel (উপন্যাস) and essay (প্রবন্ধ) for Bankim Chandra Chattopadhyay. The low level tier comprises of individual items. All of the writings are stored in the relevant collection as an individual item. Available information about those items, i.e. metadata, is also stored with them. There are different sets of metadata for different collections. A brief overview of the complete data structure along with available metadata for individual items in respective collection is given in Figure 1.

4. Architecture of Anwesan

Anwesan is a digital library as well as search engine for the Rabindra Rachanabali collection. Anwesan is developed on the framework of DSpace\(^2\) version 1.5.2. Like any other search engines, in Anwesan also, the complete functionality consists of two basic parts. First part is indexing and the second one is retrieval (Manning, Raghavan, and Schütze 2008). In the first part, to create an index file from the contents of all the documents, we need to extract every token from documents. A tokenizer is used for that, and from every token, root words are extracted using a stemmer (Xu and Croft 1998, Majumder, Mitra, Parui et al. 2007). These root words are used to create the index file. We use a rule-based stemmer (Sarkar and Bandyopadhyay 2008, Das and Mitra 2011) for this stemming procedure. The whole indexing is done in administrator’s side. All the actions in second part are done in users’ side. In this part, users search for their query. The search engine looks for the user query from that index file created in first part and gives the correct results, that is, relevant to the user query. In DSpace framework, indexing of full text as well as metadata entries can be done. So every word in the repository as well as the metadata entries are searchable in Anwesan using Lucene\(^3\), which is a high-performance, full-featured text search engine library written entirely in Java and used by DSpace.

<table>
<thead>
<tr>
<th>Genres</th>
<th>No. of documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>উপন্যাস</td>
<td>14</td>
</tr>
<tr>
<td>প্রবন্ধ</td>
<td>179</td>
</tr>
</tbody>
</table>

\(^1\) SNLTR: http://www.nltr.org/SNLTR/
\(^2\) DSpace, open source digital library framework: http://www.dspace.org/
\(^3\) Apache Lucene, open-source search engine library: http://lucene.apache.org/
In the architecture of Anwesan, there are many supporting modules for searching procedure. There are tokenizer, stemmer, database and database updating module for indexing procedure. The index file is created for Bengali literary works using these modules. In the retrieval part, the users type their query and get search result through user interface. Searching is done using application module and the rank score module is used to return the search result after sorting them by relevancy measure. Each and every module is written in Java, JSP or servlet. The whole architecture of Anwesan is given in Figure 2.

From the architectural view of Anwesan, it is seen that the whole system consists of mainly three modules. The basic three components are as follows.

- **Display Module**: The display module directly interacts with the users. This module maintains the displaying pages, where a user can enter queries or the pages where the search results are displayed. The user queries are sent to the application module for further processing and the search results are received from the application module to display using this module. The whole module consists of JSP and servlets.

- **Application Module**: The application module is the most important module in the system. Primarily while indexing, this module reads the documents one by one and redirects the contents of the documents to the database updating module. At the time of searching, it takes user query as input, sends it to database updating module to extract root words and finally searches for that root word in database. If there are more than one search results for a query, then all the search results are sorted according to the relevancy.

![Figure 1 Data structure and available metadata for respective collections](image-url)
score and this sorted result is sent to the display module.

- **Database Updating Module**: The efficiency and correctness of the system depends upon this module. Measures like precision and recall are highly depend upon how efficiently this module can extract root words to store in database and to match with database entries. The contents of the documents and user queries are the input to this module. Tokenizer is used to split a string into individual tokens according to a set of delimiters. The resulting tokens are then passed on to stemmer for further processing of the input string. The process can be considered a sub-task of passing input.

5. Implementation in DSpace

The whole system of Anwesan is developed on the framework of DSpace version 1.5.2. Some major changes have been made on this open-source digital repository DSpace, according to our requirements. Though DSpace version 1.5.2 supports Unicode, but there were no available tokenizer or stemmer for Bengali language in the existing system of DSpace. We developed a tokenizer and a rule-based stemmer for Bengali, and integrated them with our system. After integrating these modules with our system, the correctness, i.e. recall value of the search result increased significantly. The rule-based stemmer was also tested on Bengali collection of the FIRE 2010 data set with 50 queries using Lucene as the search engine and it gives 96.27% recall value (Das and Mitra 2011). Besides that, numeric values, like page numbers, dates, and so on, are displayed in Bengali alphabets in our system. Each and every web pages in the system are very simple and user-friendly, so that the users can understand the content of the page easily as well as pages are made light-weight to load through web and display quickly in web browser. Perhaps this has been the most extensive exercise in extending DSpace to an Indian language.

Many new metadata fields are added with existing DSpace system according to our requirements. We had tried to populate every metadata field on the basis of availability of information. The available metadata entries for every document are displayed in tabular format so that users can find results according to their requirement easily.

A user feedback form with Bengali interface is also provided, so that users can post their comments in Bengali. This is very helpful as we
get many ideas about the up-gradation of our system from these user feedbacks.

6. Advanced Features

In any digital repository system, browsing and searching is the main criteria for its robustness. In Anwesan, we strongly emphasise on the searching criteria. Our basic aim is to search from the whole repository according to users’ query. For that, the functionality for searching full text as well as searching from metadata entries is provided in Anwesan. Apart from that, some advanced search features are included in the system. In advanced search, users can search from metadata values, and the most useful feature is that user can build complex query by merging maximum three metadata field values by logical operator like AND (\(\land\)), OR (\(\lor\)), NOT (\(\neg\)). In this feature, the relevant metadata are listed for every collection. For example, the metadata entry for ‘প্রধান পুরুষ চরিত্র’ is ‘মহেন্দ্র’ and for ‘প্রধান নারী চরিত্র’ is ‘বিনোদিনী’ and they are nested using AND operator, then the query will be like ‘(প্রধান পুরুষ চরিত্র: মহেন্দ্র) AND (প্রধান নারী চরিত্র: বিনোদিনী)’. It will give ‘চোখের বালি’ as search result.

The snapshot of the advanced search page is given in Figure 3.

Some special search features are included in the system of Anwesan for the benefit of users. Users can also search a query only within some specialized collection, excluding the remaining part of the whole community. This feature makes the whole searching space smaller and makes the search procedure faster. If the user knows the collection’s name, in which the required document exists, this searching criteria will be helpful.

In the search result page, the list of retrieved documents is listed along with the collection name, in which it exists. All the information about the retrieved document is displayed in two phase. The whole set of metadata for the document is displayed robustly in first phase. A link, named ‘নামাঙ্কিত মাল্টি প্যাকেজ’, is given here to redirect to the second phase. It will display the content of the whole document with query words highlighted. Using this feature, user can be aware about the exact location of query words in the document. This feature of highlighted page is invoked in Anwesan, which is not present in the DSpace system. A highlighted page of the document ‘দুই উপমা’ for query ‘যে নদী হারায়ে স্রোত চলিতে না পারে’ is given in Figure 4.

As there is no cross-lingual support in Anwesan, users must type their query in Bengali alphabets only. But typing the queries in Bengali alphabets is a problem for users as they do not have a Bengali keyboard. Hence, a virtual Bengali keyboard is provided in the web page of Anwesan. But it is seen from the statistics of user logs that many users have been typing query in English alphabets. So recently a transliteration (translation) engine is also provided.
provided, so that an user can type Bengali words in English alphabets phonetically.

7. Deployment

The system is currently a complete collection of Rabindranath Tagore’s works. A part of Bankim Chandra Chattopadhyay’s work has been uploaded in the database of Anwesan. The website for Anwesan (anwesan.iitkgp.ernet.in) was released during Kolkata Book Fair 2010 only with Rabindra Rachanabali collection. After that also, continuous up-gradation with modifications and additional new features to the website is in progress.

After the website was open for public, we logged the user query details. The number of search performed, number of items viewed, user query and many other features have been logged in the statistic file of the system. Number of search performed per month from the month of February 2010 to February 2011 for only Rabindra Rachanabali collection is displayed in Figure 5. From this figure, it is seen that there was a huge number of hits just after Anwesan was publicly open, and after that there is a steady continuation of the same.

8. Future Development

Due to huge response from the users, we have plans for adding the literary works of other authors, like Bankim Chandra Chattopadhyay, Sarat Chandra Chattopadhyay etc. in Anwesan.
The writings of Bankim Chandra Chattopadhyay is now being uploaded in the database of Anwesan. Besides, to make the search result more reliable, it is advantageous to show some snippets along with the entries in search result page. Hence, snippets can be added with search results. To create the repository, the administrator has to submit all the documents first in the repository. The existing DSpace system does not support any batch submission. So submitting each and every document along with the metadata entries is too time consuming and costly. An automatic submission procedure in DSpace system, having the feature of submitting the whole collection to the repository in one click, which can make administrators work easier, is being developed.

References


Suggested readings
Friedrich Summann and Norbert Lossau Search Engine Technology and Digital Libraries, Moving from Theory to Practice, D-Lib Magazine September 2004, Volume 10 Number 9, ISSN 1082-9873


Ratna Sanyal Sanyal, Kushal Keshri and Vidyanand, Importance of Retrieving Noun Phrases and Named Entities from Digital Library Content, Journal of Zhejiang University Science C (Computers and Electronics), Vol. 11, No. 11, pp. 835-930


Adding Value to an Abstracting and Indexing System: The Case of MyAIS, Malaysia

Zainab A N
Professor, Digital Library Research Group, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
E-mail: zainab@um.edu.my

Abstract
There are problems in (a) identifying the entire scholarly journals published in Malaysia, (b) locating and accessing the journals, (c) knowing who has published what and in which journals; (d) assessing the indexation status of the journals; and the citation received and general impact count of the journals. The Soft Systems Methodology (SSM) is used to frame the study, especially in determining the problem situations, extracting the problems and proposing a solution in rich pictures, initiating the prototype, comparing existing process with those of the proposed system and making improvements from feedback. The prototype MyAIS or Malaysian Abstracting and Indexing System is described, especially the features that adds value to the system and motivate journal publishers and authors to contribute to the system in order to increase availability and accessibility.

Keywords: Indexing, Abstracting, Malaysia, Citations, General impact, Scholarly journals.
1. The problems

Online indexes and abstracting services are fast becoming the first choice of libraries because of the added value they provide in giving full access to journals’ contents and allowing access remotely, which seems to work well with today’s new user generation (Katz, 2002). This leads us to the first problem. While access to these abstracting and indexing services are true for mainstream foreign journals, it is not so for most national journals published in developing countries. For these countries, which includes Malaysia, scholarly works published by academic publishers and professional associations remain inaccessible and sadly unused as well as cited. As such, an open access abstracting and indexing system could help solve this problem. This is exactly what MyAIS or the Malaysian Abstracting and Indexing System provides for Malaysian scholarly works.

Also, in Malaysia, it is difficult to estimate with certainty the exact number of scholarly journals in existence. There is no single directory available that provides such a list, and, if there is one, digital born electronic journals very often would be missed (Roosfa, 2006). This applies to both print and electronic scholarly journals in Malaysia. Md Sidin (1997) appended a list of 214 journal titles to a paper presented at a conference, and indicated that 59.3% of the titles were in the Arts, Humanities, Social Sciences and 40.7% were in Science, Technology and Medicine. Normah (1999) reported a total of 284 titles recorded through an ISSN listing issued by the National Library of Malaysia. A more recent journal audit was carried out by Zainab et al. (2012), which reported a total of 464 journal titles tracked through library catalogues of the National Library of Malaysia (a legal repository) and the University of Malaya Library (the oldest university library in Malaysia). Most Malaysian journals are published in single volumes annually and about a third published twice a year. Knowing what titles exist is extremely important to estimate just how large is the corpus of Malaysian scholarly journal article publications available for reference to the Malaysian research community, that needs to be covered by a national abstracting and indexing system.

Problem is also faced by Malaysian scholars of knowing which journals are best suited to disseminate research results. It cannot be denied that technology can both empower or exclude people by determining their access to knowledge (Adams, Blandford and Luntm, 2005). Malaysian scholars often do not know what their colleagues have published in their respective fields and therefore seldom cite each other (Science and Technology, 2004). This has resulted in the flow of quality articles out of the country and low submissions to local academic journals. This situation inevitably affects the attempt of Malaysian scholarly journals to sustain their content quality. It is therefore extremely important to provide them with an avenue where they can not only self-archive their issues but also make their articles more visible to the global educational community.

Conversations with editors of academic journals reveal that the editorials are often so bogged down with work with minimal help from support staff, that they have no time to strategize about marketing or improving the status of their journals. They are often unaware of how to gauge the quality of their journals in terms of international standards. To enable publishers to periodically self-assess the status of their journals in terms of quality would be an added bonus for academic publishers so that they could roughly estimate how to alleviate their journals toward an international status. Internationalization is related to the indexation status of Malaysian journals. An indicator which shows that a journal has reached an international status lies in its ability to get itself included in an indexing and abstracting service both at national and international levels. A national indexing system would help spearhead the indexation status by international databases such as Scopus, the ISI
databases and other subject-based indexing services. This article focuses on the initiative to develop a Malaysian abstracting and indexing system or MyAIS for scholarly journals and value added features provided by the system to improve availability, visibility and citability of Malaysian scholarly works. Ultimately, the initiative would contribute to enrich the corpus of Malaysia’s scholarly e-content.

2. Objectives
The objectives of this initiative are:

i. To identify the problems faced by stakeholders (journal publishers, librarians and information professionals, researchers and authors) in locating, accessing and using scholarly journal contents in Malaysia.

ii. To find out the status of Malaysian journals in terms of its availability in listings either printed or electronically and their indexation status by Malaysian or international databases.

iii. To provide an indexing and abstracting system which empowers journal publishers and individual scholars to archive their journals and scholarly works and at the same time provide access and visibility of contributions over the Internet.

iv. To add value to the system by providing citation and impact information for both authors and journals.

3. Methodology
The soft approach was used to identify problem situations extracted from discussions reported in literature surrounding the issues of access to Malaysian scholarly works. The soft method approach is found suitable when dealing with problems that are difficult to define, problems that involve social, political and human activity components. The social and educational implications of the situation dealing with the availability, visibility of scholarly journals is suspected. We know the problem exists but it is difficult to pinpoint exactly the extent of the problem and the possible systems required to alleviate the situation. The soft approach was developed by Checkland (1981, 1998) and Checkland and Scholes (2001), who proposed that all type of problems can be solved using the soft method. The method emphasizes on identifying and expressing problems inherent in any situation, considering the views of stakeholders and proposing a possible solution. This approach helps to vie problems from various angles and ensures that the solution works for all stakeholders involved. Even though the method proposed has seven stages (Figure 1) Checkland informed that one can start at any stage and may even need to combine some stages. In this project, at various stages the data gathering approach used vary, comprising discussions and exchange of ideas at workshops and conferences with journal publishers, writers and librarians as well as information obtained from published literature.

Stages 1 and 2 comprise identifying problems and expressing it in a rich picture using information extracted from discussions and exchange of ideas from these occasions (a) 30 minute presentations and 20 minute discussion each at two scholarly journal publishing workshops and two seminars on the internationalization of Malaysian academic journals attended by academic journal publishers and (b) problems expressed in published literature. Stage 3 will identify the roles and estimate the requirements of stakeholders, in this case the authors/researchers/academicians, the journal publishers and the librarians. The proposed requirements are mapped to the proposed system.

Stage 4 will indicate the proposed conceptual model of the system and Stage 5 will compare the proposed solution with the current problem situations. Stage 6 and 7 comprise actual proposed designs of a prototype and populating the system with sufficient content to simulate the running of the system.
4. The results of the soft approach

4.1. Stages 1 and 2: Problems identified and expressed

Stages 1 and 2 entail finding out the problems from three groups of stakeholders, the journal publishers and editorials, the librarians and information professionals, and Malaysian researchers and academicians.

(a) Problems expressed by publishers and editorials

The problems expressed by journal publishers and editors are extracted from dialogues and exchanges of ideas during two workshops on scholarly journal publishing (Workshop on the Editorial Proses, February 2009; Workshop on Managing Scholarly Journals, 13-16 January 2008) and two conferences on internationalization of Malaysian scholarly journals (Regional conference on Scholarly Journals, March 2006 and National Conference on Scholarly Publishing, April 2009). In the two workshops, the author conducted two 30 minute sessions with academic publishers where during and after the talk editors voiced their constraints and limitations experienced when publishing their journals. Problems are also identified from two conferences where, after the author’s presentation, the editors from the floor voiced their problems and opinions. Out of these venues a number of problem issues were extracted. This is indicated in Figure 2.

In summary, journal editors are facing poor submission problems and this makes it difficult for them to maintain frequency punctuality. Very few get financial help to enable them to engage support personnel to help manage the journals. Also, very few journal editors are aware of their role in getting their journals indexed by an abstracting and indexing database both nationally or internationally as a strategy to improve their visibility. As a result, many of the journals were either institutional or faculty based where in most cases the articles were mainly submissions from academics of the same institutions and faculties. They were also unaware of the process of getting their journals indexed either by the ISIS databases, Scopus, or other discipline-based databases. Very few provide an electronic version of their journals.

Figure 1 The Soft System Map
As a result, the publishers have little idea of the total number of scholarly journals published, who are publishing what and who are citing the articles. The publishers expressed willingness to contribute to the system to ensure their journals are indexed and to increase their visibility globally.

(b) Problems expressed by librarians and researchers in literature

Problems in determining the corpus of Malaysian scholarly journals are abstracted from published literature. Md Sidin (1997) and Roosfa (2006) both indicated this problem. Currently, no comprehensive listing of scholarly Malaysian journals can be located. The two listings compiled by the Malaysian National Library are outdated (Senarai induk terbitan bersiri = Malaysian union list of serials, 1976; Indeks majalah Malaysia = Malaysian periodical index, 1973 – 1990). MyULIS or Malaysian Union list of serials compiled and published by PERPUN or The Conference of Academic Librarians and the National Library covers total serials holdings including those published outside Malaysia and no separate Malaysian journal listing is given (MyULIS, 2008). Roosfa (2006) provided a listing in 2006 and mentioned the elusive and mushrooming journal titles from the newer universities as well as those published solely in the electronic form. All these situations indicate that there is no single current list of Malaysian journals and this poses a problem not only to
librarians to keep track but also researchers and authors to support their referencing needs. This is exacerbated by the lack of a national indexing and abstracting services which should cover these scholarly journals as the indexing initiative carried out by the National Library of Malaysia have ceased (*Indeks majalah Malaysia = Malaysian periodical index, 1973 – 1990*) and moreover no abstracts and citation information was provided by this index. Zainab (1997) who studied the indexation status of Malaysian scholarly journals in 23 indexing and abstracting services, which represented the major disciplines, found that at that time only 54 Malaysian journals were indexed by subject-based international databases. Out of this number 43 were in the STM and 11 titles were in the AHSS. A more recent study indicated that there are currently 49 Malaysian journals indexed by Scopus and 9 titles by Thomson Reuters, Web of Science (Abu Bakar, et al., 2009; Zainab et al, 2012). These situations revealed that the total coverage of Malaysian journals in international indexes was very small and the contents of the majority of Malaysian journals remain inaccessible and invisible. Stages 1 and 2 of the SSM provide answers to objectives 1 and 2 of the study.

4.2 Stage 3: Roles and determining requirements

This stage involves identifying the needs and roles of stakeholders, in this case the authors, researchers, the journal publishers and the librarian and environment where the problem is situated as well as the underlying criteria of the proposed system. Their roles are determined and a conceptual model of how the situation can be improved is proposed (Table 1).

4.3 Stage 4: Proposed conceptual models

Funding is an important issue in this context. The initial amount of fund received was too small to contract the development of the proposed solution from scratch. As a result, the initiative proposed the adoption of an open software which can be customised to support the requirements. The software should also support an open access architecture and is Open Archive Initiative (OAI) compliant because (a) this is encouraged by the Malaysian government; (b) it will easily allow search engines to harvest the systems metadata; and (c) will increase the accessibility and visibility of Malaysian articles on the Web. The *E-print* software from Southampton University was chosen and customized to support the requirements of MyAIS. This software provides an efficient and dynamic indexing system that returns fast feedback when querying its ‘browse and search’ modules. The repository function of the system is enhanced and added in value through a statistics module which provides citation counts, impact scores, h index, viewing and download information. It is this statistic features that would hopefully motivate journal publishers to contribute, to make their presence visible, and to obtain citation counts for their journals as well as assess the general impact of their journal at least nationally. The conceptual models which would help improve the situation is indicated in Figure 3 and the main concepts are as follows:

- **The people involved and the support for empowerment:** The system allows publishers/ authors/ librarians to self-archive their own articles/journals and hence ensuring an archive collection is available on an alternative site. This will sustain their journal’s availability and visibility in the long run. The system should be able to provide a dynamic masterlist of Malaysian scholarly journals.

- **The Process:** The system should provide information about the number of article citations /views/ downloads. The system should also be able to extract from the articles’ contents and provide the information about 1) the citation count for authors (number of articles citing an author’s work); 2) the number of citations obtained by journals available
Table 1 What the Proposed System should Support

<table>
<thead>
<tr>
<th>Current Roles</th>
<th>Requirements Extracted for Future Roles</th>
</tr>
</thead>
</table>
| **(1) Editors:** ‘Publishers’ | Manage editorial and publishing process electronically.  
 current roles: Manage publication process; ensure punctuality of publication; maintain publication style, distribute issues to libraries, individuals and bookshops.  
 Fast and easy publication delivery process and shorter submission to publication and delivery stage.  
 Able to handle scanned images of journal articles.  
 Easy and intuitive interface and instructions for non-technical personnel.  
 Able to control access - determine which issues are freely accessible and which issues or articles are not.  
 Would not incur cost and time in maintaining and monitoring the system  
 Support uploads from back issues as well as current (an archival repository)  
 Able to provide bibliographic and citation information to the global learning community.  
 Able to handle storage, protection, preservation problems of their journals.  
 Able to know who are citing articles published in Malaysian journals and the national impact of their journal. |
| **(2) Authors and researchers**  
 Current role: Submit articles by mail or email attached files to editors; priority given for sending articles abroad because local titles could not be easily located and irregularly published; visit the library to browse and search for required articles. | Able to submit and archive their own published or unpublished works.  
 Able to know their total works available in the system instantaneously.  
 Make known their works to Malaysian and foreign researchers in the same field.  
 Able to know who are citing their works.  
 Able to browse and search for who else and in which journals their peers are publishing in Malaysian publications.  
 Able to access works published in local scholarly journals and cite relevant works.  
 Able to get their works in an indexing and abstracting system and disseminate their works over the web, at least accessible in Google Scholar.  
 Able to determine which of their works are open for access and which are close for access. |
| **(3) Librarians and Information Centre Managers**  
 Current roles: Compile listings of journals upon request; compile indexes without abstracts or full-text. | Able to determine the status of Malaysian scholarly journals (their number and bibliographic details).  
 Able to update current journal information so that the scholarly journal listing becomes dynamic.  
 Empowered to upload, index, and abstract own institutional journals with the option to provide full-text access to published articles (institutional repository).  
 Empowered to archive their institutional journals even those which have ceased publication.  
 Able to edit, delete their own records entries.  
 Efficiently handle multi-collections in a union digital library. |
| **(4) Systems Administrators**  
 Current roles: Not available in the current role. | Create, edit, delete new editors and users.  
 Able to administer and troubleshoot access problems on site and remotely.  
 Support metadata fields for different types of information resource. |
within the system and this will allow for a rough calculation of each journal’s general impact factor and h index; and 3) the number of articles that an author has published in national journals. The system will be populated with sufficient number of contents to simulate the various functions of modules in the system. The target is around 7,000 articles from Malaysian journals. Also, the system should be able to provide added information about the traffic of usage and users.

- Quality assessment: The system should be able to provide journal publisher with a self-assessment journal quality tool to gauge the status of their journals. This is also an add-on feature.

4.4. Stage 5: Compare stage 4 and 2
Table 2 compares the current problematic situations with the proposed solution (comparisons between stages 4 and 2).

4.5. Stage 6 and 7: Developing modules and initiate improvements
Stages 6 and 7 involve detailing the features of the proposed system. MyAIS currently holds over 12,000 articles abstracts and full-text documents since its inception in early 2007. The articles are published in a few Malaysian journals and local conference proceedings. Details of the proposed system Client/User modules is given in Table 3 and the Administrators module is given in Table 4. Stages 3 to 7 provides answers to objectives 3 and 4 of the study.

5. The project outcome

5.1 Empowering stakeholders
The project outcome is an open access solution named MyAIS (Malaysian Indexing and Abstracting System). It was developed to support a national repository for Malaysian scholarly...
<table>
<thead>
<tr>
<th>Current Problem Situations</th>
<th>Proposed Solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Journal Publishers / Editorials</strong></td>
<td><strong>MyAIS</strong></td>
</tr>
<tr>
<td>1. Low quality submissions</td>
<td>1. Indexed nationally and internationally when contributing content. MyAIS is an open access OAI compliant system accessible through the internet and indexed by Google Scholar</td>
</tr>
<tr>
<td>2. Not indexed nationally and internationally</td>
<td>2. Increase accessibility and visibility. Accessibility is one of the criteria used by indexing agencies such as Scopus and ISI databases when reviewing indexation requests</td>
</tr>
<tr>
<td>3. Not easily accessible and visible</td>
<td>3. Gives citation counts for both authors’ works and journals</td>
</tr>
<tr>
<td>4. Not aware who are citing articles published in journals</td>
<td>4. Free use of the MyAIS system to support journal archiving and current publications</td>
</tr>
<tr>
<td>5. None access to an indexing and abstracting system locally to request for coverage</td>
<td>5. Provide self-assessment journal quality tool to estimate the current quality of a journal</td>
</tr>
<tr>
<td>6. Not aware of the journal’s impact and quality</td>
<td>6. Allow publishers to publish and archive their journals without having the technical know-how on how to maintain the system</td>
</tr>
<tr>
<td>7. No funding for support of electronic publishing,</td>
<td></td>
</tr>
<tr>
<td>8. No technical know-how on how to publish electronically</td>
<td></td>
</tr>
<tr>
<td><strong>Librarians / Information Professionals</strong></td>
<td><strong>MyAIS</strong></td>
</tr>
<tr>
<td>1. Could not effectively control scholarly journals bibliographically</td>
<td>1. A dynamic listing that can add, edit, delete journals titles and the information is available to users immediately</td>
</tr>
<tr>
<td>2. There is no national directory of Malaysian scholarly journals</td>
<td>2. The listing is only on Malaysian journals even though authors may archive their works published in other non-Malaysian journals</td>
</tr>
<tr>
<td>3. The listings compiled by individuals and institutions tend to be outdated</td>
<td></td>
</tr>
<tr>
<td><strong>Researchers / Academicians (Users)</strong></td>
<td><strong>MyAIS</strong></td>
</tr>
<tr>
<td>1. Do not know how to search for peers publishing in the discipline in local journals</td>
<td>1. Users can browse and locate articles by authors’ names, keywords, institutions, year, journal title, and format</td>
</tr>
<tr>
<td>2. Do not know who is publishing what and where in local journals</td>
<td>2. Users can view information about the citation and impact factor of their works</td>
</tr>
<tr>
<td>3. Difficult to access local journals</td>
<td>3. Users can know total viewing, downloads, and citation of their works</td>
</tr>
<tr>
<td>4. Difficult to know the indexation status of the journals.</td>
<td>4. Users can self-archive their works which will be made accessible nationally and internationally over the web and google scholar.</td>
</tr>
<tr>
<td>5. Cannot determine whether the journals located are still current</td>
<td>5. Can decide which of their works can be given full access and which can be denied access to full-text</td>
</tr>
<tr>
<td>6. Is there an electronic version of local journals in my subject area?</td>
<td></td>
</tr>
</tbody>
</table>
### Table 3 Client / User Modules for MyAIS

<table>
<thead>
<tr>
<th>Modules</th>
<th>Menu Item</th>
<th>Sub-Menu Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Information</td>
<td>Home (Return to the main page)</td>
<td>About MyAis</td>
<td>Returns to the default page of MyAIS. Client views general description and identify number of items indexed and number of full-text uploaded into the system.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Background</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Objectives of the research</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Work Teams</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Funding</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Contact information</td>
</tr>
<tr>
<td>Search Engine</td>
<td>Word Search</td>
<td>Author search</td>
<td>Indicate quantity and lists works by an author's name - most recent first</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Subject search</td>
<td>Lists Works by a subject chronologically - most recent first</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Keywords</td>
<td>Lists works by a keyword chronologically - most recent first</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Citing search</td>
<td>Lists works that cite the author, includes self-citation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Advance Search</td>
<td>Options to search full-text, title, creators, abstract, keywords, subject (controlled), item type, department, editors, status, references, journal titles, date, and specify order of listing (most recent, oldest first, author, title).</td>
</tr>
<tr>
<td>Browse Options</td>
<td>Browse</td>
<td>By Author, Editor</td>
<td>Lists Authors' and editors' names, number of items indexed is indicated</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Author by Alphabet</td>
<td>Lists alphabet A-Z</td>
</tr>
<tr>
<td></td>
<td></td>
<td>By Journal</td>
<td>Lists journal titles A-Z, number of items indexed is indicated</td>
</tr>
<tr>
<td></td>
<td></td>
<td>By Subject</td>
<td>Lists LC broad subject categories, number of items in each category is indicated</td>
</tr>
<tr>
<td></td>
<td></td>
<td>By Affiliation</td>
<td>Lists affiliations A-Z, number of items indexed under each affiliation is provided</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Malaysian Journals</td>
<td>Journal Master list Directory</td>
</tr>
<tr>
<td></td>
<td></td>
<td>By Type</td>
<td>Lists types of items, journals, conferences, books, book chapters, thesis, and monographs. Number of items indexed under each type is provided.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>By Year</td>
<td>Lists years, number of items indexed under each year is provided.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>By Latest Addition</td>
<td>List items of three days back</td>
</tr>
<tr>
<td>Browse Statistics</td>
<td>Browse</td>
<td>Author Statistics</td>
<td>Total articles, total downloads, total views, total citations received</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Top 20 Authors</td>
<td>Rank, authors’ names, total views, and downloads</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Top 20 Articles</td>
<td>Rank, titles, total views, and downloads</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Top 20 Journals</td>
<td>Ranked journal titles, total citations received</td>
</tr>
</tbody>
</table>
Table 3 Contd..

| Top 20 Cited | Rank, titles of articles, total citations received |
| Top 20 Authors Cited | Rank, authors’ names, total citations received |
| Journal Statistics | Total cites, total articles, H index, 5 year impact, 2 year impact, Immediacy Index, cited half life, RM index |
| Usage Statistics | Daily, monthly usage statistics, total visits, pages, sites, total URLs, top countries accessing |

Registered User

- Register: Register in order to deposit items
- User area
  - Manage deposits - new items: Specify item type, upload item, input item details, specify item subjects, save items, send items for review
  - Manage deposit - import items: Upload imported record, test run imports, send items for review
- Login: Enables user to login as a registered member
- Logout: Enables user to logout as a registered member

Table 4 Administrator’s Module (System Administrator, Systems editors)

<table>
<thead>
<tr>
<th>Modules</th>
<th>Menu Items</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrator</td>
<td>Status</td>
<td>Monitor status information for this server</td>
</tr>
<tr>
<td></td>
<td>Search items</td>
<td>Search for items in the repository to effect amendments</td>
</tr>
<tr>
<td></td>
<td>Search users</td>
<td>Search for users of the repository</td>
</tr>
<tr>
<td></td>
<td>Search history</td>
<td>Search for activities that has occurred on items in the repository</td>
</tr>
<tr>
<td></td>
<td>Extra admin</td>
<td>May add extra admin menu, example journal authority name</td>
</tr>
<tr>
<td></td>
<td>Create user</td>
<td>Create new user account</td>
</tr>
<tr>
<td></td>
<td>Stops indexer</td>
<td>Stops full-text indexer</td>
</tr>
<tr>
<td></td>
<td>Edit subject</td>
<td>Edit, delete subject tree(s)</td>
</tr>
<tr>
<td>Review</td>
<td>Review</td>
<td>View, edit, save submissions from depositors for review</td>
</tr>
<tr>
<td>Saved searches</td>
<td>Saved searches</td>
<td>View saved searches and request to receive periodic email updates, if anything in the search changes.</td>
</tr>
<tr>
<td>Profile</td>
<td>Modify profile</td>
<td>Edit account information</td>
</tr>
<tr>
<td></td>
<td>Saved searches</td>
<td>View and configure saved searches</td>
</tr>
<tr>
<td></td>
<td>User history</td>
<td>View all changes to items caused by a user</td>
</tr>
<tr>
<td></td>
<td>Administrators Account</td>
<td>Change username and user type</td>
</tr>
<tr>
<td></td>
<td>Delete user</td>
<td>Remove user from the system</td>
</tr>
<tr>
<td>Manage Deposits</td>
<td>New items</td>
<td>Specify item type, upload item, input item details, specify item subjects, save items, send items for review, save items to live archive, send items to retire.</td>
</tr>
<tr>
<td></td>
<td>Import item</td>
<td>Cut and paste, upload imported record, test run imports, save items to live archive, send items to retire.</td>
</tr>
<tr>
<td>Access</td>
<td>Login</td>
<td>Assists administrator to log in</td>
</tr>
<tr>
<td></td>
<td>Logout</td>
<td>Enables administrator to log out</td>
</tr>
</tbody>
</table>
journal contents. It is a national rather than an institutional repository. MyAIS is an open access system for abstracts and indexes of articles focusing on those published in referred scholarly Malaysian journals and conferences (Figure 4).

The system provides the following information:

- Bibliographic information about articles published in Malaysian scholarly journals (include information such as title, author, affiliation, contact information, abstract, keywords, references listed at the end of each article to generate the citation count).
- Supports full-text uploads by publishers and authors who wish to archive and share works and journal contents with other users.
- Supports dynamic links to full-text articles residing in other repository, websites or portals as well as supports dynamic links to other related articles available within the system.
- Lists publication total by any authors’ names, affiliations, format, and by journals.
- Lists articles published in each journals arranged by year.
- Provide a master list of scholarly journals published in Malaysia regardless of whether they are indexed or not in MyAIS
- Provide the total number of views, downloads and citations obtained by an author, an article and a journal covered by the system.
- Provide the total number of citations and general impact factor score received by Malaysian journals included in the system.
- Provide statistical information about total usage of the system, by month and from which country.

It is expected that the data would be richer and more meaningful when all articles published in scholarly journals or in Malaysia are included in the system. MyAIS has empowered stakeholders to locate, access, and use scholarly journal contents in Malaysia.

5.2. Bibliographic control of Malaysian scholarly journals

MyAIS supports a small module that allows repository administrators or any assigned
librarian to update, edit or delete the journal masterlist of Malaysian scholarly journals under the module 'Malaysian Journals' (Figure 5). Currently, a total 464 scholarly journals are listed and this list can be updated and viewed instantly.

5.3. Access and visibility over the internet

MyAIS provides not only an indexing and abstracting service but also offers full-text access to scholarly articles in journals published in Malaysia or abroad in two ways. Firstly, through dynamic links to full text that resides in other digital collections and secondly through voluntary uploads by publishers, editors, and authors. An increasing number of editors or scholarly journals have requested for indexation. This is a non-profit making project that attempts to serve Malaysian educational and research community with information about what is published in Malaysian referred journals for each year and in various disciplines. In other words, it empowers users by allowing them to contribute contents and access to full-text articles. These features of providing access and successfully making its contents indexed by Google scholar seems to motivate journal publishers to participate in the initiative, especially those in medicine and related fields. Another motivating feature is the provision of statistical information such top authors.
top articles, top journal views/downloads, top journals cite, and journal as well as author citation counts. Figure 6 indicates the usage statistics of MyAIS which shows an increasing number of users accessing MyAIS for the past 12 months. This means that there is equal chance for any article indexed by MyAIS to be picked up by researchers that could be used or cited.

5.4. Adding value by giving citation counts for authors and journals

MyAIS adds value to submissions by providing citation information for authors as well as for journals.

Figure 7 indicates a print screen of a search under author ‘Abrizah’. The results show citations received by Abrizah and the highlighted records are self citations. A summary count of the citations received is displayed at the top of the display list. In Figure 8 display of the results under ‘top journal cited’ is given, listing the journals by the number of citations received. Figure 9 shows an example of the sub-module ‘Journal Citation Statistics’. Under this option, results can be ordered by the 2 year impact factor, h-index and RM (relative measure) Index, which is an index specifically formulated under this research to evaluate Malaysian journals.

6. Conclusion

There are numerous open source software available for different purposes (pre-prints, journals, theses, government publications) that can be plucked, enhanced, and used. The e-print is an example of an open software. This type of open source software have changed scholarship and the scholarly communication process. The spirit supported by the open systems are voluntary contributions and the mutual sharing of resources. In the field of medicine, physics and chemistry open systems are welcomed. In Malaysia, many still regard the idea with caution. They are unnecessarily worried about copyright issues and this has resulted in Malaysia lagging behind in the open systems initiatives. An open system means...
Adding Value to an Abstracting and Indexing System: The Case of MyAIS, Malaysia

Figure 7 Citation Received by an Author

Figure 8 Results Displayed under “Top Journal Cited”
there are no restrictions in terms of access to resources and systems. They are developed to support a collaborative environment. In summary, MyAIS has made it possible 1) to gauge Malaysian publication performance at both individual and national levels; 2) to motivate Malaysian academic publishers to contribute and archive their journals in order to improve the availability and visibility of their scholarly journals; 3) to contribute to enriching national research content; 4) to have better knowledge of the nation’s research outputs funded by universities and research institutions; and 5) to improve information access to end-users. MyAIS empowers the scholarly community of users and can support utility at the national level. This would provide access to published Malaysian scholarly works and it may trigger collaboration and citations.

Acknowledgement
This project was made possible from two fundings; Fundamental grant (FRGS) provided by the Ministry of Higher Education through the University of Malaya in 2006: FP056/2006A and Research University Grant (PPJ) in 2007 to 2008.

References
Adding Value to an Abstracting and Indexing System: The Case of MyAIS, Malaysia


FORM IV
(as per Rule 8)

Statement regarding ownership and other particulars about: WDL (World Digital Libraries: An International Journal)

1. Place of Publication : New Delhi

2. Periodicity of its Publication : Half-yearly

3. Printer’s Name
   Whether Citizen of India : Dr R K Pachauri
   Yes
   Address : The Energy and Resources Institute
   Darbari Seth Block
   IHC Complex, Lodhi Road
   New Delhi – 110003

4. Publisher’s Name
   Whether Citizen of India : Dr R K Pachauri
   Yes
   Address : The Energy and Resources Institute
   Darbari Seth Block
   IHC Complex, Lodhi Road
   New Delhi – 110003

5. Editor’s Name
   Whether Citizen of India : Debal Chandra Kar
   Yes
   Address : The Energy and Resources Institute
   Darbari Seth Block
   IHC Complex, Lodhi Road
   New Delhi – 110003

6. Name and address of individuals who own the newspaper and partners or shareholders holding more than one per cent of the total capital : The Energy and Resources Institute
Darbari Seth Block
IHC Complex, Lodhi Road
New Delhi – 110003

I, R K Pachauri, hereby declare that the particulars given above are true to the best of my knowledge and belief.

Signature of Publisher
(Sd/-)

Date: 1 July 2012
Dr R K Pachauri
Abstract
This paper attempts to discuss key issues surrounding the collection management of electronic books (e-books) with a developing country perspective. The libraries have witnessed the emergence of e-books in recent years by major publishers and content providers. With the expansion of electronic publishing, research and academic libraries have started to complement their print holdings with electronic publications. The market for electronic publication seems to be in constant change and collection developers and acquisition managers, especially in libraries, find themselves disillusioned as new methods of electronic publishing and accessing are emerging. Libraries are in a transition period from conventional to digital formats and have not yet developed common practices and traditions on policy for digital material.

The paper discussed the evolution, chronology, collection development, evaluation, delivery, and archiving of e-books. It focuses on contemporary management from the perspective of a librarian concerned with Collection Management. Status of e-books collection of Higher Education Commission (HEC) of Pakistan’s Digital Library Program (HEC-DLP), a milestone project in the scholastic history of the country, has also been discussed. Moreover, some practical steps for local creation of e-books have been proposed. The paper concludes by highlighting major challenges faced by libraries in collection management and management of e-books.

The paper is an outcome of the review of published literature in the reputed LIS journals on the topic, and writer’s 20 years’ first hand experience of administrating the prestigious academic libraries including the oldest Punjab University and modern GC University Lahore.

Key Words: Electronic Books; Collection Management; Academic Libraries; Developing Countries; Pakistan.
1. Introduction

During the last three decades various terms (selection, acquisition, collection building, collection development, and collection management) have been used to denote knowledge acquisition and information sources in the libraries. During 1980s, the terms collection development (CD) and collection management (CM) were used synonymously. However, since 1990s, the literature presented different aspects of these two terms because of the adoption of technological applications in the libraries (Ameen, 2004). A detailed definition presented in ‘International Encyclopedia of Information and Library Science’ is as follows:

“CM is a broad term that has replaced the narrower ‘collection building’ and ‘collection development’ of former decades…it thus encompasses the activities traditionally associated with collection development—the selection and acquisition of library material—but is also far more comprehensive; it also includes the systematic maintenance of library’s collection, covering resource allocation, technical processing, preservation and storage, weeding and discarding of stock, and the monitoring and encouragement of collection use”. (Feather, et al., eds., 2003, p. 81)

Hence, Collection Management (CM) may be considered as ‘a strategy and a process to settle on what to purchase’; ‘a technique to choose materials for purchase’; ‘estimation of what to select or retain based upon worth, usage, cost effectiveness, support of organizational mission, and quality of access’; ‘assessment of how selected materials are used through user input and response’; ‘budget and funding strategies’; ‘marketing strategies and setting of mission, goals, and objectives’.

Library collection management is further subdivided into diverse areas, such as acquisitions, technical processing, usage and making available the archiving of publications.

With the advancement in Information and Communication Technology (ICT), the difference between various groups in collection management department is in its last leg, and work-areas happen to be all the time more inter-dependent. Electronic publishing has opened up new possibilities for business, research, and scholarship. With the expansion of electronic publishing, research and academic libraries have started to complement their print holdings with electronic publications. These changeovers commence with scientific journals, and are now advancing into academic and scholarly books, as well. In the past few years, corporate and government libraries have also begun acquiring e-books along with print holdings. The market for electronic publication seems to be in constant change and collection developers and acquisition managers, especially in libraries, find themselves disillusioned as new methods of electronic publishing and accessing are emerging. Libraries are in a transition period from conventional to digital formats and have not yet developed common practices and traditions on policy for digital material.

2. Definition of the e-book

Nelson (2008) has defined an e-book as, ‘an electronic book that can be read digitally on a computer screen, a special e-book reader, a Personal Digital Assistant (PDA), or even a mobile phone. In other words, e-books are consumed on a screen rather than on paper’.

Wikipedia defines e-book as ‘an e-text that forms the digital media equivalent of a conventional printed book, often protected with a digital rights management system’. General and specialized are two kinds of e-books. The examples of these two are ‘netLibrary’ or ‘ebrary’ and ‘Books 24X7’, ‘Knovel’.
3. Evolution of E-books

A huge amount of articles about development of e-books is available in LIS literature. Sinha (2008) traced a brief overview on the evolution of e-books. She described that in the early 90s, electronic content became popular as publishers sent print books accompanied by floppy disks or CD-ROMs. These added value for students and researchers but caused numerous problems for libraries, from processing to circulation. Ultimately, such issues became doubtful since most librarians were double minded that these formats would endure for scientific and scholarly content. With web-based access, however, it is rare to find a library that does not have significant numbers of electronic resources even as it continues to build its print collections. Over the last 15 years or so, both content providers and libraries have overcome, or at least learned to deal with, the many challenges e-books present in terms of pricing, licensing, access, and workflow. The electronic book, however, in many respects, seems to be more problematic, and more difficult to incorporate into library collections and processes.

The history of electronic publishing dates back to six decades. One of the first non-military applications of the computer technology took place with the project of Father Roberto Busa on the works of St Thomas Aquinas. The American Chemical Society began alerting scholars to new developments through electronic means as far back as 1962. In 1971, MEDLINE was launched and Project Gutenberg issued its first electronic text of ‘The Declaration of Independence’. In 1970s the large business organizations that focused on electronic publications came to the fore. Dialog and ORBIT emerged in 1972 while LEXIS in 1973. Since then digital products have mushroomed throughout the world.

In 1990s, for the first time it became cheaper to publish a reference book on CD-ROM than it did in print format (Lee & Boyle, 2004). Meyers (1996) has given a reference of an article written by Lancaster titled ‘The Evolution of Electronic Publishing’ (Library Trends, Spring, 1995). Lancaster notes that ‘electronic publishing can be considered to have evolved gradually over a period of about thirty years, the evolution having the following manifestations:

1. Use of computers to generate conventional print-on-paper publications;
2. The distribution of text in electronic form, where the electronic version is the exact equivalent of a paper version and maybe used to generate the paper version;
3. Distribution in electronic form only but with the publication being little more than print on paper displayed electronically. Nevertheless, it may have various ‘value added’ features, including search, data manipulation, and alerting (through profile matching) capabilities;
4. The generation of completely new publications that exploits the true capabilities of electronics.

At present, it is safe to say that all four modes of operation are being employed, and experts are experiencing the number 3 and 4 modes of operations.

The International Digital Publishing Forum (IDPF) reports that wholesale trade e-book sales in the United States have increased from $1.6 million in 2002 to $11.6 million by the second quarter of 2008 (IDPF, 2008). Outside the US, as an example, the Digital Content Association of Japan (DCAJ) reported that sales of electronic books increased almost 70% from 2005 to 2006 (Anime News Network, 2007). The e-brary management conducted an international e-book survey in the spring of 2007 to know the e-books subscription by libraries. The survey results showed that out of 552 libraries, 88% owned or subscribed to e-books, with 63% responding that they owned or subscribed to over 1,000 e-books (Mullarkey, 2007).
4. Chronology of e-books

Wikipedia (2011) described the following history of e-books:

1971: Michael S. Hart launches the Gutenberg Project.

1993
- Zahur develops the first software to read digital books. Digital Book v.1 and the first e-book published is ‘On Murder’ considered as one of the Fine Arts (Thomas de Quincey).
- Digital Book offers the first 50 e-books in a Floppy disk with Digital Book Format (DBF).
- Bibliobytes, a project of free digital books online in Internet.

1995: Amazon starts to sell physical books in Internet.

1996: The Project Gutenberg reached 1,000 titles. The target is 1,000,000

1998
- Websites selling e-books in English.


2000: Stephen King offers his book ‘Riding Bullet’ in digital file; it only can be read in computer.

2001: Open ‘Todoebook.com’ the first website selling e-books in Spanish.

2002: Random House and HarperCollins start to sell digital versions of their titles in English.

2005: Amazon bought Mobi-pocket like strategic positioning.

2006
- Sony presents the Sony Reader with e-ink.
- Libre Digital launched Book Browse as an online reader for publisher content.

2007: Amazon launched Kindle in US.

2008
- Adobe and Sony agreed to share their technologies (Reader and DRM).
- Sony sells the Sony Reader PRS-505 in UK and France.
- Books on board is first to sell e-books for iPhones.

2009
- Bookeen releases the Cybook Opus in the US and Europe.
- Amazon releases Kindle2.
- Amazon releases Kindle DX in the US and others.

2010
- Amazon releases Kindle DX International edition worldwide.
- Bookeen reveals the Cybook Orizo at CES.[26]
- TurboSquid magazine announces first magazine publication using Apple’s iTunes LP format, however, this project was cancelled before it reached the market.
- Apple releases the iPad with an e-book app called iBooks. Between its release in April 2010 to October, Apple had sold 7 million iPads.
- Kobo Inc. releases its Kobo eReader to be sold at Indigo/Chapters in Canada and Borders in the United States.
- Amazon.com reported that its e-book sales outnumbered sales of hardcover books for the first time ever during the second quarter of 2010.
- Amazon releases the third generation kindle, available in 3G+Wi-Fi and Wi-Fi versions.
- Kobo Inc. releases an updated Kobo eReader which now includes Wi-Fi.
- Barnes and Noble releases the new NOOKcolor.
PocketBook expands its successful line of e-readers in the ever-growing market.

Google launches Google eBooks.

2011

- Barnes and Noble releases the new Nook - The Simple Touch Reader.
- Amazon.com announces in May that its e-book sales now exceed all of its printed book sales.
- Bookeen launches its own e-books store: BookeenStore.com and starts to sell digital versions of its titles in French.
- The e-reader market grows up in Spain and companies like Telefonica, Fnac and Casa del Libro (the most important Spanish bookshop) launches their e-readers with the Spanish brand bq reader.

5. Collection development of e-books in libraries

E-Book acquisition continues to increase, though most libraries agree that it will be many years, if ever, before their collections become electronic only. However libraries are committed to a long-term electronic goal only. Towards that end, they do not acquire print copies parallel to the electronic versions, except in cases where users explicitly demand print copies. There are a number of sites published primarily by libraries or library organizations that are intended to aid in acquisitions and collection development work. Many library experts like Uta (1998), Herzog (2004), Armstrong (2005) and Twiss-Brooks (2007), discussed collection development issues and challenges about e-books in detail. Uta emphasized a clear collection development policy for e-books. He has listed the following steps of collection development process as under:

5.1. Budget and funds

To make a decision about the subscription renewal or cancellation of the e-books, it is required to establish a budget and maintain record of funds. The cost to acquire and process of resources including the technology costs to provide access (e.g., network requirements, local portal interface, workstations, and bandwidth) shall be considered at this stage.

5.2. Awareness of publication

With new websites being announced each day, it can be difficult to stay informed about available resources on the Internet. So one of the most obvious problem with regard to e-books’ collection development is awareness on daily basis of what is available is sometimes very difficult. Therefore, annotated lists of new electronic publications and information resources by subject librarian are integral and highly appreciated.

6. Evaluation of e-books

E-books must be evaluated with respect to the following technical features and usability aspects:

6.1. Technical implications

E-books should be usable with widely available software like Internet Explorer, Mozilla Firefox etc. If special hardware or proprietary software is required, like in Ebrary or Taylor and Francis e-books, you need user’s training to use this special software for accessing e-books.

6.2. Technical standards

Standards for setting up and maintaining e-resources are evolving and should be applied to e-books in order to allow seamless interconnection with other information systems and to be prepared for technological changes. Standardized Location Description (SLD), metadata (for describing the e-books), and archiving (to guarantee future access) are the required areas for standardization.

6.3. Accuracy and currency of maintenance

The provider of e-books must be well-maintained and should update his website.
regularly. Long-term commitment regarding maintenance of the offered e-books is a critical selection criterion for the acquisition librarians.

6.4. Usage conditions
In process of selection of e-books, we should review the usage conditions carefully. Unacceptable restrictions like no facility of downloading books or time limit selection for reading of e-books can prevent libraries from buying or subscribing to collection of e-books. During the selection process priority must be given to common as well as potential users.

6.5. Prioritization of the publication
Among the available collection of e-books, we must select the collection according to information needs of our library users.

6.6. Purchasing/subscribing to the publication
Obtaining e-books is much more complex than the purchasing of printed books. Access and usage conditions can vary for different publishers and vendors. A number of issues need to be considered.

6.7. Access conditions
Librarians must be aware that ‘obtaining e-book or e-books’ usually means leasing access rights for a given time rather than obtaining physical items. At the end of a leasing period, libraries may be left without anything. Future access to these e-books will depend on the way the librarian archived and made it accessible to his/her users. Archiving should be done by institutions with long-term commitment to this issue, and librarians should advocate a stable and reliable solution.

6.8. Usage conditions
Attention should be paid to the following questions while reviewing the agreement about the use of e-books: How are ‘use’ and ‘user’ defined in the agreement? Will librarians be held accountable for how users use (or misuse) the publication? Can the publication be used for inter-library loan? Will the ‘fair use’ concept be adhered to? Will the publication be archived and if so, who will be responsible for archiving it, librarians or vendors? Is future access to titles published during the subscription period covered by the agreement even if a subscription has been cancelled in the meantime?

6.9. Costs
Libraries may have an option to choose among various pricing models based on different access conditions (number of simultaneous users, total number of logins etc.). Access through library consortia seems to be an advantageous solution. In addition, the costs involved in e-books are not limited to the obvious access fee, be it through subscription rates, on a pay-per-view basis, or through consortia. A reliable and fast internet connection with standard telecommunications infrastructure is required to keep up with the rising demand.

7. Processing of e-books
The cataloging of the e-books is the main point of consideration while subscribing or purchasing e-books. In order to use e-books, they must be retrievable through well-known information sources like library catalogues. It would be most efficient to use existing library procedures to create catalogue records. As is the case with print publications, catalog entries for electronic documents must correspond with existing cataloging formats, for instance MARC, Dublin Core and AACR2 etc. E-books should contain the standard bibliographic details which are typically found on the title page of print publications (author, title, publication date, etc.) as well as descriptions of the content (keywords, thesaurus terms, etc.). In addition to formal and contents-related data, catalog entries for electronic documents must contain information regarding the network address (or network name) of the e-publication, the mode of access (via WWW, gopher, telnet,
etc.), required tools (necessary hardware and software), information on access rights and any bar required as per copyright rules. During assigning metadata applications, one must ensure that information about the document is available to end-users. The examples of metadata systems is the Dublin Core set of metadata.

8. Delivery of the e-books to the reader

Typically, printed books still receive more attention in our libraries than their counterpart’s e-books. How can we announce, present and promote e-books? For those libraries equipped with a web-based catalog, an obvious solution is to use the MARC 856 tag to include a publication’s network address. Users will then be able to directly ‘click through’ to an e-book after having retrieved the catalog entry. The library may also decide to create a web page that lists e-books, preferably by title, author, or editor rather than by publisher as most users hardly ever remember which publisher issues which publication. The availability of newly acquired e-books can be announced in the ‘New Arrival or News’ section of the library’s web pages or distributed to users by e-mail.

8.1. Monitoring usage of the publication

In order to monitor the actual usage, librarians should have access to statistics collected by publishers. Small numbers may indicate that users do not need the resource, or that they are not yet sufficiently aware of its existence. Large numbers can help librarians to justify expenditure on a given resource as well as to argue for more funding.

8.2. Subscription renewal or cancellation

On the basis of data collected from the publishers or e-books vendors, the decision about the renewal or cancellation of subscription may be made by the library acquiring e-books for its users.

Archiving e-books

The topic of archiving e-books still bears more questions than answers. Up to now, it has not been decided who is or will be responsible for archiving, what exactly shall be archived, and where the archived material will be stored. It can be assumed that solutions will vary largely in different countries.

As electronic media can be changed more easily than documents on paper, they are more vulnerable than print publications. Archival data for electronic publications therefore must include the following information (Garrett and Waters 1996):

- Content of e-books (i.e., the information contained)
- Fixity (content stored in a medium as opposed to continuously updated “dynamic” documents)
- Reference (reliable systems for locating and citing)
- Provenance (a record of the document’s origin and chain of custody)
- Context (links within a document and/or between various sources)

Many institutions are trying to find appropriate answers to the question of how e-books can be prepared today for the (still unknown) technology of tomorrow. In this context, we need to address the uncertain durability and future availability of current storage media like CD-ROM and magnetic tape and the respective reading devices.

Deterioration of the storage media is a distinct problem, but an even greater danger lies in the foreseeable technological obsolescence of the equipment needed to use their content. As evolving technologies do not necessarily replace previous ones, libraries theoretically have to provide all kinds of reading devices in parallel and incomplete task.

It is crucial for all libraries as well as information users that access to e-books is guaranteed over time. As archiving e-books is
a very expensive undertaking, small libraries probably will not have sufficient funds to continue archiving publications as they have done in the print environment. National libraries are possible candidates for becoming archiving institutions, probably in collaboration with the publishers or vendors. For instance, national library of Pakistan could act as mirror sites as long as a given publisher/vendors still maintains the e-books. In case the publisher disappears, the national library could solely take over responsibility in order to guarantee future access.

10. Possibility of e-books creation in Pakistan

Higher Education Commission of Pakistan started National Digital Library programme in February, 2004. The databases were mostly focused on science and technology, and included peer reviewed journals and abstracting services. Other resources included subject and multidisciplinary databases, resources in social sciences and humanities. HEC also launched e-brary and McGraw Hill Collections to provide 45,000 online books. The e-books support programme allowed researchers to access most of the important text and reference books electronically in a variety of subject areas. It mentions that ‘the budget allocation for the HECDDL has doubled after each financial year and the total amount for 2005-06 is Rs. 200,000,000 for Pakistan (approximately US $3,278,688) and 5% of the budget is allocated for technical support’.

The full text availability of HEC e-books is not user friendly. It is very hectic for readers to read the e-books. It is also not possible for them to take print or save the same in PDF or HTML format. If we compare the issue with traditional books we may conclude that reading e-book provided by HEC is time consuming, as only some pages may be downloaded at one time. In no way can the users open any chapter directly. At times print facility is not made available. The download facility is available only on the University’s LANs. Some universities have initiated digital project using PDF.

Until the Collection Management policies of universities are not finalized and implemented completely, the following steps may be taken as first step for local creation of e-books.

10.1 Dissertations

Mainly those written in Oriental languages, because currently no system exists in this part of world through which a scholar can check the duplication. Such a system may be established on the pattern of DELNET. DELNET was started at the India International Centre Library in January 1988 and was registered as a society in 1992. It was initially supported by the National Information System for Science and Technology (NISSAT), Department of Scientific and Industrial Research, Government of India.

DELNET has been established with the major goal of promoting resource sharing among the libraries developing a network of libraries. It aims to collect, store, and disseminate information besides offering computerized services to users, to coordinate efforts for suitable collection development. It has been actively engaged with the compilation of various Union Catalogues of the resources available in member-libraries. It has already created the Union Catalogue of Books, Union List of Current Periodicals, Union Catalogue of Periodicals, CD-ROM Database, Database of Indian Specialists, Database of Periodical Articles, Union List of Video Recordings, Urdu Manuscripts’ Database, Database of Thesis and Dissertations, sample databases of language publications using GIST technology and several other databases. The data is being updated in each of these databases and is growing rapidly. In this context, DELNET is very much fruitful. It not only provides access to bibliographic information but also avoids duplication of effort.

Higher Education Commission of Pakistan is the authorized funding institute in Pakistan to
grant funds to the universities. It should provide maximum funds to the varsities of Pakistan so that they may digitize their collections for the feasibility of library clients. HEC may also establish a union catalogue of all important libraries collections of Pakistan to meet the challenges of the present age that is the age of Information and Communication Technologies (ICTs). In this context, African Journals Online is a fine example that may be followed by HEC Pakistan to provide an easy access to all library clients. Recently, HEC has uploaded about four thousand dissertations in its website to provide an easy access to researchers to complete their research projects but there is also a dire need to establish a Union Catalogue on the pattern of DELNET so that literary works may also be maintained side by side with the uploading of thesis and dissertations. By Union Catalogue, duplication of work and plagiarism may also be prevented and research culture may flourish.

10.2. Rarely held materials
A large number of books available in the libraries of public sector universities have been written in Urdu, Persian, Arabic, and Sanskrit etc. which were published before 1900. Many titles published in English may be considered in this category. After digitization, many titles may be exchanged with the libraries of western countries.

10.3. Unique manuscripts
Many manuscripts written in Arabic and Persian languages may be exchanged with Arab countries and Iran. More than twenty thousand manuscripts are available only in the University of the Punjab Library, Lahore. Many public libraries also have a similar number of manuscripts.

10.4. Miscellaneous
The libraries of various universities may also digitize books recommended for higher degrees courses. The process will not only preserve the valuable books but also be helpful for users on low cost. However, in this category of books fair-use approach as per provisions of copyright law of Pakistan shall be observed.

Government College University (GCU) Lahore recently completed a digital project of 4000 letters written by different literary personalities. Digitization of The Ravi through Green Stone software is also in progress. History of Govt. College Lahore 1864-1964 has been shaped into digital form. However it is available with contents in PDF format. We are of the view at GCU to use open standards (HTML, ASCII, JPEG, PNG, XML, etc.), as the library aims to make information available to the widest audience possible. At this stage there is a dire need to develop growing network/consortia among the universities.

11. Major Challenges
As awareness and sale of e-books mark a significant escalation, challenges for all kinds of libraries abound, price of e-books is an overriding factor in making decisions pertaining to Collection Management no matter whether they are purchased as collections or title by title. Libraries, which are in a process of transformation, bear the brunt more than those who have already put into practice policies that do not consent selectors to replicate formats, purchasing the e-book, and print book of the same title.

A few e-book platforms have been set-up in recent past which are user-friendly. Yet the cardinal quandary remains unaddressed i.e., there are too many platforms to choose from. It poses a serious challenge at the phase of short-listing the required platforms. While it is imperative to focus on subject-matter, libraries are restricted in their choice of e-book titles on account of the platform or aggregator they may have earlier chartered. A sole and effortless access model for e-books is also considered necessary, particularly for the aggregators who have a single-user or two-user models. As far as libraries are concerned, such options are unfeasible since they rout the rationale of online access to
subject-matter. Such restrictions are essentially retrogressive in the sense that they date back to the early days of electronic resources when content providers deemed it mandatory for users to be actually present in a library since it was not possible for users to access content in a library from their desktop across the lane or in a remote part of the world.

A large variety of price and purchase models offered by publishers and e-book aggregators have weighed down the libraries. While they prefer to be the sole owners of their purchased materials, the publishers and aggregators offer numerous choices including subscription or lease, purchase to own, perpetual access, where ‘perpetual’ is to be determined by the publisher. There are chances that the access could be denied utterly if e-books get removed from an aggregator platform, in case publishers set their mind to it. As a final point, while moveable readers grow to be all the rage, libraries have to determine the role they have to perform in academic, specialized, and other libraries. Per chance these are most well-matched for public libraries but readers expect to have access to scholarly and technical book material on their PDAs, Blackberries and cellphones. One such example is found in Japan where mobile phone e-book sales shot up to more than 300% in 2006 with no indications of a downward trend. (Anirne News Network, 2007). So far as the aggregator platforms are concerned, the interest of authors, publishers, or those producing devices on long term basis will not be served by limiting e-book content to specific devices. Unrestricted access ought to be the goal; devices must not limit as to which book readers need to access or not.

12. Conclusion

E-books have a bearing on all spheres of library management, and traditional collection management should be improved to stay valuable. It is not possible to adopt a uniform set of rules pertaining to electronic publications which can be equally applicable to all libraries; each librarian has to find indigenous way to facilitate local users. A relatively market-oriented approach in evaluating information resources has been adopted on account of the electronic environment. Librarians can decide from among an assortment of access conditions while dealing with e-publications. This newly found elasticity paves way for an extremely augmented liability for the individual librarian.

While it is imperative to discover exclusive solutions, we ought to prevail over our isolation, and engage in a discourse with our colleagues so as to gain knowledge of other methods to manage a swiftly shifting information environment. We can benefit from the solutions our contemporaries have already practiced. It will also enable us to cleave to a stronger position while negotiating with vendors and publishers. Joint efforts can reap better dividends in case we put into operation user-friendly access conditions.

Scores of matters pertaining to electronic publishing still require thrashing out. Most publishers have not yet given up trying out promising solutions. Thus, they are keen to lend a hand to librarians. It can only be helpful for the entire community of information providers, mediators, and users, if e-books are dealt with thoughtfully.

Inclusion of e-books as an essential part of library collections has turned out to be quite beneficial in recent years. Whereas libraries were initially skeptical about the utility and desirability of this format for library users, and were reluctant to relinquish Collection Management activities to publishers and aggregators offering E-book collections, libraries are more frequently engaged in appraising print and electronic books simultaneously for developing subject collections. Now that the publishers offer the electronic editions along with print editions, the task to develop e-libraries has become easier. Now-a-days, vendors - in partnership with publishers and e-book aggregators - distribute electronic books by regulating and boosting their existing operations and services to accommodate this new format rather than the previous practice according to which they traditionally distributed print books.
via approval plans and direct orders. Does it imply that with these recent managements in the realm of electronic books and an increasing number of options for libraries the printed books will no longer be in vogue? Certainly not. Electronic books merely offer an alternative. They cannot swap print book collection entirely. In fact, e-books fulfill the needs of some users and are suitable in a few situations and disciplines but not all. As in the past, libraries and librarians will go on responding to the wants and needs of their users, offering an array of resources in manifold formats.

References/Bibliography
Christopher C. Brown and Elizabeth S. Meagher, metadata and materials processing librarian, are preparing for publication an article on the process of gathering usages statistics for free online resources through the library’s OPAC.


Integrated Pest Management

Strategies for Onion and Garlic

R K Mishra • Alok Adholeya • H R Sardana

2012 • ISBN: 9788179934968
Pages: 88 • Binding: Paperback
Size: 140 × 210 mm • Price: ₹395.00

Onion (Allium cepa L.) and garlic (Allium sativum L.) are the most important Allium species cultivated worldwide and are used as vegetable and spice in our daily diet. These crops are widely cultivated for domestic consumption as well as for export purposes. They are attacked by many insect pests and diseases, which vary with region, season, and variety. This lowers the quality and yield, thereby increasing the cost of production and reduces export potential.

Integrated Pest Management: strategies for onion and garlic discusses methods and tools used to minimize the incidence and severity caused by diseases and insect pests. It also focuses on the symptoms of diseases caused by various pathogens.

Table of contents
- Introduction
- Integrated Pest Management: key strategies
- Major Diseases and their Management
- Major Insect Pest and their Management
- Nematodes and Mites and their Management
Abstract
The world is witnessing a considerable transformation from print to electronic-based formats thanks to advanced computing technology. This has had a profound impact on the dissemination of nearly all previous formats of publication into digital formats on computer networks. Text, still and moving images, sound tracks, music, and almost all known formats can be stored and retrieved on computer magnetic disks. Over the last two decades, a number of special libraries and information centres in the Arab world have introduced electronic resources into their library services. Very few have implemented automated and integrated systems. Despite the importance of designing digital libraries, not merely for accessing or retrieving information but for the provision of electronic services, hardly any special library has started the design of digital library services.

The aim of this paper is to provide a general guideline for the design of a low-cost digital library that provides services that are most frequently required by various categories of special library users in developing countries. This paper also aims at illustrating strategies and method approaches that can be adopted for building such projects. Seeing the importance of designing an inexpensive digital library as a basic principle for the design, the utilisation of today’s ICTs and freely available open source software is the right path for accomplishing such a goal. The paper intends to describe the phases and stages required for building such projects from scratch. It also aims at highlighting the barriers and obstacles facing Arabic content and such problems could be overcome.

Keywords: Digital libraries, Electronic library services, Special libraries, Greenstone system, Open source software.
1. Introduction

This research originates from a belief that information services for special library users in Libya are poor and not well developed because they are suffering from a general weakness in their principal foundations, equipment, ICT infrastructure, resources, and cadres, making them unable to comply with their duties and obligations towards their end users. Developing special library services and activities in Libya should be achievable through implementing and exploiting today’s technology. Therefore, it is anticipated that designing a low-cost digital library service and introducing electronic services will assist in solving a great deal of problems, in addition to meeting the needs and requirements of the users. Furthermore, relying on conventional technical procedures in libraries in general, and in special libraries in particular, has proved to be the wrong path to success. For that reason, a shift to information technology and its applications would seem to be essential to give these institutions the capability of providing reliable services to targeted end users. It is also important to point out in this context that information and communication technologies present new opportunities and challenges for libraries in general and special libraries in particular. Special libraries worldwide are passing through a phase of dramatic change due to the on-going challenges being posed by information and communication technologies. Accordingly, they are challenged to explore new ways to accept and implement the changes made by IT to be fit enough to serve experts, scientists, and research workers, who always demand perfect, faultless, and accurate specialised services. Most scientific information today is available in a variety of formats such as CDs, DVD, the Internet etc., which requires knowledge of specialised information handling techniques. Therefore, qualified well-trained staff and modern technological equipment should be available to boost technical procedures and services to an effective level, and to enhance and foster the productivity and expansion of research activity. It can be assumed that special libraries in developing countries need to modernise and implement their ICT infrastructure, and articulate information policies that assure the optimum exploitation of information resources to increase national productivity. Improving IT education and research environment may be expected to create new business and facilitate industrial research. This in turn will foster related human development and enhance the total economy. Seeing the importance of today’s technology and its capability for improving library services in addition to realising the recommendation of the World Summit On Information Society (WSIS, 2005) regarding exploiting modern ICT to foster productivity and expansion of research activity, the researcher come to the decision to build a low-cost digital library for the National Oil Corporation as a module.

2. Digital libraries

There is no doubt that there are many different views in the literature as to the actual nature of digital libraries. This paper does not intend to provide a comprehensive collection of definitions of the digital library, but rather a number of representative definitions. Various terms are still used interchangeably worldwide such as electronic library, hybrid library, library without walls, cyber library, virtual library etc. Arms (2000) views digital libraries as “managed collection of information with associated services, where the information is stored in digital formats and accessible over a network.” Witten (2003) defines the digital library as a focused collection of digital objects, including text, video, and audio along with methods for access and retrieval, and for selection, organisation and maintenance of the collection. The digital library federation (DLF) define digital libraries as “organisations that provide the resources – including the specialised staff –
to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.” According to Jeng (2006), a digital library in the 21st century has the following characteristics:

- It is an organised and managed collection of digital objects
- It is accessible over the internet or servers
- It has global information infrastructure and;
- Should offer service.

The last point indicates that there is a difference between a digital collection and a digital library, in that a digital library should offer service to end users. Therefore, a digital library is considered a collection of information objects and a collection of services that should be provided by the digital library. “The definition of a digital library that came up in March 1994 in Digital Library Workshop emphasised that a full service digital library must accomplish all the essential services of traditional libraries and also exploit the well-known advantages of digital storage, searching, and communication.

(Chowdhury, 2002) Leiner (2009) report that “There are a large and varied set of services, including services to support management of collections, services to provide replicated and reliable storage, services to aid in query formulation and execution, services to assist in name resolution and location, etc.”

3. Design Framework

This section focuses on the design of a proposed low-cost digital library for the National Oil Corporation, which is one of the leaders in the energy sector in Libya (in fact all local and foreign companies in the oil sector in Libya run their business under the NOC). It discusses the methodology used for designing the proposed digital library, the framework, and the steps implemented to reach the final goal. Designing a digital library usually comprises several stages. These phases are shown in the following flow diagram.

4. Objectives

The first crucial step is to decide the objectives of designing the digital library. This step has to be considered early – before any further steps are taken. The whole project will be worthless

![Figure 1 Flow diagram](image-url)

without clear objectives. The objectives of this design are summarised as follows:

- Developing the standard of services presented to targeted end users to enhance the level of research activities within the organisation.
- Using and putting into practice, modern ICT to improve the current level of services and to solve the problems that the central library at the National Oil Corporation faces.
- Producing a design for a low-cost digital library that provides not only access to and retrieval of information, but also the services that are most frequently required by the users.

5. Methodology

An experimental system design approach was used for the designing of the proposed digital library. The BNET dictionary defines this approach as “the planning of the procedures to be used in an experimental study” (BNET dictionary). This approach was applied because it was found to be the most appropriate method for such projects. The design of a low-cost digital library encompasses various stages. The phases that follow are part of the system design approach. Each phase consists of a number of stages, each stage is subdivided into steps, and each step further contains a series of tasks. Thus, the total work is broken down into manageable portion.

6. Planning

The planning stage comprised a survey of the current state of special libraries and information centres in the energy sector. The energy sector was chosen due to its significant role in supporting the Libyan economy, which relies mainly on the revenues from the oil sector, which account for almost all of Libya’s export earnings and about 25 percent of its GDP. Oil is currently said to be providing the government with its main source of revenue and constitutes 99% of Libya’s exports. The National Oil Corporation was chosen for this model from a number of local and foreign oil companies in the country because the NOC is considered the corporation under which all oil enterprises in Libya run their business. In fact, all oil has to be sold through the NOC, which carry out marketing operations of oil and gas, locally and abroad. For this purpose, the NOC has its own fully owned companies, which carry out exploration, development, and production operations, in addition to local and international marketing.

The planning phase is considered a fundamental process of understanding why a digital library should be built and determining how it should be built. Questions examine economic, technical, and organisational feasibility i.e.

- What exactly is the project? Is it possible? Is it practicable? Can it be done?
- If we build a digital library, will it be used?
- Economic feasibility:
  - Are the benefits greater than the costs? If benefits outweigh costs, then the decision is made to design the library.
- Technical feasibility:
  - Do we have the technology? If not, can we get it?
  - What is the current scenario of ICTs within the organisation?

The aforementioned questions should be answered in this phase to determine the expected value of designing a digital library service. A feasibility study is a short assessment of the proposed digital library to determine whether the projected library can effectively meet the specified business requirements of the organisation as well as to identify constrains, budget, time, ICT infrastructure, and resources.

The first and second surveys, which comprised of questionnaires and interviews with librarians and managers of information units in a number of oil companies, have focused on the following:

6.1 Policy in relation to:
- Computers
- Networks
• Internet: Type and speed
• Intranet

6.2 Policy related to content, focusing on
• Electronic resources
• Budget
• Access control

6.3 Policy related to management:
• Familiarity
• Willingness to provide efficient and effective service
• Vision: e.g. visions with regard to developing ICT policy.
• Plans: E.g., any plan for moving towards an electronic library service.

6.4 Policy related to training
• Users’ training
• Librarians’ training
• Cost and type of training programmes

Seventeen libraries were surveyed to get information pertaining to current state of services. The first and second surveys, which were conducted in July 2007 in order to reveal the state of library services in different sectors, showed that special library services in Libya are poor and not well developed. Therefore, designing low-cost digital library services may possibly help in solving some of the problems faced by special libraries, and strengthen the services presented to targeted end users. The survey also revealed that special library users are eager and willing to use electronic library services, and the demand for such services is enormous. Because special libraries are supposed to provide services for experts, scientists, and research workers, the design of a low-cost digital library could assist in providing services that satisfy users’ diverse needs and requirements. Regarding ICT infrastructure in the energy sector, the survey has revealed that this sector is well equipped with ICT infrastructure. Therefore, designing a low-cost digital library is likely to be much easier within the existing ICT infrastructure because the way for such developments is already paved. Additionally, the number of computers connected to the Internet in corporations attached to the energy sector is higher than in other sectors. Therefore, it is hoped that the design of the proposed digital library will be successful and users will benefit more from the services, which will be offered through the digital library.

Computer literacy is also considered a crucial issue. Establishing the feasibility of developing the proposed digital library was the main concern of this study. The survey explored how end users use computers and the Internet, as both tools are considered the backbone of today’s electronic services. In terms of the capability of users to use the proposed digital library, the responses indicated that the chance of successfully exploiting extensive information resources through the digital library is quite promising. Firstly, in terms of use of computers, responses showed that half of the users surveyed (50%) had a wide knowledge of computers, and under a fifth (17%) considered themselves as experts in the use of computers and associated software. Only a tiny percentage (2%) of special library users reported a total lack of knowledge in the use of computers. Secondly, in terms of use of the Internet, responses showed that over half of all users surveyed (56%) had a wide knowledge in the use of the Internet and under a fifth of them (15%) considered themselves as experts in the use of the Internet. Only a small percentage (6%) of special library users reported that they possessed absolutely no knowledge of the use of the Internet. These figures indicate that the proposed digital library may be a great opportunity for special library users and the creation of electronic services may be beneficial to the targeted users.

As regards to policy related to computers, networks, the Internet, intranets, and access controls, the survey revealed that oil companies in general, and the NOC in particular, had a
written policy. Some oil companies have issued and released specific resolutions to regulate the use of the Internet, intranet, and access control. Management policy was also investigated relating to the following:

• Familiarity with the technology.
• Willingness to provide efficient and effective services.
• Vision: with regard to developing ICT policy.
• Plans: for moving towards an electronic library service.

Responses revealed that a number of oil companies, and the NOC in particular, had plans to provide digital library services in the near future. The scientific library of the Arabian Gulf Company for oil exploration has established a new project for the establishment of a technical database. The central library of the National Oil Corporation will have, as soon as possible, a link to the parent organisation’s website. Librarians at the central library of the NOC have decided to provide electronic services such as current awareness and SDI services, in addition to online services through this link. It can be said, therefore, that there is a trend among oil companies towards moving to electronic library services. However, libraries in oil companies are still in the early stages of a move towards electronic services.

As regards policy related to user training, librarian training, types of training courses, and the costs of training courses responses showed that just under half of the selected libraries (47%) had a document regarding the training of human resources and the type and cost of training courses.

7. Requirements Analysis

Developing services that meet the expectations of users and customers is critical to success. Requirements analysis is the foundation of a user-centred approach, and helps to create projects and services that appeal to and meet user needs (D’Hertefelt, 2009). User requirements analysis is not about asking users what type of services they want. User requirements analysis is about understanding users’ current practices and the problems they encounter. This stage is comprised mainly of the analysis of fieldwork, which has been done to allow an availability check of the required components needed in the design of the digital library. Requirement analysis shows what elements and components

![Figure 2](image-url)  

**Figure 2** How knowledgeable are special library users in the use of computers
are necessary for the proposed digital library. This stage should provide answers to questions such as:

- For whom is the digital library going to be designed?
- Who are the key users?
  - Administrators – Internal research staff –
  - External researchers – Academics - Other
- How many users are going to benefit from the proposed digital library?
- What is the actual status of the ICT infrastructure in the organisation?
- What type of material is already in digital form?
- What sort of materials need to be digitised?
- What items are distributed within the organisation or the institution?
- What items or resources are outside the organisation or the institution?

In brief, the requirements analysis should also indicate the need for the following:

- Internal documents which include:
  - Reports
  - Statistics
  - Standards
  - Technical data
  - Publications by people within the organisation
- External documents which include;
  - Purchased materials (in either hard copy or electronic format)
  - Documents obtained from other resources

### 8. Analysis

Three semi-structured interviews were conducted to gather information related to ICT infrastructure currently used IT applications, and information resources and services in the energy sector. The first interview was conducted with the Manager of Information Unit in the National Oil Corporation (NOC) and the second interview was conducted with the librarian responsible for running the central library of the NOC. The third interview was conducted with Manager of System Support in the same organization. These semi-structured interviews have assisted in creating the digital library and getting the aforementioned questions answered.

The intention of these interviews was to identify the most needed and frequently used resources, most common problems, etc. The result of these interviews was a list of the most frequently used resources and common problems. In addition to this, it was possible to recognize how users would prefer to use these resources, and what approach was currently being used for the retrieval of information resources. From this, it was possible to establish the metadata and the requirements needed for designing a digital library in this situation. The interviews conducted also focused on identifying frequently used resources and the format of these resources or documents, whether these resources were in hard copy or in electronic form, and whether they had specified internal resources, which would consist of reports and publications, created by employees within the targeted organization. External resources were also identified and the means used for acquiring these resources were identified. The abovementioned interviews assisted the developer in creating the proposed digital library in an effective way, as many questions were answered at this stage.

As regards the first and second questions, the key users of the proposed digital library were identified as the NOC administrators, employees, internal research staff, and external researchers in addition to academics. The number of users who maybe benefit from making use of the services, which the library will provide, was also identified as 1000 users at least. As regards the actual status of the ICT infrastructure in the organization, it was recognized that ICT was used for a range of office work and other applications in the organization. These tasks generally included e-mail, word processing, spreadsheets, and database management systems. The central library of the NOC had good hardware facilities including scanners, printers, and CD-ROM towers, etc. The
The central library of the NOC owns a considerable number of digital materials. These digital resources can comprise the collections of the proposed digital library. The digital-born materials are in various file formats and cover a wide range of resources and documents on various subjects such as reports, production reports, maps, NOC publications, technical data, statistics, and publications by people within the organization. There are also considerable numbers of documents, which need to be digitized so that users can have easy access to these materials on the digital library. The manager of information unit in the NOC, in accordance with the company’s copyright regulation, asked the developer to restrict and control some materials such as theses and dissertations from being accessed on the library by external users without permission. This could be done by allowing only authorized users to view and browse such a collection.

9. User Studies

User studies are considered one of the most important activities in the design of a digital library (Dawson, 2009). The first survey focused on special library users and highlighted a considerable demand for information resources. The survey focused on what users really want and what they actually get. Research results have also shown that there is a huge gap between supply and demand. It was obvious that special libraries in Libya were not in a position to offer a good level of services due to the problems they face. Therefore, designing low-cost digital library services is expected to improve the services presented to end users and to satisfy, to a large degree, the diverse needs of end user.

10. User Requirements

The first survey was conducted in order to shed light on the status of special libraries and the level of user satisfaction with the offered services. The survey results revealed that users seek out heterogeneous resources. Reliance on homogeneous resources, which comprise one sort of information materials i.e. print materials, proved to be insufficient to satisfy user requirements by information specialists. Consequently continuing to provide homogeneous resources has affected the services in a negative way. The survey findings also revealed a scarcity of electronic materials. Therefore, it is essential to start providing heterogeneous content that comprise various kinds of materials and information resources. Electronic services constitute a new horizon for special library users that should be rediscovered in order to find out what is new in a world that expands every day with innovations. Provision of heterogeneous information resources as well as materials suitable for teaching, learning, and research should enrich research activities in many ways and should improve the condition of the current services. Moreover, access to a multi-content digital library should contribute to the development of on-going research and supply end users with new possibilities. In order to recognize what sort of information resource is most required from a user’s viewpoint, the first survey included a question on this particular issue. The responses showed that electronic journals ranked first (95%), followed by electronic books (90%). Official publications and reports were also ranked at the top at 95% and 90% respectively; statistical data was also popular at 90%. Newspapers and patents ranked last at just 40% and 50% respectively with literature reviews at 60% and specifications at 65%. The figures also illustrate other information materials such as literature reviews, standards, and specifications whose demand goes up and down according to user preferences. As previously stated, there are various information resources but only some of them were found to be essential and, from the users’ perspectives, to be obtainable in electronic form in a digital library.
10.1 Design

Once the requirements analysis is concluded, the next step is to design the proposed digital library. The design phase decides how the system will operate, in terms of hardware, software, and available infrastructure. The previous mentioned interview, which was conducted with the system support manager in the National Oil Corporation, revealed that the ICT infrastructure in the energy sector in general and in the NOC in particular has reached a good level and that this sector is well equipped with ICT infrastructure. Therefore, designing a low-cost digital library is likely to be much easier within the existing ICT infrastructure because the way for such development is already paved. The first step in the design phase is to develop the design strategy or the framework. In fact, there are several points that need to be considered before starting the physical design of the proposed digital library. These are as follows:

11. Library Purpose and Philosophy

The purpose of designing the proposed digital library originates from the actual need of users to have access to information resources that support their research work and everyday jobs. This was evident from the first survey and the lessons learned which have shown that there is great need for designing electronic services to meet user needs and requirements. The proposed digital library should support research activities in different ways. The digital library must be accessible to all research workers and company staff wherever they are regardless of their location. This means that company workforces, either in the main office or in regional branches, should benefit from accessing available resources any time, 24x7. For example, workers in the oil sector need access to seismic technology. Because of the importance of seismic technology, its data needs to be available all the time for research workers in oil fields and engineers working in oil fields who will benefit from accessing the required data through the digital library, so long as they have the basic infrastructure. Since oil fields are already connected to the Internet through satellites; accessing, browsing, and searching the digital library will not be a problem in this sense. Users who want to have access to sensitive materials such as the data centre collection (technical library) of the corporation should also have the ability to access any sort of document from anywhere in the country regardless of their location (security settings allowing). The digitization of the corporation’s technical reports and production reports will also be of great importance as it will preserve and safeguard sensitive documents from loss and damage and at the same time will mean they can be accessed in the most convenient way possible.

12. Software

There are various types of software available for the creation of the proposed digital library. However, because the main principle of this research is to design a low-cost digital library, selecting software from diverse, freely available software is a complicated question. Moreover, due to the availability of a variety of open source software that could serve the purpose of the proposed library, selection of appropriate software was a vital issue. Greenstone software was selected due to its good reputation and its support for different languages, which is an essential aspect for the design of bilingual content. In addition, Greenstone is capable of displaying the user’s interface in multiple languages and handling collections of text, pictures, audio, and video files. The program also offers flexible browsing facilities and can run on Windows and UNIX. Despite the availability of various open source software programmes required for building digital libraries such as Dspace, (jointly developed by MIT Libraries and Hewlett-Packard Labs), Eprints, Open repository, Open publication system, Fedora, DigiTool, etc., many organizations and institutions worldwide have built their digital libraries using Greenstone software, for example:
12.1 iArchives
The iArchives is several demo collections of digitized newspapers that have been set up for clients, including British Library Newspapers, the Californian, and the Dublin Evening Telegraph.

12.2 MOST Digital Library (UNESCO)
The MOST Digital Library contains results from research carried out during the first ten years of the MOST Programme. The themes covered include drugs, globalization and governance, international migration, multicultural societies, poverty eradication, and social transformations.

12.3 Oxford Digital Library
A dozen or so collections on an eclectic variety of subjects: ornithology and entomology, prints and maps, educational games from the 18th and 19th century, classic geological literature, medieval science, and many others.

12.4 National University of Science and Technology
A prototype Greenstone digital library for the National University of Science and Technology in Zimbabwe.

However, the most important reason for selecting Greenstone software lies in its capability for building bilingual content, as this feature is not available in any other open source software. None of the previously mentioned software is capable of building Arabic content or displaying the user interface in Arabic. Greenstone software is capable of that because the software is supported by UNESCO, which encourages developing countries to participate in the current information revolution by adopting modern ICT systems. Designing a bilingual digital library that could support both Arabic and English languages was a crucial aspect due to the following:

- English is the second most used language in Libya.
- Most of the literature written in any branch of knowledge can be found in English.
- A considerable number of research workers, especially in the energy sector in local and foreign companies in Libya, are non-Arabic speakers.
- A substantial number of Libyan research workers are acquainted with English.

13. Testing Greenstone
The objective of this test is to ensure that each element of the application meets the functional requirements required for the proposed digital library. The first phase of testing carried out to insure the systems installation and support group, prior to implementing the system. The second phase of testing carried out to prove that all areas of the system interface with each other correctly and that there are no gaps in the data flow. The third phase of testing carried out to ensure that the system provides acceptable response times for retrieval of information. In this test, some documents were retrieved from the system and the response times proved encouraging. Greenstone has two separate interfaces, the librarian interface, and the user interface. Both interfaces were tested to make sure that the system runs as it should.

14. Content
As content is considered the key to success for any digital library, further attention was given to this issue in order to guarantee the achievement of research goals and the success of the proposed digital library. As there are a number of foreign research workers and employees in the NOC it was decided to create bilingual content in the form of Arabic and English collections. It was also decided to display the user interface in two languages so that users could easily browse and search the constructed library in both languages. This means that the proposed digital library can serve users whose first language is English (native English speakers) and users whose first language is Arabic (native Arabic speakers). By so doing, the proposed library should be highly
effective and efficient in providing a variety of electronic services to end users.

15. Content Creation
This step covers the digitization processes and selections of materials and decisions, in addition to the subsequent manipulation and management processes (Dawson, 2003). Various decisions have to be made on various aspects before starting the actual design. For example, should optical character recognition (OCR) be applied to create text files and should files be held as images and text. If OCR is found to be technologically and economically feasible, the output will have to be checked carefully before dissemination. Choice of formats was another challenge, so it was determined to choose the format according to the nature of original documents. Decisions also have to be made on aspects such as whether to use HTML or rich text to display plain text. For this, it was determined to use HTML for displaying plain text. Other issues related to PDF files such as whether to use PDF or other plug-ins such as the Microsoft Word plug-in for displaying documents. A PDF plug-in is important because it is ideal for retaining complex layouts. There are also various formats for image and even more choices. Whilst JPEG is a common format, PNG is spreading quickly, and TIFF can be possibly recommended for printing or preservation. (Dawson, 2003).

16. Digitization
Before starting the digitization process, it was crucial to find out the types of available information resources that would make up the collections. Therefore, the interviews conducted with librarians and people at the central library of the NOC focused on this matter. Suggested collections were divided into two groups. The first set comprised those already in electronic form (which were created as digital files) or those converted to digital form (for example scanned documents). The second set comprised traditional documents (printed documents). Suggested documents were also divided into Arabic and English documents.

17. English Collections
Materials in English may already be in digital form, but in different formats, e.g. PDF, rich text, and MS Word. It is worth mentioning in this context that a decision has to be made before starting to design the digital library whether to display the content of the library in the original format of the document or in HTML format. It is desirable to display the library’s content in more than one format to give users the choice of viewing the content in various formats. PDF format is popular for downloading purposes (Dawson, 2003). Traditional print documents, scanned using OCR, are easily displayed in HTML without any problem. However, the Greenstone system is not capable of retrieving such materials when a search is performed. To overcome the problem all materials converted into PDF using OCR were transformed into searchable PDF using a transformation programme such as ABBYY PDF Transformer 2.0, Adobe PDF professional, or Nitro PDF professional. After transformation, Greenstone software was able to retrieve target content when a search was performed. Regarding documents in rich text and MS Word, no problems were faced in displaying such documents as HTML.

18. Arabic Collections
Considerable problems were faced in the process of building Arabic collections. In fact, the problem with the Arabic language is related to the poor performance of OCR systems in this domain. Born-digital documents in the Arabic language are easily displayed and retrieved as HTML by the Greenstone system when a search is performed. However, either documents scanned and converted to PDF or MS Word can be neither displayed nor retrieved by Greenstone.
Moreover, in this case the Greenstone system is not capable of performing a full text search. This means that the Greenstone system is unable to extract words and terms occurring in the text if digitized materials are entered as scanned images. Despite the availability of many specialized OCR programs that claim the capability of converting Arabic from digital image to digital text format, such as Readris 11, none of these programs is capable of dealing with Arabic characters and of converting documents 100 percent. Experience has shown that no programme in this domain is able to give reasonable results, as the percentage of success did not exceed 50 percent. Even the retrieval of digital-generated documents in PDF format with this system is difficult when a search is performed, because they are simply not searchable. The Greenstone system cannot recognize and search characters that are not written in ASCII code, which is a standard code used so that data can be moved between computers that use different programmes. It is worth mentioning in this regard that born-digital documents in Arabic can be displayed and retrieved by the Greenstone system provided they are in MS Word, rich text, or searchable PDF. To overcome problems associated with the retrieval of objects in PDF format, documents of such types were carefully reviewed and more attention was given to their associated metadata so that users could exploit enabled browsing facilities instead of search facilities for retrieval purposes. Developers should therefore avoid scanning Arabic content using OCR systems. Nevertheless, conversion of Arabic content already in digital form from rich text or MS Word to PDF proved to be practical. As regards the format for image, there are also various formats for image and even more choices, as mentioned previously. However, for the NOC digital library, the JPG file format was used for images.

19. Information Organization: Classification and Indexing

In order to make searching more reliable, and allow users to browse across collections, there is a need for controlled methods of information classification. Therefore, the Library of Congress Subject Headings was proposed as a means of linking diverse collections into a coherent information structure, in addition to Arabic subject headings as a primary means for linking collections in Arabic. The terms of both subject headings needed to be included in the metadata and were used in contexts where international compatibility was required. In order to keep consistency among collections, controlled terminology for place names and organization names, as well as for technical terms, is necessary where appropriate as authority files, e.g. names of oil fields, names of concessions, and names of areas and locations.

20. Metadata

Regarding metadata, the question of which metadata standard to use was a key issue. Because of the availability of diverse metadata standards, there was a need to find out which standard could serve the purpose of building the collections of the proposed library. The Dublin core metadata standard was expected to meet the requirements of building the required collections. However, the standard was tested by retrieving some specific files of the content of the digital library. Some metadata elements therefore had to be added to enhance retrieval process and to correspond with the methods of information retrieval by users. Metadata elements were added for example to the collections of technical library (data centre collection) to correspond with the method of information retrieval by users, e.g. concession and location.

21. Information Retrieval

As web users have become used to the simplest possible search interfaces (Google interface for example) and very fast results, there is no need in the short-term to provide complex search facilities for the proposed digital library. Enabling users to search fields such as title, creator, and subject terms, and then to search other metadata fields if
no matches are found, then continue with full-text searching or cross-searching only if no matches are found in any metadata fields, was considered sufficient at this stage. It is worth mentioning in this context that Greenstone software enables users to search collections in different ways. Specifically, the software offers users the capability of doing advanced searches across different collections (cross-collection search) by adding a mixture of collections before performing a search.

22. Authentication or Access Control
This issue is concerned with policies for controlling access to different types of content; therefore, it raises different questions. For example, should all digital content be accessible to all users across the Internet? Is it necessary to control who can access specific collections and make use of it? Should the content of the proposed digital library contain sensitive documents that would harm the company business if they were accessible over the Internet? These issues were discussed and studied before starting the actual design of the digital library. Greenstone software has a built-in access control mechanism, which allows collections, and even individual documents, to be restricted to authorized users using a password protection scheme. This mechanism can be applied if the company prefers to apply restrictions to some specific content. A policy for controlling access to different digital content should be written by the managers of information units in companies seeking to build digital libraries, who should take into consideration the negative and positive impact of open access. According to the recommendations of the manager of the information unit in the National Oil Corporation, it was decided to protect some sensitive collections such as the data centre (Technical library) collection, the thesis collection, and e-journals.

23. Standards
Initially, the core standards applied to the proposed digital library were Dublin Core and MARC 21 for metadata; Library of Congress subject headings, and Arabic subject headings for subject vocabulary; AACR2 for resource descriptions; JPG for image. These standards are important for the consistency of resource description across collections and for assisting information retrieval.

24. User Interface
The user interface is considered an important issue and it is an essential element of the proposed digital library. The design of the prototype had to take into consideration simplicity, consistency, and flexibility. The main issue before starting the design of the user interface is to understand the main priorities from the user’s perspective. Therefore, how the interface should be designed depends on user opinions and can be decided after conducting some interviews with people in charge and with a number of potential users. Because the user interface can be designed in different ways, it is crucial to understand primary preferences in order to produce a practical design. The second survey and the interviews carried out with librarians, and some potential users from the NOC, shed light on their preferences. This was an essential step before starting the design of the prototype. The most significant issue that was faced within the design stage was in relation to producing a practical design with two different interfaces. On the one hand, it was crucial to design an Arabic and English interface to display collections in Arabic and English. On the other hand, it was also important to build the same collections in a flexible way that could be displayed exactly the same, on any interface. In fact, the Greenstone system has the ability to display multiple interfaces without difficulty. However, producing a practical design that enables users to browse and search collections on different interfaces is not an easy task, as both interfaces should have the same features to display collections effectively and efficiently.
25. Legal Issues

“Intellectual property and copyright is a considerable issue for nearly all digital libraries therefore it is important to understand the legal issues and to follow legal requirements”. (Dawson, 2003). In the short term, legal issues were not considered to be a serious restriction as most of the documents that were to be digitized belong to the National Oil Corporation and most of these documents are internal correspondence, statistical data, reports, and publications by research workers and employees of the National Oil Corporation, in addition to a CDS/ISIS database which is a bibliographic database in Arabic and English for the holdings of the central library of the NOC. The proposed digital library will also include links to free portals in the oil industries that are not under copyright. In the long term, the NOC has a documented policy regarding copyright issues. This policy regulates the use of copyright materials as well as intellectual property.

26. Prototype

After designing the proposed digital library, the library was tested in order to discover if there were any problems with the design. The prototype was taken back to Libya in order to get feedback on the design and to evaluate the library. The prototype was useful in many ways: it was built to test the function of the proposed library and to solve unexpected problems. The point of the prototype was to assist in building the full system.

27. Evaluation

This is the last stage in designing a digital library. This phase is an important stage as the evaluation will assist in developing the proposed digital library and lead to an improved implementation based on the information collated in the evaluation.

Evaluating the prototype digital library is considered an important stage towards developing a full-scale NOC digital library. Different approaches have been used by digital library developers depending on their evaluation goals (Solis, 2005). Buchanan (2009) states that usability and usefulness can be readily combined, and that questionnaire and observation are valid multi-method approaches. For the NOC digital library, a usability–centred approach was applied to evaluate user perceptions as to the ease of the use of the prototype library and to assess user satisfaction. To achieve the evaluation of the prototype library, the work has been divided into different phases, each stage is subdivided into steps, and each step further contains a series of tasks. Thus, the total work is broken down into manageable portion.

28. Objectives

The purpose of this stage is to assist in answering the following questions:

• What problems do library patrons face in using the prototype library?
• To what extent was the prototype library efficient in assisting users to get the most from it?
• To what extent was the user interface successful in assisting library users to find and discover relevant resources in the most convenient way?
• To what extent were browsing and searching facilities suitable for end users? Is there any need to modify these facilities to meet user needs and requirements?

23. Generic Stages

This stage clarifies the main activities, purposes, and assessment activities. It is divided into the following steps:

• Drawing up an evaluation plan
• Data gathering and recording
• Data analysis and interpretation of results
• Presentation of findings

24. Drawing up an evaluation plan

Before designing the questionnaire, a plan has to be established which clarifies what sort of
data should be collected and for what purpose. Saracevic (2004) views digital libraries as complex social, institutional, and technical systems. No evaluation can possibly address all of these aspects together. Thus, different approaches have been used by other researchers to achieve different evaluation goals (the desired knowledge about a system, its performance, or usability) e.g. a systems-centred approach, a human-centred approach, a sociological approach, an economic approach, and a usability-centred approach. Chowdhury (2006) concludes, “Digital libraries differ significantly from one another in terms of their nature, content, target users, access mechanisms, etc., and consequently it is difficult to measure the usability of such diverse digital libraries through one set of universally accepted tools and benchmarks.” Fuhr (2007) states that “digital libraries are complex systems; they can be, and are, viewed from different perspectives. The methods and metrics for the evaluation [of Digital Libraries] may vary according to whether they are viewed as institutions, as information systems, as new technologies, as collections, or as new services.” The International Standards Organization (ISO) defines usability “as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.” As usability of the constructed digital library is considered to be one of the most important issues and can reflect to great extent the success of the prototype library, the researcher assessed the following elements:

25. Interface Features
- Search facilities (e.g. simple vs. expert search)
- Attractiveness and consistency.
- Language of the interface.
- Visual appearance: use of colours, typography, layout and graphics, font size, and font type.
- Personalization of the interface, e.g. the ability to define the number of records on one page, sort options, etc.

26. What needs to be discovered through this evaluation?
- Ease of use. To evaluate users’ perceptions as to the ease of the use of the library.
- Quality of user’s experience.
- How satisfied are the users? Satisfaction and success.
- What changes are required?

27. Methodology
Evaluation is the systematic process of determining the merit, value, and worth of something. Evaluation method is a series of steps or actions taken to accomplish an evaluation. Saracevic (2004) states that “the range of methods used in digital library evaluations is wide. It would be hard to find a scholarly evaluation method that was not used”. A number of evaluation studies have involved several methods. Ingrid Hsieh - Yee (2005) identifies the following methods:
- Surveys, including direct questionnaires and online surveys
- Structured interviews
- Focus groups
- Observations
- Task accomplishment
- Think aloud
- Case studies
- Experimentation
- Records analysis
- Usage analysis
- Documents, meeting, communication analysis (anthropology)
- Economic analysis
- Cross-cultural assessment
- Transaction logs analysis

Saracevic (2004) states that digital libraries are complex entities, therefore many methods are appropriate. “Each method has, by definition,
certain strengths, and weaknesses. There is no one best method and there never will be one.” The evaluation method applied for this task consisted of a questionnaire survey for gathering the data for the evaluation of the digital library. Questionnaires were used as the primary method of getting user feedback during the experiments. Check-box questions were used in this questionnaire and a Likert scale was used to measure usability criteria.

28. Likert Scales

Likert scales present users with a statement and they are asked to register their level of agreement or disagreement with the statement. The intention of using the Likert scale lies in the fact that this type of scale is capable of measuring users’ tendency. A five-point scale was used instead of a seven-point or a three-point scale in the evaluation of this library. A five-point scale can measure what needs to be discovered and can fulfill the purpose of the evaluation. In fact, the seven-point scale is usually used in psychological studies where phenomena are much complicated and require in depth analysis to understand human behaviour. The score of the scale was listed at the base of the scale.

29. Administering the Scale

In order to capture the perceptions of respondents with respect to the features that they evaluated in the Likert scale, each question required the respondent to rate each feature on a response scale. For instance, they could rate each item on a 1 to 5--response scale where in the first question:

1. = very poor
2. = poor
3. = barely acceptable
4. = good
5. = very good

And in the second question:
1. = strongly disagree
2. = disagree
3. = undecided
4. = agree
5. = strongly agree

The questionnaire was made up of three general sections:
- The characteristics that users wanted the new library to have.
- Search tools and preferences.
- Quality of user experience and required changes.

30. Participants

A number of questionnaires were distributed to a sample of employees and research workers within the National Oil Corporation in different departments and administrative units. In addition, a number of questionnaires were distributed to managers and systems people in different locations of the same organization. The mechanism used for selecting a sample of special library users is described below.

According to the labour force statistics for the oil and gas sector issued by the Manpower Planning Department in Libya in September 2007, the number of employees and workers at the National Oil Corporation totalled 908 people. In order to have a manageable sample of special library users, a small sample from the total number of employees was selected (5%) and the questionnaire was then distributed to 40 employees and research workers in the organization. In addition to that, six questionnaires were distributed to managers and systems people in the corporation. According to Nielsen (2003), “in this type of surveys on usability, from the fifth user on, most of the usability problems that may be found in an information system are already identified. From then on, the results present little variability.” A small sample of users was used as it was recognized that the evaluation process would require volunteers to use the prototype digital library for a period and that a large sample would involve diverting a considerable amount of resource from core activities. Questionnaires
were distributed by hand to National Oil Corporation employees. The researcher had visited the central library of the NOC for three consecutive days. During these three days, volunteers were asked to participate in the evaluation study. In the test, volunteers were requested to freely search the prototype digital library for a period of between 15 and 20 minutes and then to answer the three sets of questions. It was assumed that through this exercise, the participants got a general understanding of this library. From the 46 questionnaires that were distributed to NOC library users, 38 questionnaires were completed representing a response rate of (83%). As this figure is acceptable in research and academic studies, the figure was judged reliable and acceptable and gave a good indication regarding the effectiveness of the design of the prototype digital library. 96% of distributed questionnaires to managers and systems people were returned complete.

32. Analysis

<table>
<thead>
<tr>
<th>Feature</th>
<th>Very poor</th>
<th>Poor</th>
<th>Barely acceptable</th>
<th>Good</th>
<th>Very good</th>
<th>Rating average</th>
<th>Response count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic interface</td>
<td>1 (2%)</td>
<td>4 (11%)</td>
<td>3 (8%)</td>
<td>19 (50%)</td>
<td>11 (29%)</td>
<td>3.92</td>
<td>38</td>
</tr>
<tr>
<td>English interface</td>
<td>2 (4%)</td>
<td>2 (4%)</td>
<td>7 (18%)</td>
<td>17 (48%)</td>
<td>10 (26%)</td>
<td>3.82</td>
<td>38</td>
</tr>
<tr>
<td>Navigation options</td>
<td>8 (21%)</td>
<td>4 (11%)</td>
<td>12 (32%)</td>
<td>8 (21%)</td>
<td>6 (15%)</td>
<td>3.00</td>
<td>38</td>
</tr>
<tr>
<td>Cross-collection search</td>
<td>7 (18%)</td>
<td>9 (24%)</td>
<td>8 (21%)</td>
<td>10 (26%)</td>
<td>4 (11%)</td>
<td>2.87</td>
<td>38</td>
</tr>
<tr>
<td>Simple search</td>
<td>2 (4%)</td>
<td>4 (11%)</td>
<td>6 (16%)</td>
<td>12 (32%)</td>
<td>14 (37%)</td>
<td>3.84</td>
<td>38</td>
</tr>
<tr>
<td>Expert search</td>
<td>2 (5%)</td>
<td>2 (5%)</td>
<td>3 (8%)</td>
<td>23 (61%)</td>
<td>8 (21%)</td>
<td>3.87</td>
<td>38</td>
</tr>
<tr>
<td>Full text searching</td>
<td>9 (24%)</td>
<td>11 (29%)</td>
<td>3 (8%)</td>
<td>7 (18%)</td>
<td>8 (21%)</td>
<td>2.84</td>
<td>38</td>
</tr>
<tr>
<td>Clear and readable text</td>
<td>7 (18%)</td>
<td>13 (35%)</td>
<td>3 (8%)</td>
<td>5 (13%)</td>
<td>10 (26%)</td>
<td>2.95</td>
<td>38</td>
</tr>
<tr>
<td>Font size and type</td>
<td>8 (21%)</td>
<td>11 (29%)</td>
<td>6 (16%)</td>
<td>9 (24%)</td>
<td>4 (11%)</td>
<td>2.74</td>
<td>38</td>
</tr>
<tr>
<td>Visual appearance</td>
<td>2 (5%)</td>
<td>7 (18%)</td>
<td>3 (8%)</td>
<td>14 (37%)</td>
<td>12 (32%)</td>
<td>3.71</td>
<td>38</td>
</tr>
<tr>
<td>Preferences</td>
<td>2 (5%)</td>
<td>4 (11%)</td>
<td>8 (21%)</td>
<td>12 (32%)</td>
<td>12 (32%)</td>
<td>3.66</td>
<td>38</td>
</tr>
<tr>
<td>Personalization of inter-</td>
<td>8 (21%)</td>
<td>4 (11%)</td>
<td>2 (5%)</td>
<td>14 (37%)</td>
<td>10 (26%)</td>
<td>3.37</td>
<td>38</td>
</tr>
</tbody>
</table>
Here is how the rating average for the first row in the first question Arabic interface was calculated.

\[ 1 (1) + 4 (2) + 3 (3) + 19 (4) + 11(5) = 149 \]

Sum of selected choices or calculated frequencies:

\[ 1 + 4 + 3 + 19 + 11 = 38 \]

Sum of frequency times column weight divided by the sum of selected choices:

\[ \frac{149}{38} = 3.92 \]

The rating average is 3.92

This means that the respondents selected between columns 3 and 4 or between barely acceptable and good, but very closer to good or 4.

Overall grade or indicator for the first question = 3.38

Here is how the rating average for the first row in the second question was calculated.

\[ 3 \times 1 + 3 \times 2 + 6 \times 3 + 12 \times 4 + 14 \times 5 = 145 \]

Sum of selected choices or calculated frequencies:

\[ 3 + 3 + 6 + 12 + 14 = 38 \]

Sum of frequency times column weight divided by the sum of selected choices:

\[ \frac{145}{38} = 3.82 \]

The rating average is 3.82

This means that the respondents selected between columns 3 and 4 or between barely acceptable and good, but very closer to good or 4.

Overall grade or indicator for the second question = 3.41

### Table 2 Calculations of responses for the second question

<table>
<thead>
<tr>
<th>Statements</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Tend to agree</th>
<th>Agree</th>
<th>Strongly agree</th>
<th>Rating average</th>
<th>Response count</th>
</tr>
</thead>
<tbody>
<tr>
<td>My experience with the library today was quite pleasurable</td>
<td>3 (8%)</td>
<td>3 (8%)</td>
<td>6 (16%)</td>
<td>12 (31%)</td>
<td>14 (37%)</td>
<td>3.82</td>
<td>38</td>
</tr>
<tr>
<td>I was able to navigate easily in the digital library</td>
<td>6 (16%)</td>
<td>8 (21%)</td>
<td>10 (26%)</td>
<td>10 (26%)</td>
<td>4 (11%)</td>
<td>2.95</td>
<td>38</td>
</tr>
<tr>
<td>The results I obtained in my searches were relevant</td>
<td>9 (24%)</td>
<td>5 (13%)</td>
<td>7 (18%)</td>
<td>14 (37%)</td>
<td>3 (8%)</td>
<td>2.92</td>
<td>38</td>
</tr>
<tr>
<td>The design of the library is clear, simple, and consistent</td>
<td>2 (5%)</td>
<td>2 (5%)</td>
<td>8 (21%)</td>
<td>12 (32%)</td>
<td>14 (37%)</td>
<td>3.89</td>
<td>38</td>
</tr>
<tr>
<td>Search and browse facilities are good enough for retrieval of information</td>
<td>2 (5%)</td>
<td>4 (11%)</td>
<td>10 (26%)</td>
<td>12 (32%)</td>
<td>10 (26%)</td>
<td>3.63</td>
<td>38</td>
</tr>
<tr>
<td>The library works fine in both languages</td>
<td>7 (18%)</td>
<td>9 (24%)</td>
<td>4 (11%)</td>
<td>8 (21%)</td>
<td>10 (26%)</td>
<td>3.21</td>
<td>38</td>
</tr>
<tr>
<td>Personalization of the interface made my experience more enjoyable</td>
<td>7 (18%)</td>
<td>3 (8%)</td>
<td>11 (29%)</td>
<td>9 (24%)</td>
<td>8 (21%)</td>
<td>3.21</td>
<td>38</td>
</tr>
<tr>
<td>The digital library offers related links to other electronic resources</td>
<td>6 (16%)</td>
<td>4 (11%)</td>
<td>2 (5%)</td>
<td>12 (31%)</td>
<td>14 (37%)</td>
<td>3.63</td>
<td>38</td>
</tr>
</tbody>
</table>
33. Discussion
Above and below indicator: It would be rather difficult to extract any conclusion from these numbers, but it gives a general idea on how to compare those questions that were graded either below or above this number. Below is a list of indicators that have larger deviations from the indicator calculated for the first question.

33.1 Above the indicator 3.38
• Arabic interface (3.92)
• English interface (3.82)
• Simple search (3.84)
• Expert search (3.87)
• Visual appearance (3.71)

33.2 Below the indicator 3.38
• Cross-collection search (2.87)
• Full text searching (2.84)
• Clear and readable text (2.95)
• Font size and font type (2.74)

Below is a list of indicators that have larger deviations from the indicator calculated for the second question.

33.3 Above the indicator 3.41
• My experience with the library today was quite pleasurable (3.82)
• The design of the library is clear, simple, and consistent (3.89)
• Search and browse facilities are good enough for retrieval of information (3.63)
• The digital library offers relevant links to other electronic resources (3.63)

33.4 Below the indicator 3.41
• I was able to navigate easily in the digital library (2.95)
• The results I obtained in my searches were relevant (2.92)
• The library works fine in both languages (3.21)
• Personalization of the interface made my experience more enjoyable (3.21)

The first block shows a high performance in the design of the Arabic and English interface in addition to high performance in the design of the visual appearance and search and browse facilities. The features with a poor grade reflect some design problems. The grades given to font size, font type and clear and readable text may reflect some problems that need to be fixed and improved.

In the second question, again, according to the above indicators the first block shows a high performance in terms of general accessibility and user satisfaction but the grade given to relevant search results may also reflect a problem with advanced search facilities especially in the Arabic interface where stemming does not work as it should be as the recall of relevant documents in the set of all documents returned by a search was too low.

34. Results
The overall grade returned for the first and second question is quite reasonable, as the highest grade was 5. This indicates that a considerable number of the library’s characteristics seem to be efficient and effective.

The design of the Arabic and English interfaces was quite good as (50%) and (48%) respectively of returned responses demonstrated that the design was quite good.

Simple and advanced search was also quite reasonable according to user feedback (3.84) and (3.87).

Font size and type of font reflected some design problems and need to be fixed by changing both the type of the font as well as the size.

35. An overview of the National Oil Corporation digital library
This is a snapshot of the NOC digital library that has been designed with the aim of developing the services currently provided for end users. The following screenshot shows the collections of the library. As can be seen, the digital library consists
of fifteen different collections. Preferences and help options are located at the upper right corner. Users may set their preferences before starting to browse the library and can consult help options to get more information about how to use the digital library. The digital library home page shows different icons. In order to navigate the library, users have to select a collection by clicking the preferred icon. Once the icon is clicked, the user will be directed to the appropriate collection and a new page will open.

The next screenshot shows the thesis collection. This collection is authenticated and requires users to sign in, which means that users have to enter a username and a password in order to view the collection. The NOC digital library contains some other collections that are restricted to authorized users—especially collections that contain sensitive documents. The collection’s name is placed in the upper left corner of the banner. In the upper right corner home, help, and preferences are placed. The home button if clicked will return the user to the home page to select a collection.

The following screenshot shows the CDS/ISIS collection. This is a bibliographic database of the holdings of the central library of the National Oil Corporation. This collection consists of bibliographic details of about 3,400 records taken from the NOC CDS/ISIS database. One can

![Figure 3 Library interface](image)

**Figure 3** Library interface

![Figure 4 Sign in page](image)

**Figure 4** Sign in page
browse the publications by title, author, editor, keyword, type of source, and location. Advanced searches can be performed using title, creator, keyword, year of publication, Dewey class, and terms occurring in the text. The search terms can be combined using Boolean operators. By default, search is case sensitive, although the preferences can be changed. Browse facilities can be clearly seen under the banner. The user can browse the publications by title, author, editor, keyword, type of information source, and by location of the source in the central library. In fact, the CDS/ISIS database, which is a standalone database, was converted to web interface using the Greenstone facility so that users can search and browse the central library catalogue from anywhere, provided they have Internet access and basic ITC infrastructure.

The following screenshot displays the advanced search facility. Users can enter a query directly into the space provided and click on “run query” or use “advanced search” which can be performed using title, creator, and keyword, year of publication, Dewey class, and terms occurring in the text. The search terms can be combined using Boolean operators. Users have to set their preferences to “advanced search” in order to be able to perform this task as the ordinary search is set by default to plain search. The preferences button also enables users to change the search preferences concerning query mode, query style, case differences, word ending, accent differences, and search history.

The following screenshot displays titles from the CDS/ISIS collection in browse mode. It can be seen that titles are arranged

---

**Figure 5** CDS/ISIS database

**Figure 6** Advanced search facility in NOC digital library
alphabetically from A – Z and then from 1 – 9. Users can browse the titles under particular letters. Users could also browse the database using author, editor, and keyword in the same way.

36. Conclusion

This paper has shown that today’s technology can be put into practice for developing special library services. The outcomes of the design have proven encouraging, as the design has been shown to be cost effective. The design has also demonstrated that the digital library developed in this research can assist, to a large degree, in developing the services provided to special library users in Libya. This paper has also shown that digital libraries ought not to be seen as an end in themselves, but rather as a means for enabling end users to access a variety and wide range of frequently required services. The adoption and exploitation of today’s technology in a time of shrinking budgets has become more important than ever. This paper has proposed a new strategy for developing special library services by adopting and utilizing Open Source Software (OSS) since building repositories and digital libraries by using freely available software can assist to a large degree in enhancing the level of the services provided to users.

References


Figure 7 Browse facility


Witten, I. and Bainbridge, D. 2007. How to build a digital library. San Francisco: Morgan Kaufmann
Rapid progress has been made in the discipline of biochemical engineering and biotechnology for bioprocess development during the last 50 years. Process Biotechnology: theory and practice has been written with the consideration that tutorial practice is as important as understanding the subject theoretically. This book is an introductory tutorial book involving multidisciplinary principles. Principal innovations that have been made in biosystem-related developments have been emphasized through tutorials in this book. The first few chapters cover theoretical aspects of biochemical and chemical engineering concerns in biotechnological advances in a concise manner. The rest have been dedicated to the tutorial aspects of this multidisciplinary subject.

This book covers biological, ecological, chemical, and biochemical engineering topics related to the subject. It provides much needed theory-based solved numerical problems for practice in quantitative evaluation of various parameters relevant to process biotechnology. It will be useful for students who would like to further their careers as biotechnologists and can be used as a self-study text for practicing engineers, biotechnologists, microbiologists, and scientists involved in bioprocessing research and other related fields.

**Key features**

- Introduction to the NADE and RADE theory of evolution of life forms.
- Coverage of theoretical aspects of biological, ecological, biochemical, and chemical engineering systems in process biotechnology.
- Inclusion of research-generated data based on several theory-based solved numerical problems.
- Provision of experimental data-based formulated numerical problems for tutorial practice.
- Inclusion of exceptional colour plates to explain complex concepts and material.

**Table of contents**

**SECTION I: BACKGROUND ON LIFE FORMATION**
- Prebiotic science of chemicals for life
- Biotic chemistry
- Cell biology and microbiology
- Chemical and biochemical engineering principles

**SECTION II: BIOPROCESS ENGINEERING SYSTEM: TUTORIALS**
- Upstream process biotechnology: solved numerical problems
- Enzyme engineering and technology: solved numerical problems
- Process recovery and purification: solved numerical problems
- Unsolved numerical problems
Conservation and Preservation of Archives: A Case Study of Punjab Digital Library

Ms Balbir Kaur Pandher
Librarian, Guru Nanak Khalsa College for Women
Model Town (Gujarkhan Campus)
Ludhiana, Punjab
E-mail: balbirpandher@gmail.com

Abstract

Archives contain millions of documents like manuscripts, rare books, paintings, photographs, and historical records. All these constitute human heritage. Paper manuscripts, with the passage of time, are getting fragile and brittle because of various reasons. Lamination does not seem to be a permanent solution of preserving this cultural heritage for posterity. Digitization technology brings with it untold benefits for heritage preservation and access. Once a document has been properly digitized, it becomes immortal and can remain accessible long after the original has ceased to exist. The option of digital access further aids in preservation of originals through reduced need for physical handling.

The Punjab Digital Library is a new approach to preserving Punjab’s heritage for future generations. The central digital archive, developed over the last six years, grants electronic access using powerful searching and browsing capabilities for the researcher to sift through vast amounts of data. PDL has been instrumental in digitally preserving over 2.5 million folios from 3,400 manuscripts, 2,200 book, 1,990 issues of periodicals, 5,578 issues of newspapers, 3,152 photographs, 248,000 legal documents, and some 168 hours of video recordings. The current collection of data amounts to about 15,000 GB.

Keywords: Archives, Punjab Digital Library, Digitization, Manuscripts, Heritage, Artefacts
1. Introduction

“Understanding the past is the key to our future.” – Professor Stuart Macintyre.

Human past is splendid and full of wonders. The past is known by records, archives, books, and other library materials that constitute documentary sources. There are also other things like monuments, buildings, art objects, and other artefacts. All these constitute human heritage. They need conservation and preservation. Librarians, archivists, curators, chemists, and archaeologist the world over are concerned with the problems of conservation and preservation.

The documentary sources which are in the form of manuscripts, books, records, archives, films, CDs, etc, are fragile and prone to decay and deterioration. These are damaged because of weather, fire, insects, rodents, floods, virus, and many other factors. Sometimes the documentary sources are also destroyed because of wars, holocaust, disasters, and religious animosity and intolerance. The examples are of the burning of the great libraries at Alexandria and Nalanda in Egypt and India respectively, and the recent destruction of Bamiyan Buddha in Afghanistan.

Preservation is a part of conservation. Preservation is concerned with problems like the repair, dusting, fumigation, de-acidification, air-conditioning, lamination, binding, and storage of manuscripts, books, films, disks, and optical materials.

2. Preservation and digital preservation

The American Institute for Conservation of Historic and Artistic Works (AIC) Definitions of conservation defines these three terms as follows:

- **Preservation** — the protection of cultural property through activities that minimize chemical and physical deterioration and damage and that prevent loss of informational content. The primary goal of preservation is to prolong the existence of cultural property.

- **Conservation** — the profession devoted to the preservation of cultural property for the future. Conservation activities include examination, documentation, treatment, and preventive care, supported by research and education.

- **Restoration** — Treatment procedures intended to return cultural property to a known or assumed state, often through the addition of non-original material. Informational content. The primary goal of preservation is to prolong the existence of cultural property.

Preservation, by definition, is “activities associated with maintaining library and archival materials for use either in their original physical form or in some other usable way.”

Preservation involves keeping a balance between collection-level activities such as environmental control, which can be difficult and/or costly to manage but provide the greatest long-term benefit for most materials, and item-level activities such as conservation treatment, which are often more easily understood and managed but can have limited effect, especially if the items are returned to a damaging environment. Nowadays due to information and communication technology, preservation techniques are changed and digital preservation has come up.

Digital preservation is the management of digital information over time. Preservation of digital information is widely considered to require more constant and ongoing attention than preservation of other media.

Digital preservation combines policies, strategies, and actions to ensure the accurate rendering of authenticated content over time, regardless of the challenges of media failure and technological change. Digital preservation applies to both born digital and reformatted content.

Digital preservation policies document an organization’s commitment to preserve digital content for future use; specify file formats to
be preserved and the level of preservation to be provided; and ensure compliance with standards and best practices for responsible stewardship of digital information. Digital preservation strategies and actions address content creation, integrity, and maintenance.

2.1 Content creation includes
- Clear and complete technical specifications
- Production of reliable master files
- Sufficient descriptive, administrative, and structural metadata to ensure future access
- Detailed quality control of processes

2.2 Content integrity includes
- Documentation of all policies, strategies, and procedures
- Use of persistent identifiers
- Recorded provenance and change history for all objects
- Verification mechanisms
- Attention to security requirements
- Routine audits

2.3 Content maintenance includes
- A robust computing and networking infrastructure
- Storage and synchronization of files at multiple sites
- Continuous monitoring and management of files
- Programs for refreshing, migration, and emulation
- Creation and testing of disaster prevention and recovery plans
- Periodic review and updating of policies and procedures

3. Manuscripts at a glance
The term ‘manuscript’ is derived from the Latin manu scriptus, which mean ‘hand written’. The term may also be used for text written in other ways such as chiselling or scratching with a pointed device over a recording surface. In the modern context, manuscript also refers to a hand-written text submitted to a printer for publication.

In earlier times, all recordable text knowledge was in the literal manuscript form, due to the non-existence of modern print technology. Even when the printing became available later, manuscript was the still favoured form in South Asia, as printing often remained a prohibitively expensive affair. This kept the majority of the scribes at bay, and manuscripts remained in vogue until the late nineteenth century, when typewriters came onto the public scene. Information has travelled through human generations in various forms over the centuries. Initially all the information was passed on orally, which later took the shape of the scribal tradition of hand written recordings. Print technology took over from the manuscripts, allowing for mass-scale distribution, eventually evolving in to the electronic medium that has revolutionized and come to define the modern age of communication and transmission.

Manuscript writing was a laborious and time-consuming task that would take months or sometimes years to complete. Some manuscripts were considered especially valuable because of the addition of illustrations. The pages were usually embellished with colourful works, which could contain enriched depictions, multi-inked decorated borders, boldly written, brightly coloured lettering, or the insertion of full-page images. Since the creation of these pieces of art was a task requiring indulgence great deal of skill and labour, illuminated pages also show a considerable emotional relationship between such work and the scribe.
Due to high cost of paper and ink, most of the manuscripts, in South Asia, were written without spaces between the words where the script allowed such arrangement. This makes the reading of such material quite difficult for a novice. Manuscripts are defined by the inscription technique as explained above, and not by their content, which may variously contain text, graphics, maps, figures, and illustrations. Manuscripts may be in the form of loose pages tied together, bound or in codex format. The study of the writing in manuscripts is termed palaeography and the study of manuscripts themselves manuscriptology.

3.1. Significance
Manuscripts, like any other treasure of literature, are the mirrors of that age, more so as they include pieces of history beyond a plain pattern of printed text. Manuscripts are livelier than any other piece of written literature and art. They give us a wide glimpse into their age of inscription, encompassing the culture, history, art, script, calligraphy, language, social, political, and economic states of the time. They are also indirect records of the state of technology, paper quality, ink, embellishment, binding material, and techniques popular and available and the time, and give us an inside view of the stages of development and advancements in the thought and literature in their particular epoch. Such a rich repository of the past definitively deserves a well-defined policy and practice of handling and preservation in order to prolong their viability as long as possible.

4. Need for digitization
Like any materials, manuscripts are prone to ageing and deterioration. Most of the manuscripts are paper based, which has a naturally limited life despite preservation practices. Manuscripts are affected by myriad factors such as time, weather, temperature, humidity, fire, water, insects, bugs, and human negligence or destruction. Although the exact spirit and techniques of the originals cannot be genuinely replicated in physical form, the information contained in the form of text and graphics can be preserved in the digital form. The practice of digitization transforms the physical analog data into a virtual digital equivalent of the same that can be displayed on a computer screen or reproduced on any other physical material through printing. This allows for the information contained in a manuscript to survive long after the original is gone, while also increasing the ways in which anyone can access and utilize that information. This is where digitization has the most important role to play.

4.1 Area of the study
Considering the importance of digitization in the present context, the study is intended to know the digitization process of various materials in Punjab Digital Library, Chandigarh.

4.2 Objectives of the study
The objectives of the study are:
1. To facilitate the long-term management and preservation of manuscripts
2. To facilitate new forms of access and use
3. To enhance access for improving services to an expanding user group
4. To reduce the handling and use of fragile or heavily used original material and create a backup copy for endangered material

5. Punjab Digital Library
Punjab is a state in northwest India, forming part of the larger Punjab region. The state is bordered by the Indian states of Himachal Pradesh and Jammu and Kashmir to the northeast, Haryana to the south and southeast, Rajasthan to the southwest, and the Union Territory of Chandigarh to the east, as well as the Pakistani province of Punjab to the west. The state capital is Chandigarh, which is administered separately as a Union Territory because it is also the capital of Haryana. The major cities of Punjab include Mohali, Ludhiana, Amritsar, Bhatinda, Patiala,
and Jalandhar. After the partition of India in 1947, the Punjab province of British India was divided between India and Pakistan. The Indian Punjab was divided in 1966 with the formation of the new states of Haryana and Himachal Pradesh as well as the current state of Punjab. PDL began operations in 2003 as a small organization tasked with a gigantic mission. Starting with one agenda, one desktop, and one employee, PDL has come a long way since. The Punjab Digital Library is a NGO that is digitizing and preserving the cultural heritage of Punjab. There are many historically significant documents stored and made available online. Its scope covers Sikh and Punjabi culture. The library funded by The Nanakshahi Trust was finally launched online in August 2009. It is located at Chandigarh. The project emerged from the concerns shared by a group of individuals for the fast-disappearing or already-lost heritage of Sikhs and Punjab. In many such cases, valuable material, rare literature, architecture, and signs of much celebrated memories were completely destroyed, whether by ageing, accident, or human aggression. Despite these losses, it was clear there was still a lot left that worth saving. This vision of preservation was the primary motivation for the project. After much deliberation and brainstorming, the project was established with a focus on archiving the endangered invaluable manuscripts and other literature to conserve and defend the heritage, culture, and language of the Punjab state.

6. Miniatures from Bhagvata Puran

Heritage includes manuscripts, books, official records, newspapers, miniatures, murals, periodicals, paintings, pictures, architecture, which are being digitized at PDL. The rich heritage of Punjab was rapidly dwindling and degenerating due to the destructive effects of time, age, and human negligence. Many rare manuscripts or invaluable pieces of history are destroyed during wars, invasions, and natural disasters. The idea of digitizing text is the best possible solution to this loss of information. PDL went through rigorous research and groundwork initially to establish commonalities and fix mutual priorities. With each passing day, the organization grew in skill, talent, experience, and number. The budding library acquired equipment, and skills in preservation, digitization, data management, and archiving. Projects around the globe were closely studied to build an understanding of what would be needed to establish a successful digitization project. Internationally recognized benchmarks were referred to and complied with. In the process, the organization also created benchmarks and documentation relevant to the realities of South Asia that conformed to the international standards.
PDL seeks to promote an entirely new culture of awareness, where the masses contribute to the safeguarding of old texts.

Digitized materials are as following:
1. Manuscripts
2. Books
3. Magazines
4. Photographs

Launched in 2003 under Nanakshahi Trust, the Punjab Digital Library was a result of the early phase of the digital revolution in Punjab. While most saw Nanakshahi as a small digitization organization, or as an assemblage of some unknown youth working towards capturing some manuscripts on their digital cameras, its founders saw it as a cornerstone of a fundamentally new approach to preserving Punjab’s heritage for future generations. In the shadow of search engines, a Semantic Web approach thought of in the early 2003 reached maturity in 2006. This was when the organization planned to expand its operations from a mere three-employee organization to one of the leading NGOs working in the field of digital preservation all over India. Digitized collections include manuscripts held by the:
1. Punjab Languages Department
2. Government Museum, Chandigarh
3. Art Gallery, Chandigarh
4. Chief Khalsa Diwan, SGPC, DSGMC and manuscripts in the Jawaharlal Nehru Library of Kurukshetra University
5. It also includes hundreds of personal collections

With over five million pages digitized, it is the biggest repository of digital data in Punjab. The PDL collection includes manuscripts dated from the early fifteenth century, with varying subjects that include theology, history, philosophy, medicine, and dictionaries. The genre includes texts in Gurumukhi, Sharda, Devanagri, Punjabi, Urdu, Sanskrit, Arabic, and Farsi scripts and languages.

7. Techniques of digitization

Different types of documents require different techniques. International standards are followed. Digital preservation is the most viable, and the only major technological alternative available to us for safeguarding our fast diminishing heritage. In order to contribute to any preservation effort and bring about a positive change, technological advances have to be met with equal amount of concern, enthusiasm, and will to take concrete steps. PDL, from its very conception, has been at the helm of preservation efforts aimed at salvaging the position of Sikhs and Punjabis as a cultural identity. Providing digitization services to individuals and institutions since 2003, it has been instrumental in digitally archiving thousand of originals of manuscripts, rare books, and other literature.

PDL provides digitization services both on site and off site depending upon the constraints.

7.1 Services provided by PDL
1. PDL Discussion Forum
2. Digitization
3. Data Mining
4. Inter-library loan
5. Exhibitions
6. Library shop
7. Preservation
8. Turning point
8. Funding
PDL Gets support from
1 Nanakshahi Trust
2 Sikh Reference Institute
3 Individual Donors

9. Working groups
PDL assembled 15 Working Groups in 2008, drawing specialists from their respective fields. Of these original 15, eight were made permanent Working Groups. Group members consist of programmers, professionals, and researchers who work and collaborate on an interdisciplinary basis to complete a given project. The functions of a working group also include review of existing applications, evaluating goals, effectiveness, usability, and design, in order to provide feedback for further improving presentation and functionality. It is on the professional expertise of these working groups that the output and success of PDL depends. The life span of each working group lasts anywhere between a few months to several years depending upon specificity and size of the project.

1. PDL Content Selection Working Group
2. PDL Copyright & IPR Issues Working Group
3. PDL Digital Data Management Working Group
4. PDL Finance Management Working Group
5. PDL Marketing Working Group
6. PDL Library Planning Working Group
7. PDL Digitization Standards (Review) Working Group
8. PDL Metadata Working Group
9. PDL Technology, Standards and Architecture Working Group
10. PDL Digital Library Metrics Working Group
11. PDL Exhibits Working Group
12. PDL Products Working Group
13. PDL Design Working Group
14. PDL Data Migration Working Group
15. PDL Back-up Working Group.

10. Staffing
Human resource is the most essential part of any institutional effort. To have a trained staff that can handle a variety of tasks and take informed decision under different situation is vital to any organization. PDL has acquired and developed much of its required skilled work force to sustain the digital project. The PDL staff includes

1. A dedicated group of programmers that has developed its online digital library and maintains it as well.
2. Computer professionals, highly skilled in digitization and post digitization, and processing staff to handle a capacity of up to 10,000 folios per day.
3. There is also a group of professors, language and library experts working part to full time for metadata creation and research purposes.
4. PDL also hires short-term local employees for assistance in on-site digitization projects in other cities. These people are adequately trained to help in various digitization processes, usually generating a class of digitization literate people who can be readily employed to adapt to other digital projects without much investments.

11. Infrastructure and equipments in use
PDL continues to expand its infrastructure to match the service quality it wishes to provide to its user base. Most of the infrastructure is based on the scalable and interoperable models. With about 30 installed workstations, PDL has the highest number of data processing power in the region. A good amount of digitization power is in used, which include:
• Multiple cameras
• Lighting equipments for varying and specific digitization needs
Flatbed scanners with capacity ranging needs from A4 size papers to 42 wide formats. PDL has developed on its own much of the in-house digitization equipment in use.

**Scanners/Cameras**
- Wide format scanner that can scan document width up to 42 inches
- 10 Digital SLR cameras
- 2 scanners 8.5”x14”
- 4 scanners 8.5”x11.5”
- 2 book scanners with V shaped cradle
- 6 dark rooms with lighting equipment
- Servers: IBM server with an installed storage of 20TB, Scalable up to 48TB
- Backup equipment: LTOP Tap drives, hard drives, DVD writers
- Computers: 25 desktops, 10 Laptops

PDL has independently developed much of the support equipment required for digitization. Lately, it has successfully created specially customized digitization worktables, a complete digitization apparatus with integrated lighting system and computer systems for wide screen live view, and real time simultaneous data transfer with the installed equipment. PDL intends to increase both the output and quality of the digitization work undertaken, and to digitally preserve more manuscripts in short span of time.

**The Manthan Award South Asia 2010** (Digital inclusion for Development) has been awarded to PDL for its services.

**12. Advantages of digitization**
- Remote access: Connecting people globally by providing continued free online access
- Multiple accesses: One of the most important advantage digitization is multiple access documents. If a manuscript is in physical format, then it can be handled by only one person in the specific period, but through the digitization process, several users can access the specified document at a time.
- Preservation and conservation: Addressing heritage loss due to environment, ignorance, and destruction
- Dissemination and promotion: Saving invaluable treasures to enrich the present and enlighten the future. Dissemination of knowledge and culture via the internet is a 21st century phenomenon.

**10.1. Barriers in Digitizing**
- Initially no one understood what digitization was, and so everyone was hesitating in submitting his or her manuscript for digitization.
- There is no doubt that in present times 27 million documents from the government, universities, and personal collection are to be digitized.
- Initially they used to send requests to the public to send their manuscripts for digitization but now the scenario has changed; the request is being sent by the public.
- It will take approximately ten years to fulfil the requests that are pending.
- There is a great need to create awareness among the masses about the preservation of the archives so that this heritage can be handed over to the future generation.
- Digitization requires a huge amount of money. Therefore, they have introduced a new scheme of “adopt one book” for digitization.

**13. New projects under digitization**

**13.1 Treasures with SGPC opened for digital preservation**

PDL continues to extend the legacy of digital preservation by bagging new projects of significance. The latest in the series is the collection of Shiromani Guru Daura Parbandhak Committee (SGPC), Amritsar. The collection includes a huge collection of manuscripts, books, periodicals, newspapers, and photographs concerning Sikhs history, theology, and philosophy.
PDL has been permitted to digitize all the 735 manuscripts, 2217 rare books, and about 10,000 historical photographs and newspapers since 1923 that cover significant historical events including World War II, the Partition of Punjab (1947), the Punjabi Suba Movement, and later events. The project commenced on 1 December 2008, and is currently underway in the Darbar Sahib Gurudwara (the Golden Temple), Amritsar. PDL has deployed nine employees, four digital cameras and a flatbed scanner for the task.

13.2 PDL moves its frontiers to USA

Making its reach global, PDL announces the initiation of a digitization project for the collection of Dr. Gurpal Singh Bhuller, Richmond, Virginia. Dr. Bhuller is an active member of the Sikh community. After moving to the US in 1980, he has been involved in several Sikh organizations and charities. He is a collector of Sikh artefacts and coins, and has authored The Enquiring Guru, published in 2006.

Dr. Bhuller’s collection includes three Guru Granth Sahib Birs (collection of scriptures), fourteen other pothis (books), about 400 Sikh coins and old Temple Tokens, a few paintings, three hukamnamas (prayers) by Guru Tegh Bahadur, one hukamnama (prayer) by Guru Gobind Singh, a few pictures, one illustrated diary, and a few out of print books.

PDL has acquired the equipment in the United States, and plans to start the digitization of the collection in the second week of August. It hopes to finish it before the month ends. PDL welcomes interests in further helping the project in North America and have more people get their collections digitized.

14. Conclusion

PDL is an effort to digitally preserve and reveal the rich heritage of the Punjab state. As time passes, historical treasures have suffered immensely due to human apathy as to the value of their preservation. PDL is not only digitally archiving these items in their original form, colour, graphics, and texture, but it will also change the way the public is able to access and utilize these texts. PDL will redefine the role and scope of a global library. Scholars and the public would have easy access to this online digital library with its powerful searching and browsing capability. This holds great promise for research, education, and awareness, and the conservation of time and financial resources. What the PDL is doing, in a macro sense, is breaking down the traditional barriers between a major library, university or academic institution and the general citizenry.

Special Thanks: Mr Davinder Pal Singh (Executive Director) PDL.
References


http://www.digitallibraries.com
http://www.digitalpreservation.com

Hodge Gail. 2005. Preservation of and permanent access to Electronic Information Resources system Perspective information services & use


Chopra H S. Need of electronic Preservation of Archival Records and Manuscripts: An Indian Scenario

Satyanarayana V V V. Preservation and Archival storage of print forms – digital way

**News**

**e-Granthalaya Software launched**
A Network of Libraries, powered by e-Granthalaya Software, has been launched by National Informatics Centre (NIC), Government of India. NIC is providing e-Granthalaya Software at ZERO cost to the libraries in the country, and now a union catalogue of libraries being made available online, accessible in public domain. Each library using e-Granthalaya has been given a home page at no cost to publish / access their catalogue. The software is available at http://www.eglibnet.gov.in for comments and feedback.

**Government to set up an online Braille library**
Using technology to empower the physically challenged, the National Institute for Visually Handicapped (NIVH), Dehradun, has set up an Online Braille Library to make books and reading material in digital format accessible to visually impaired persons. The Library has books that may be accessed in real time from any location and an amount of Rs 47.50 lakh has been spent on this project so far. An online catalogue of Braille books has also been hosted by NIVH which enables visually-impaired persons to ascertain the location and availability of a desired Braille book, without going to Braille libraries or Press. An amount of Rs 17 lakh has already been spent on this project so far. Besides, NIVH has also launched an Online Directory of Services for the Visually Impaired Persons. Social Justice and Empowerment Ministry has also made its website accessible to persons with disabilities which were launched in January 2010.

*Source: The Times of India, 15 March 2012*

**India to link 9000 libraries**
India will modernize and digitally link nearly 9,000 libraries across the country in a bid to provide readers access to books and information. The move under the National Mission for Libraries aims to cover 9,000 libraries in cities, towns and villages of the country for the next three years.

The Mission will conduct a national census on libraries, work towards upgradation of infrastructure of reading resources, and seek to modernize and promote the networking of libraries across the country. The mission will ensure that a young reader sitting in his village public library would be able to access books and information from across the world. However, the mission can only be successful with the joint efforts of the government, community, private sector, and non-governmental organizations.

The National Mission for Libraries will focus on improvement of the public library system of the country particularly concentrating on the states where library development is lagging behind.

*Source: The Times of India, 21 March 2012*

**Boise Library promotes eBooks with ‘digital’ bookmobile**
The Boise Library at Capitol Blvd, Idaho, USA has a programme that allows library cardholders to browse through, and start reading books without ever leaving their homes. Anyone with a Boise Library card, an Internet connection, and a computer can instantly access thousands of eBooks. Library members can check out eBooks through the library’s website. The eBooks may be browsed...
for seven to 14 days. Once that time has expired, eBooks disappear from the device that they were downloaded on. The Boise Library is adding new eBooks to their collection each day.

Source: www.nwcn.com, 10 May 2012

LexisNexis Digital Library
Lexis Nexis Digital Library is an innovative new service that offers legal professionals access to the largest collection of authoritative legal eBook content on all major mobile devices and desktop platforms. It also enables organizations to share individual eBook titles among multiple users, purchase eBooks centrally and manage their library efficiently. With the service, organizations can significantly reduce the costs associated with storing, filing and distributing traditional print books. LexisNexis Digital Library is customized for each law firm, corporation or organization, with flexible licensing to fit the specific needs of each individual customer.

Source: http://www.marketwatch.com, 23 April 2012

Arabic manuscripts from Jerusalem libraries goes online
Thousands of ancient Arabic-language documents, manuscripts and books from five Arab libraries in Jerusalem’s Old City will soon be available online, under an EU-sponsored initiative that aims to preserve protect Arab heritage, language and Palestinian heritage for future generations. The Arabic Manuscripts Digital Library of Jerusalem, which was inaugurated recently in Jerusalem and Ramallah, will see the development of an online platform showcasing collections from the Khalidi Library, the Budeiri Library, Al-Aksa Library, Al-Ansari Library and the Wafk Restoration Center in the Old City. The project stated goal is to make this heritage accessible to all with the sole caveat of an Internet connection. A corpus of digital documents from various Mediterranean countries has already been created – taking papers and artifacts from libraries, archives and museums, as well as from public and private collections in the region. However, there are no immediate plans to digitalize ancient Hebrew documents from Jerusalem.

Source: Jpost, 26 June 2012

Library and Information Professionals Summit (LIPS) 2012 on “Leadership, Ethics, Accountability and Professionalism in Library Services
16-17 March 2012 at New Delhi, India
A two days Conference on Leadership, Ethics, Accountability and Professionalism in Library Services, organized by Society for Library Professionals in association with Institute of Economic Growth, Asian Chapter-SLA and Ambedkar University, Delhi was organized during March 16-17, 2012 at Institute of Economic Growth, Delhi, India.

About 200 participants ranging from Librarians, Activists, Information Scientists, consultants, publishers, CEOs and entrepreneurs both from academia and industry representing several countries like USA, UK, Sri Lanka etc. Conference was marked by lively invited talks, papers, product presentations, penal discussion, and debate. Experts from across the profession grappled with issues and offered wide ranging opinions and views. Spread over two days, 5 technical sessions, 2 Keynote address, 6 invited talks, nearly 40 paper contribution, product presentations, exhibition and penal discussion conference witnessed plethora of important sessions involving a galaxy of eminent resource persons. The conference was inaugurated by Mr Brent Mai, President, Special Library Association, USA. Ms. Sheila Rosenthal, Manager of Library Services, Carnegie Mellon University, USA, Mr. Sudhir K Arora, University Librarian, IGNOU, Dr Pradeepa Wijetunge, Librarian, University of Peradeniya, Prof J K Mitra, Professor, Faculty of Management Studies, University of Delhi, Dr P R Goswami, Director, Central Secretariat Library, Ministry of Culture, New Delhi, were also present in the conference.
## Article Index

<table>
<thead>
<tr>
<th>Issue</th>
<th>Authors</th>
<th>Title of the Article</th>
<th>Page No</th>
</tr>
</thead>
<tbody>
<tr>
<td>June 2008</td>
<td>Pejtersen AM, Hansen P, and</td>
<td>A work-centred approach to evaluation of a digital library: annotation system for</td>
<td>1 - 17</td>
</tr>
<tr>
<td></td>
<td>Albrechtsen H</td>
<td>knowledge sharing in collaborative work</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Taha A</td>
<td>E-research: a new genre of digital library services</td>
<td>19 - 26</td>
</tr>
<tr>
<td></td>
<td>Comment JM</td>
<td>Digitizing cultural heritage and history archives: Case Studies from Russia (Comintern</td>
<td>27 - 33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Archives) and Albania (National Archives)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poulter A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Balsubramanian A and Jawahar CV</td>
<td>Recognition-free search in graphics stream of PDF</td>
<td>47 - 59</td>
</tr>
<tr>
<td></td>
<td>Sreekumar MG</td>
<td>Open Access: the new frontier connecting the learning commons through hassle-free and</td>
<td>61 - 76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>seamless scholarly communication</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Amrou A, Malay K, and Zubair M</td>
<td>Freelib: a peer-to-peer-based digital library architecture</td>
<td>101 - 120</td>
</tr>
<tr>
<td></td>
<td>Sanyal R, Raj D, Gupta DK, and</td>
<td>DLworm: a system for workflow and repository management for digital libraries</td>
<td>121 - 136</td>
</tr>
<tr>
<td></td>
<td>Verma V</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Seadle M</td>
<td>Copy right and risk: how to judge what to do</td>
<td>137 - 146</td>
</tr>
<tr>
<td>Issue</td>
<td>Authors</td>
<td>Title of the Article</td>
<td>Page No</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>June 2009</td>
<td>Bhatt RK</td>
<td>March towards digitization of information resources in India: issues and initiative</td>
<td>147-164</td>
</tr>
<tr>
<td></td>
<td>Hormia-Poutanen K</td>
<td>Digital library infrastructure in Finland-political decisions empowering development</td>
<td>23-32</td>
</tr>
<tr>
<td></td>
<td>Veltman KH</td>
<td>The Semantic Web: past and future</td>
<td>33-49</td>
</tr>
<tr>
<td></td>
<td>Brown H</td>
<td>Who guards the guards: meeting the challenges of digital preservation</td>
<td>51-68</td>
</tr>
<tr>
<td></td>
<td>Petersen JH and Thorhauge J</td>
<td>Information provision to knowledge creation: Danish digital libraries strategy</td>
<td>69-82</td>
</tr>
<tr>
<td>December 2009</td>
<td>Fagbami OO, Akintola BA, and Pelemo OJ</td>
<td>Cybercafes use by the research scientists in agricultural research institutes in Ibadan, Nigeria</td>
<td>93-100</td>
</tr>
<tr>
<td></td>
<td>Artini M, Candela L, Castelli D, Manghi P, Mikulicic M, Pagano P</td>
<td>Aggregative Digital Library Systems in the DRIVER Infrastructure</td>
<td>113-130</td>
</tr>
<tr>
<td></td>
<td>Poluru L</td>
<td>Building image repository of Indian cultural and heritage materials adopting Dublin Core metadata standards</td>
<td>131-144</td>
</tr>
<tr>
<td></td>
<td>Bhatt RK</td>
<td>E-Resources @ UGC-Infonet Digital Library Consortium: a breakthrough</td>
<td>145-160</td>
</tr>
<tr>
<td>June 2010</td>
<td>Altenhoner R</td>
<td>The German National Library and digital preservation: challenges and opportunities</td>
<td>1-10</td>
</tr>
<tr>
<td>Vol 3 – Issue 1</td>
<td>Zierau E and Kejser UB</td>
<td>Cross-institutional cooperation on a shared bit repository</td>
<td>11-21</td>
</tr>
<tr>
<td></td>
<td>Abrizah A</td>
<td>Piloting an institutional repository at a research-intensive university: strategies for content recruitment and the role of the library</td>
<td>23-39</td>
</tr>
<tr>
<td></td>
<td>Mishra S and Ghosh H</td>
<td>Ontology-based visualization and navigation in an online digital library</td>
<td>41-49</td>
</tr>
<tr>
<td>Issue</td>
<td>Authors</td>
<td>Title of the Article</td>
<td>Page No</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>Misra C, Bhattacharya B, and Basu A</td>
<td>A new framework to preserve Tagore songs</td>
<td>63 - 72</td>
</tr>
<tr>
<td></td>
<td>Smith II PL</td>
<td>Diatomscapes expose: how faculty and digital librarians collaborate to promote and preserve the passion of research (CP3R) for digital future</td>
<td>73 - 84</td>
</tr>
<tr>
<td></td>
<td>Falconer C</td>
<td>Project Naming: what makes it a “good digital collection”?</td>
<td>97 - 100</td>
</tr>
<tr>
<td></td>
<td>Sharma M and Urs SR</td>
<td>Editorial Board of Digital Library Journals: a social network analysis approach</td>
<td>101 - 114</td>
</tr>
<tr>
<td></td>
<td>Rafiq M and Ameen K</td>
<td>Perception of South Asian LIS community towards open source software adoption in libraries</td>
<td>115 - 124</td>
</tr>
<tr>
<td></td>
<td>Chawla R and Chawla P</td>
<td>Agent-based web intelligent tutoring system</td>
<td>125 -132</td>
</tr>
<tr>
<td></td>
<td>Luan F, Naygard M, and Mestl T</td>
<td>A survey of digital preservation strategies</td>
<td>133 - 150</td>
</tr>
<tr>
<td></td>
<td>Chaudhuri B, He L, and Juterbock D</td>
<td>Global e-library services in Novartis Knowledge Center: meeting worldwide information needs at Novartis</td>
<td>151 - 158</td>
</tr>
<tr>
<td></td>
<td>Gupta RK and Lohiya SK</td>
<td>Free/open source software for library and information management</td>
<td>159 - 168</td>
</tr>
<tr>
<td></td>
<td>Bhatt RK</td>
<td>Imperativeness of e-learning in a digital environment in an organization</td>
<td>169 - 174</td>
</tr>
<tr>
<td>June 2011 Vol 4– Issue 1</td>
<td>Doctor G</td>
<td>A simulation study for a number of simultaneous users of an institutional knowledge repository system</td>
<td>1- 8</td>
</tr>
<tr>
<td></td>
<td>Chowdhury G Prof. and Fraser M Prof.</td>
<td>Carbon footprint of the knowledge industry and ways to reduce it</td>
<td>9 - 18</td>
</tr>
<tr>
<td></td>
<td>Rathour PR and Sahu AK</td>
<td>Comparison of open source software for digital libraries</td>
<td>19 - 31</td>
</tr>
<tr>
<td></td>
<td>Md Nasiruddin</td>
<td>Vision 2021 and public libraries: an action plan to allow Bangladesh to go digital</td>
<td>33 - 47</td>
</tr>
<tr>
<td>Authors</td>
<td>Title of the Article</td>
<td>Page No</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------------------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Maceviciute E</td>
<td>Education for digital libraries: library management perspective</td>
<td>49 - 61</td>
<td></td>
</tr>
<tr>
<td>Comment JM</td>
<td>The digital repository of the Swiss Federal Archives: a case study</td>
<td>63 - 68</td>
<td></td>
</tr>
<tr>
<td>Carstens C, Rittberger M, and Wissel V</td>
<td>Information search behavior in the German Education Index</td>
<td>69 - 80</td>
<td></td>
</tr>
<tr>
<td>Kumar GHH, Srinivas V, Raju BNY, and Sudhakara S</td>
<td>India’s contribution to Ranking Web of World Repositories: A case study</td>
<td>85 - 114</td>
<td></td>
</tr>
<tr>
<td>Veeramani M and Vinayagamoorthy P</td>
<td>Analysis of Information Technology: Application in Academic Libraries in Kuwait</td>
<td>115 - 124</td>
<td></td>
</tr>
<tr>
<td>Md Rocknuzzman, Md Hasan N, and Akanda SKMEA</td>
<td>A framework for Knowledge Management Education in Digital Library Learning</td>
<td>125 - 135</td>
<td></td>
</tr>
<tr>
<td>Nieuwenhuysen P</td>
<td>Improving Information Discovery in Digital Libraries</td>
<td>137 - 151</td>
<td></td>
</tr>
<tr>
<td>Chowdhury S</td>
<td>Digital Natives and Virtual Learning Environment: A Case Study</td>
<td>153 - 163</td>
<td></td>
</tr>
</tbody>
</table>